Earth Science Informatics (2025) 18:348
https://doi.org/10.1007/s12145-025-01860-7

RESEARCH

Check for updates

Machine learning assessment of illegal mining (Galamsey) impacts on forest vegetation: a case study of Wassa Amenfi East District, Ghana

Clement Kwang[1] · Ian Kofi Afele[1] · Emmanuel Yeboah[2] · Isaac Sarfo[3,12] · Michael Batame[4] · Abraham Okrah[5] · Myint Myint Shwe[6] · Williams Siaw[7] · Dinah Boyetey[8] · Richard Odoi Larbi[9] · Augustine O. K. N. Mensah[10] · Charafa El Rhadiouini[2] · Ali Hasan Jaffry[2] · Fareeha Siddique[2] · Rukhshinda Aftab[2] · Oznur Isinkaralar[11]

Received: 1 January 2025 / Accepted: 20 March 2025 / Published online: 28 March 2025

Abstract
This study addresses the persistent land management challenge of galamsey (illegal mining) in Wassa Amenfi East by exploring the impact of these activities on vegetation through advanced machine learning techniques. A comparative analysis was conducted using four machine learning algorithms Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and Maximum Likelihood Classification (MLC) to assess their effectiveness in detecting and analyzing vegetation changes due to galamsey operations. The results highlight the Random Forest (RF) algorithm as the most effective, with overall accuracy scores of 88% for 2015 and 87% for 2023, and Kappa Coefficient values of 0.84 and 0.82, respectively, demonstrating its consistent superiority over the other methods. Findings of the study reveal significant vegetation and forest cover loss due to galamsey, driven by poverty, unemployment, poor policies, livelihood pursuits, and quick financial gains by traditional authorities. This study stresses the need for targeted interventions to mitigate the environmental impact of galamsey and suggests the adoption of advanced machine learning techniques for more accurate and effective land management strategies.

Keywords Galamsey · Machine learning techniques · Support vector machine · Artificial neural network · Maximum likelihood classification · Random

Introduction

Mining serves as a crucial source of mineral resources for numerous economies worldwide (Osman et al. 2022). It plays a vital role in driving economic development and enhancing the gross domestic product (GDP) of countries on a global scale (Ahmad et al. 2020; Ericsson and Löf 2019). Ghana, like many African nations, relies heavily on the mining sector, particularly gold, with a notable rise in small-scale mining operations (Kwang & Blagogie 2024; Aryee et al. 2003). However, mining activities pose significant environmental challenges, including the loss and degradation of vegetation (Obeng et al. 2019; Afele et al. 2022). In Ghana, the adverse impacts of mining, particularlygalamsey, have directly affected its vegetation cover and food security (Obodai et al. 2024). Galamsey is a major contributor to vegetation cover depletion in Ghana, causing persistent environmental and social impacts over decades (Afriyie et al. 2023; Kyere-Boateng and Marek 2021). The lack of formal registration or licensing of galamsey activities complicates efforts to track and monitor these operations (Owusu-Nimo et al. 2018). Galamsey operators employ indiscriminate and unscientific mining methods, lacking post-mining treatment and management of mined areas, rendering fragile ecosystems more vulnerable to environmental degradation (Kumi-Boateng et al. 2012). Between 2001 and 2021, deforestation in Ghana increase of approximately 5% (Afele et al. 2021; Biney et al. 2022). Studies, including those by Baddianaah et al. (2023), highlighted the detrimental effects of galamsey, including environmental degradation, water resource depletion, compromised air quality, destruction of ancestral homes, and significant landscape modifications. Similarly, Ocansey (2013) and Osman et al. (2022) documented the substantial harmful effects of illegal mining activities on the physical environment, resulting in land degradation, deforestation, water, and air pollution.

Communicated by: Hassan Babaie

Extended author information available on the last page of the article

The Wassa Amenfi East district, situated in Ghana's Western Region, is renowned for its abundant gold deposits. Unfortunately, in recent years, the area has experienced a surge in galamsey activities, emerging as one of the major hotspots in the country, with 223 reported sightings (Owusu-Nimo et al. 2018). This increase in galamsey operations has taken a severe toll on the local ecosystem, leading to detrimental environmental effects such as soil erosion, deforestation, water pollution, and a decline in wildlife populations, which have collectively caused unexpected and undesirable consequences (Taylor 2024). Illegal mining (galamsey) has led to significant environmental degradation, threatening sustainable development and community well-being in the Wassa Amenfi East District. Traditional monitoring approaches often lack the scale, accuracy, and timeliness required to effectively assess these impacts. While machine learning techniques offer a promising alternative, most existing research focuses on single algorithms rather than comparative evaluations. This limitation restricts the understanding of how different machine learning models perform in detecting and assessing galamsey-induced vegetation loss. Without a systematic comparison, it remains unclear which algorithm provides the most accurate and reliable results for environmental monitoring in affected regions.

To address this gap, this study aims to conduct a comprehensive comparative analysis of four machine learning algorithms Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and Maximum Likelihood Classification (MLC). By evaluating their effectiveness in mapping and assessing vegetation loss due to galamsey activities, the study seeks to determine the most suitable approach for monitoring illegal mining impacts. Additionally, the analysis integrates vegetation indices and other environmental variables to enhance model accuracy and provide a more detailed assessment of land cover changes. By identifying the most effective machine learning approach, this study contributes to advancing remote sensing applications in environmental monitoring. Recent advancements in machine learning within remote sensing have opened new possibilities for monitoring and controlling environmental concerns like the impact of galamsey activities. Machine learning techniques offer numerous advantages, such as the efficient handling of large datasets, automation of the analysis process, and the ability to derive valuable insights from complex imagery (Maxwell et al. 2018). These methods enhance accuracy and consistency, reducing the time and effort required compared to traditional manual interpretation methods (Florian, 2022). Furthermore, machine learning enables the integration of data from various sensors and remote sensing platforms (Salcedo-Sanz et al. 2020), offering a more comprehensive approach to environmental monitoring.

The objective of this study is to conduct a comprehensive comparative analysis to evaluate the performance of different machine learning techniques in assessing the impact of galamsey on vegetation in the Wassa Amenfi East District. This study not only broadens the methodological repertoire but also provides a more reliable comparison of results by utilizing four different machine learning algorithms. It offers insights into the specific impacts of galamsey on vegetation, contributing to the development of more effective and efficient approaches for monitoring and managing the environmental impact of galamsey activities, ultimately aiding in the preservation and restoration of vegetation ecosystems. The study conceptualizes the problem mathematically as k (SVM, RF, ANN, MLC), where "k" denotes the kappa values for the classifiers. Each model's performance was evaluated using the kappa coefficient (k), where k = (SVM, RF, ANN, MLC). This equation indicates that each machine learning technique is evaluated individually using its kappa value to determine its effectiveness in mapping the impact of galamsey activities on vegetation. The four algorithms; SVM, RF, ANN, and MLC are employed for their proven efficacy in handling complex environmental data. These algorithms offer flexibility, robustness, and the ability to process high-dimensional remote sensing data, making them suitable for accurately identifying the effects of galamsey activities on vegetation cover. Incorporating standardized theories related to spatial sustainability, this study aligns with concepts such as sustainable land management (SLM) and the theory of ecological modernization, which advocate for integrating technological advancements with environmental governance to achieve sustainability goals. The study also aligns with several Sustainable Development Goals (SDGs), particularly SDG 15 (Life on Land), which focuses on protecting, restoring, and promoting sustainable use of terrestrial ecosystems, and SDG 13 (Climate Action), which emphasizes the need for urgent action to combat climate change and its impacts. By addressing the environmental degradation caused by galamsey through machine learning techniques, this study contributes to achieving these SDGs, promoting more sustainable land management practices.

Study area and methodology

Study area

The Wassa Amenfi East District, located in Ghana's High Forest Zone within the Western Region, is a prominent area known for its rich natural resources and diverse ecosystems. The district's capital town is Wassa Akropong, and geographically, it spans between latitudes 5°30′N and 6°15′N, and longitudes 1°45′W and 2°11′W (Onumah 2010). Covering an area of approximately 1,729 km^2, Wassa Amenfi

East is one of the fourteen districts in the Western Region of Ghana. The district is bordered by Wassa Amenfi West to the west, Mpohor Wassa East to the east, Upper and Lower Denkyira to the north, and Wassa West District to the south (Onumah 2010). According to the Ghana Statistical Service (2021), the district has a population of 179,696, making it a moderately populated area with a significant dependence on both natural resources and agriculture for livelihood.

The district is characterized by two main vegetation types: Semi-deciduous forest to the north and tropical rainforest to the south both vital for biodiversity conservation and maintaining the ecological balance. The district experiences a bimodal rainfall pattern, one from April to July and the other from September to December (Onumah 2010). This climatic pattern supports the cultivation of various crops, particularly cocoa, which is the main agricultural activity and a significant source of income for the local population (Ezeji Onyebuchi 2014). Wassa Amenfi East is mineral-rich, with substantial deposits of gold and other valuable minerals driving legal and galamsey activities. While these mining activities contribute to the local economy, it has led to significant environmental challenges, including deforestation, soil erosion, and water pollution. The district's unique combination of rich biodiversity, significant agricultural productivity, and mineral resources makes it a critical case study area for understanding the environmental impacts of resource exploitation in Ghana Fig. 1.

The summary of the methodology used for the study is indicated in Figs. 2, 3, 4, 5 and 6. The Figure provides chronological procedures and data used to achieve the results and findings of the study.

Data and pre-processing

The Landsat 8 Operational Land Imager (OLI) satellite imagery of 2015 and 2023 were downloaded from the United Geological Survey (USGS) Earth Explorer website. Landsat 8 was chosen for this study due to its free availability. Despite its moderate spatial resolution, it effectively supported the detection and mapping of galamsey-affected areas

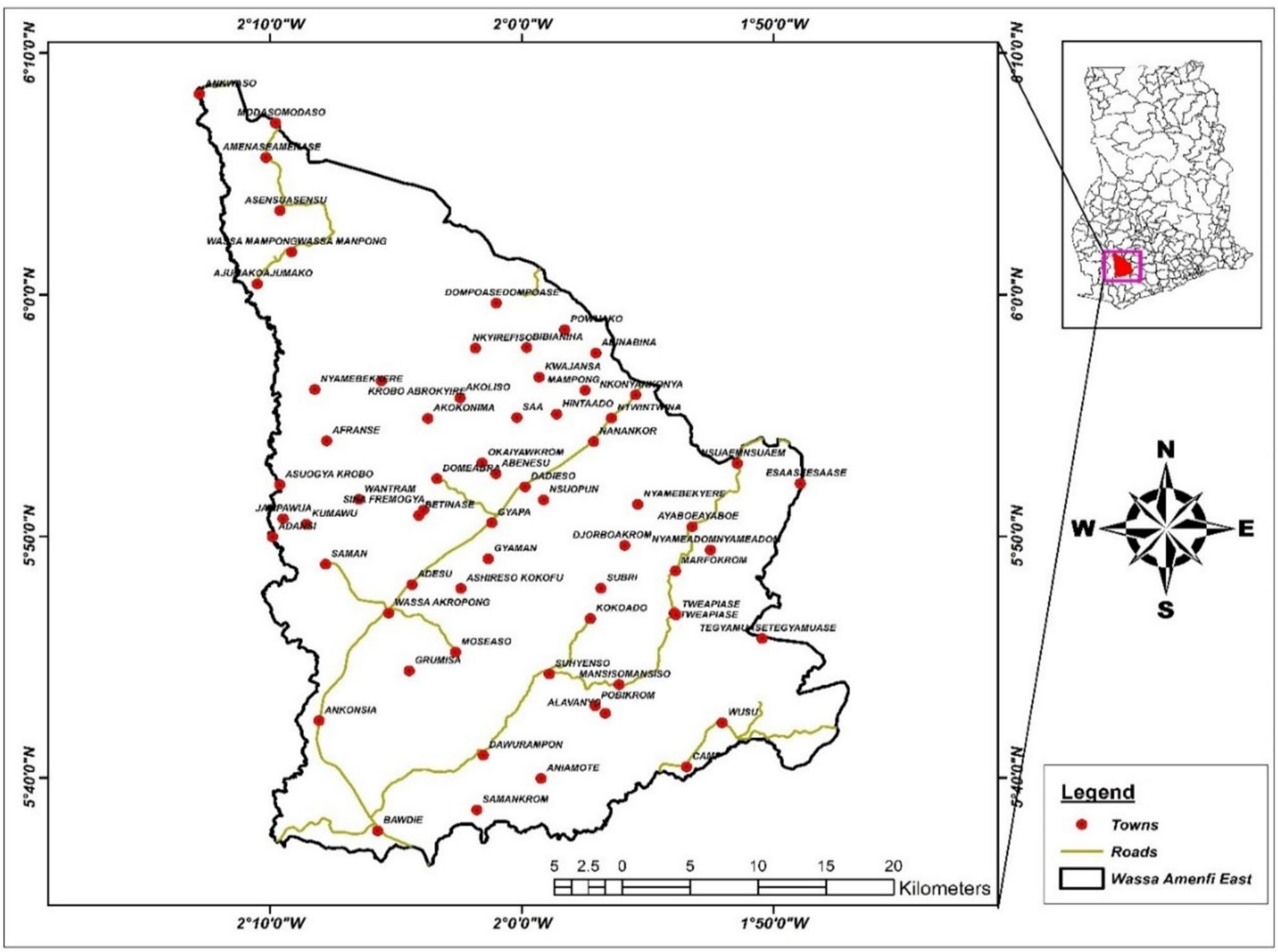

Fig. 1 Study Area map of Wassa Amenfi East District

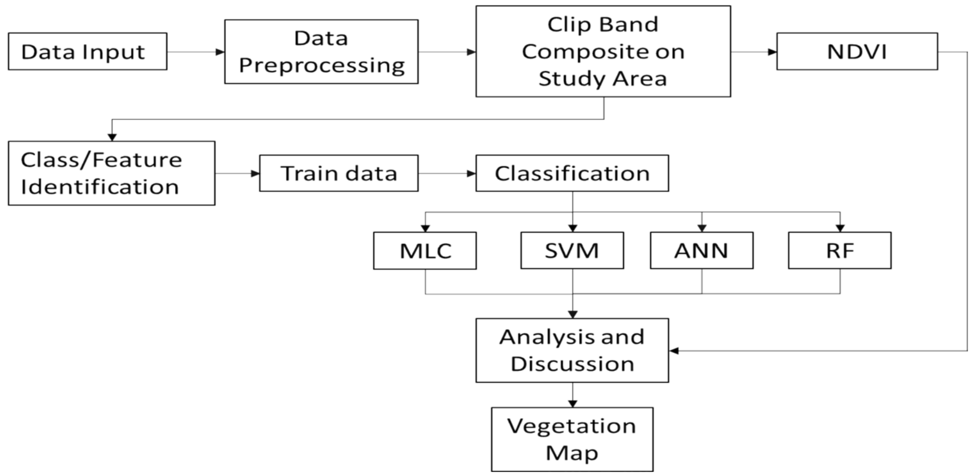

Fig. 2 Workflow of the methodology

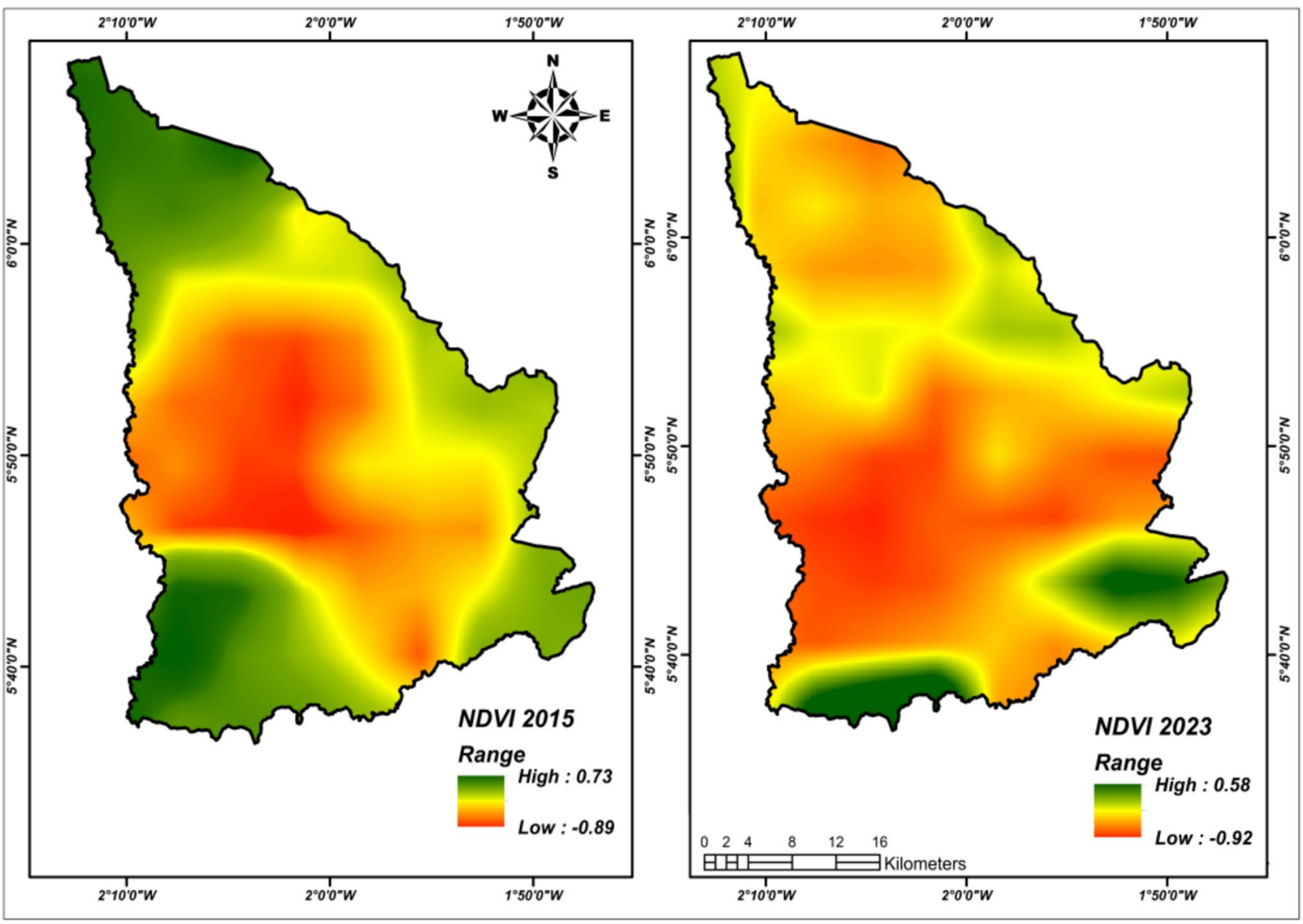

Fig. 3 NDVI results showing spatial variations of vegetation health for the years 2015 and 2023

within the vegetated zone of the study area Tables 1 and 2 and 3.

The images were pre-processed using ArcMap 10.8 and ENVI 5.3 to correct atmospheric and radiometric distortions, improving brightness, contrast, and colour balance. Bands 2, 3, 4, 5, 6, and 7, representing red, blue, green, near-infrared, and shortwave infrared wavelengths, were extracted to ensure consistent spatial resolution. These bands were used for detailed image analysis. Subsequently, the band composite images were clipped to the district's boundary using the

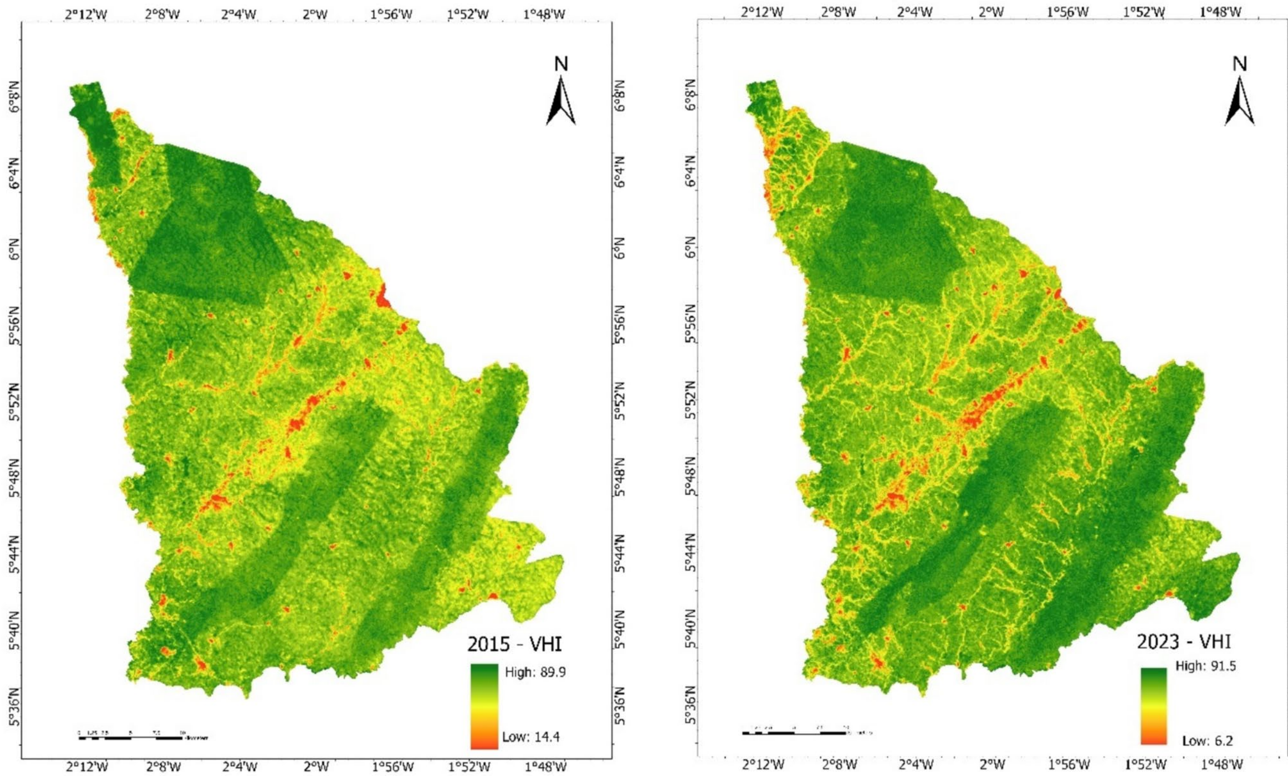

Fig. 4 VHI results showing health and stress levels of vegetation cover for the years 2015 and 2023

district's shapefile to focus the analysis on the relevant area. This pre-processing and clipping ensured that the imagery was accurately aligned and tailored for assessing the impact of galamsey on vegetation within the specified district. The Landsat 8 images underwent rigorous preprocessing to mitigate atmospheric and radiometric distortions. Radiometric calibration was first applied by converting digital number (DN) values to spectral radiance using the USGS-provided scaling factors (L = ML * Qcal + AL, where ML = 0.0003342 and AL = 0.1 for Band 2 – 7). Radiance was then converted to top-of-atmosphere (TOA) reflectance using the solar irradiance and solar zenith angle values from the metadata. Atmospheric correction was performed using ENVI 5.3's FLAASH module with the following parameters: Tropical atmospheric model (selected for the study area's equatorial latitude), Landsat-8 OLI sensor type, 30-meter pixel resolution (aligned with the native spatial resolution), and acquisition time/date derived from scene metadata. Additional FLAASH settings included the Mid-Latitude Summer aerosol model (default for tropical regions), 2-Band (K-T) aerosol retrieval method (optimized for Landsat's visible bands), and moderate spectral polishing to reduce residual noise. Post-correction quality control involved validating reflectance values against pseudo-invariant features (e.g., deep-water bodies), statistical outlier analysis (mean reflectance $\pm 3\sigma$ thresholds), and visual inspection for artifacts using ENVI's spectral profile tool. These steps ensure enhanced radiometric consistency, minimized haze contamination, and improved spectral separability for subsequent classification. Bands 2, 3, 4, 5, 6, and 7 representing red, blue, green, near-infrared, and shortwave infrared wavelengths were extracted to ensure consistent spatial resolution for detailed image analysis. These bands were then mosaiced to create a composite band of the satellite imagery. The study area shapefile was overlaid on the composite band and subsetted or clipped using ArcMap 10.8 to obtain the focus area for further analysis. This preprocessing ensured accurate alignment of the imagery, making it suitable for assessing the impact of galamsey on vegetation within the specified district.

Normalized difference vegetation index

Bands 5 (near-infrared) and 4 (red) were used to calculate the Normalized Difference Vegetation Index (NDVI). NDVI is a widely used vegetation index that helps in identifying areas of stress or damage and monitoring the growth and health of vegetation (Jabbar et al. 2021). By comparing the near-infrared and red bands, NDVI effectively highlights changes in vegetation cover over time, making it valuable for detecting alterations due to factors such as illegal small-scale mining (galamsey).

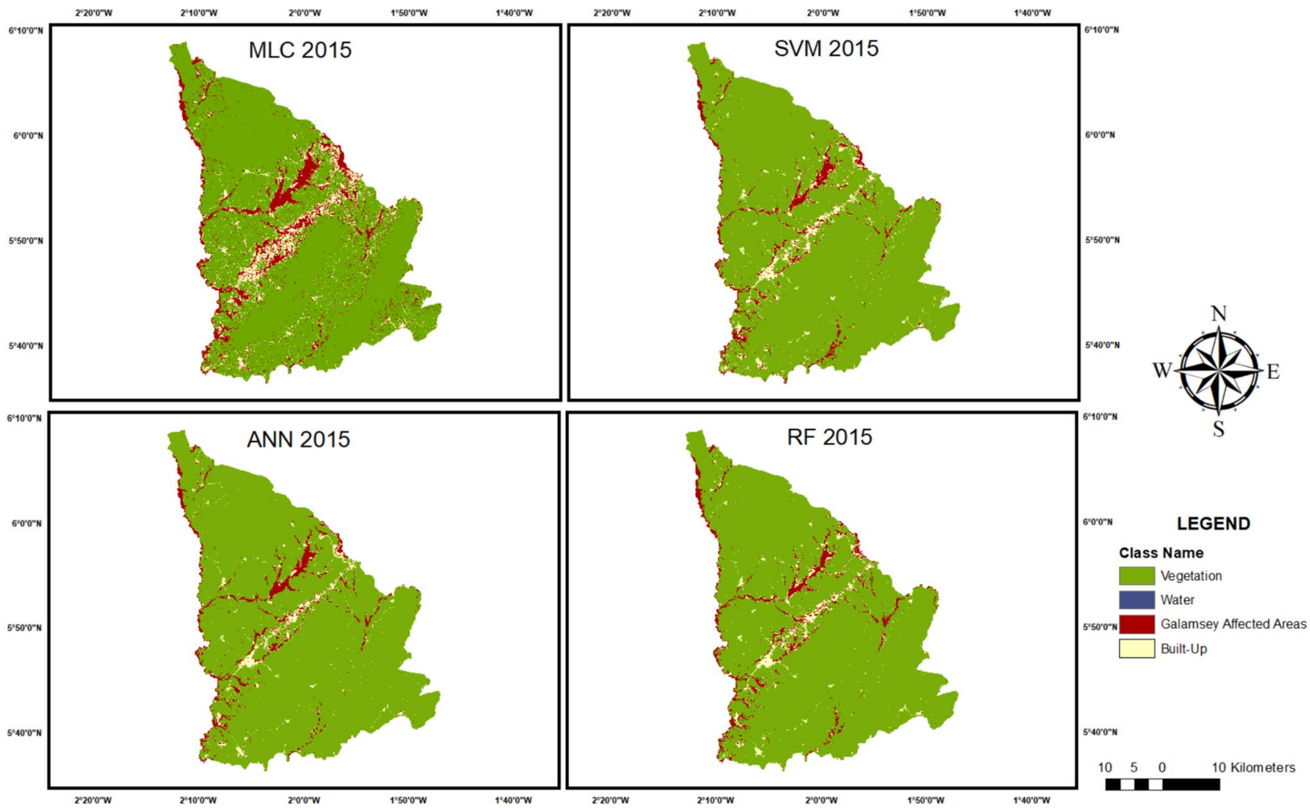

Fig. 5 2015 Land Use Classification output of the models (MCL, SVM, ANN and RF)

Additionally, NDVI aids in mapping and classifying different types of vegetation, providing insights into vegetation dynamics and land cover changes. This index is instrumental for environmental monitoring and management, as it allows for the assessment of vegetation health and the impact of anthropogenic activities on ecosystems (Jabbar and Yusoff 2022; Isinkaralar et al. 2025). NDVI is expressed mathematically as follows:

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{RED})}{(\mathrm{NIR} + \mathrm{RED})} \tag{1}$$

where RED is visible red reflectance, and NIR is near-infrared reflectance. The result ranges from −0 to 1.

Vegetation health index

The vegetation health index integrates vegetation condition index (VCI) and temperature condition index (TCI). The VCI and TCI are derived from NDVI and Land surface temperature (LST) respectively. The equations for calculating VCI, TCI and VHI are indicated as follows:

$$\mathrm{VCI} = \frac{(\mathrm{NDVI} - \mathrm{NDVI}_{\min})}{\left(\mathrm{NDVI}_{\max} - \mathrm{NDVI}_{\min}\right)} \times 100 \tag{2}$$

where NDVI , $\mathrm{NDVI}_{\max}$ and $\mathrm{NDVI}_{\min}$ are the smoothed weekly NDVI, multi-year maximum NDVI and multi-year minimum NDVI, respectively for each grid cell (Singh et al. 2003). VCI values ranges from 0 to 100, with higher values indicating healthier vegetation and less drought stress.

$$\mathrm{TCI} = \frac{(\mathrm{LST}_{\max} - \mathrm{LST})}{\left(\mathrm{LST}_{\max} - \mathrm{LST}_{\min}\right)} \times 100 \tag{3}$$

where LST, $\mathrm{LST}_{\max}$ and $\mathrm{LST}_{\min}$ are the smoothed weekly brightness temperature, multi-year maximum and multi-year minimum, respectively, for each grid cell (Singh et al. 2003). TCI values range from 0 to 100, with higher TCI values indicate better thermal conditions for vegetation

$$\mathrm{VHI} = \alpha\ \mathrm{VCI} + \alpha\ \mathrm{TCI} \tag{4}$$

where α is 0.5. VHI values range from 0 to 100, with higher values indicating healthier vegetation

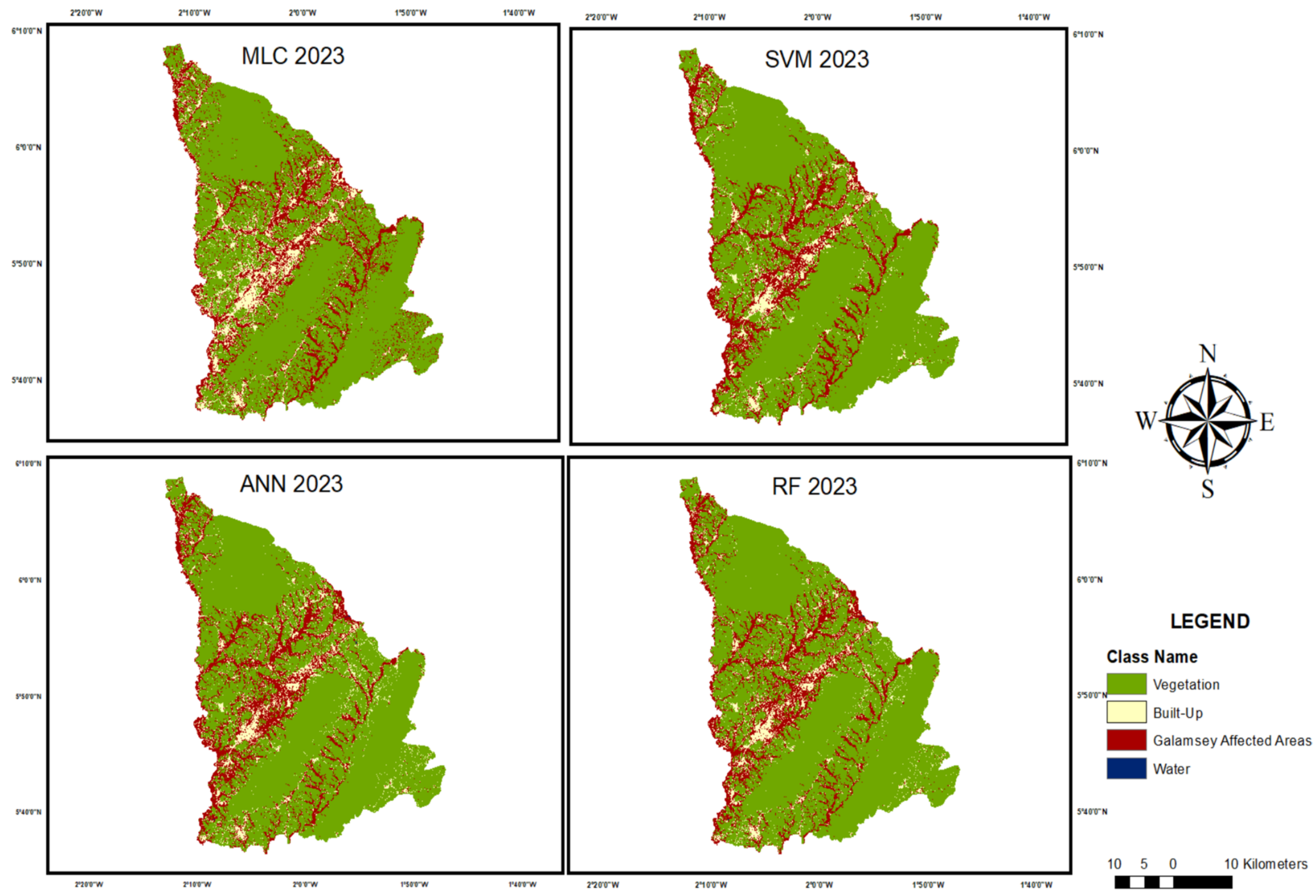

Fig. 6 2023 Land Use Classification output of the models (MCL, SVM, ANN and RF)

Table 1 dataset used in the study

Remote Sensing Data	Date Acquired	Resolution	Source
Landsat 8 OLI	29-01-2015	30 m	USGS
Landsat 8 OLI	01-01-2023	30 m	USGS

Land cover classification

The images were classified into four distinct classes: Vegetation, Built-Up, Galamsey Affected Areas, and Water. This classification was achieved by defining training areas for each class. In the study, fifty regions of interest (ROIs) were established for each classified image: fifteen for vegetation, six for built-up areas, four for water, and fifteen for galamsey affected areas. The spectral signatures from these ROIs were used to inform and improve the classification process. These ROIs provided essential data that allowed for accurate identification and differentiation of various land cover types in the classified images. The Maximum Likelihood Classification (MLC), Support Vector Machine (SVM), and Random Forest (RF) classifiers were applied using ArcMap 10.8, while the Artificial Neural Network (ANN) classifier was executed using ENVI 5.3.

Table 2 Description of land cover types identified in the study area

Land Cover	Description
Vegetation	Areas is characterized by dense clusters of trees and a lush, continuous vegetative cover, with minimal or no exposed bare soil. It also includes regions with more scattered vegetation, where trees are sparsely distributed, alongside shrubs, isolated thickets, and areas cultivated with non-tree crops.
Built-up	Residential, commercial, and industrial zones. This category also encompasses parks, gardens, playgrounds, and lorry stations located within communities.
Galamsey Affected Areas	These are areas where illegal mining activities have taken place and caused significant environmental impact.
Water	Areas include waterlogged regions, water pits, rivers, lagoons, lakes, and other similar bodies of water. .

Table 3 Temporal Changes of Classes in hectare (ha) between 2015 and 2023

LULC Type	Area (ha) 2015	Area (ha) 2023	Change (%)	Key Drivers
Vegetation	140,532	12,2343	−12.94	Galamsey encroachment
Water	7.99311	2.88844	−63.86	Mining pollution
Galamsey Affected Areas	9,381.80	27,061.9	+188.45	Galamsey expansion
Built-Up	5,820.24	6,333.16	+8.81	Urbanization, Population growth

Maximum likelihood classification

Maximum Likelihood Classification (MLC) is a conventional and widely used supervised ML technique in remote sensing (Ha et al. 2020). In MLC, each pixel is assigned to a class based on the probability of its membership to that class. This probability is determined by modelling the class distributions as normal distributions in the multispectral feature space, using the mean and covariance of each class (Watuwaya et al. 2023). MLC leverages probability theory to perform image classification, making it effective for distinguishing between different land cover types based on the statistical properties of the pixel values (Domadia & Zaveri 2011). Therefore, the equation for MLC is based on Bayesian probability theory and has been expressed as;

$$D = \ln(a_c) - \left[0.5\ln\left(\left|C_{ovc}\right|\right)\right] - \left[0.5(X - M_c)^T(C_{ovc}^{-1})(X - M_c)\right] \tag{5}$$

where D is likelihood, c represent a particular class, X indicate the measurement vector of the candidate pixel, M_c is the mean vector of the sample of class c, a_c is the percent probability that any candidate pixel is a member of class c, C_{ovc} is the covariance matrix of the pixel in the class c, $\left|C_{ovc}\right|$ represents the determinant of C_{ovc} and C_{ovc}^{-1} is the inverse of C_{ovc}

Although it is the most go-to technique, the process requires significant computational resources and is time-consuming (Watuwaya et al. 2023). A study by Sisodia et al. (2014) analysed the performance of maximum likelihood classification in image classification. The results showed a high overall accuracy of 93.75% with a kappa coefficient of 0.90 showing a high level of agreement between classified results and the actual classes. The study concluded that the technique had a low chance of misclassification. In a comparative study by Patil et al. (2012) where maximum likelihood classification and minimum distance method were both used, maximum likelihood classification scored the highest with an overall accuracy of 91% as against the minimum distance method which scored 81%.

Support vector machines

Support Vector Machines are powerful and robust classification and regression algorithms widely used in several scientific and engineering areas (Sachin 2022). The SVM method's main goal is to create a hyperplane in decision boundary space that represents the best separation of linearly separable classes (Mohan et al. 2020). According to Cervantes et al. (2020), the strengths of SVMs include their ability to obtain a subset of support vectors during the learning phase, which represents a given classification task and is formed by a small data set. Additionally, SVMs have good theoretical foundations and generalization capacity, making them one of the most used classification methods. Despite their advantages, SVMs have some limitations, including the selection of parameters, algorithmic complexity that affects the training time of the classifier in large data sets, and the development of optimal classifiers for multi-class problems and the performance of SVMs in unbalanced data sets (Cervantes et al. 2020).

The support vector machine algorithm is based on a simple linear equation model expressed as by Suthaharan, (2016):

$$y = wx' + \gamma \tag{6}$$

where y is the decision function that classifies the input image into a class label, w is the weight vector, x' is the feature vector representing the input image and γ is the bias term.

The training samples of support vector machine involve setting the parameters w and γ in Eq. 1 and orienting the hyperplane to find the optimal hyperplane that separates the data points of different classes in a high-dimensional space (Mahmoudi et al. 2012).

To classify a new image using the support vector machine algorithm, the decision function is first calculated using Eq. 2 and the sign of its result is then determined using the following expressions:

$$y \geq 0$$
$$y \leq 0$$

If $y \geq 0$, then the image is classified as belonging to one class and if $y \leq 0$, then the image is classified as belonging to another class.

Tariq and Mumtaz (2023) presented a comparative analysis of three ML algorithms, namely, Support Vector Machine

(SVM), Kernel Logistic Regression (KLR), and Naive Bayes Tree (NBT) models, for mapping and monitoring forest cover changes using optical remote sensing data. The paper found that the SVM algorithm performed better than the other two algorithms, with an overall accuracy of 82%. The study concluded by stating that the SVM algorithm is a reliable and effective method for mapping and monitoring forest cover changes using optical remote sensing data. Taati et al (2015) conducted a study aimed at creating land-use classification using the support vector machine (SVM) and maximum likelihood classifier (MLC) in Qazvin, Iran, by TM images of the Landsat 5 satellite. The results showed that the SVM algorithm with an overall accuracy of 86.67% and a kappa coefficient of 0.82 has a higher accuracy than the MLC algorithm in land-use mapping. Therefore, this algorithm has been suggested to be applied as an optimal classifier for the extraction of land-use maps due to its higher accuracy and better consistency within the study area.

Artificial neural network

For years, image classification methods used in a variety of industries have benefited from the usage of artificial neural networks (Mahmon, & Ya'acob 2014). It makes use of neural networks, which are computer models inspired by the organic neural networks of the human brain (Cichy & Kaiser 2019). ANN classification makes predictions or assigns class labels to various spatial locations or objects by using the spectral properties of data (Wang et al. 2022).

Normally the training of the network is done using the Back propagation algorithm and its procedure includes these steps (Domadia & Zaveri 2011). Training pixels are taken from each class in the image and its desired output and assigned d. The input vector of the pixel is assigned X. Initialization of weights to small random values $W_{ih}W_{hj}$. Feed input vector X_1, X_2..........X_n and compute the weighting sum coming into node, apply the sigmoid equation

$$S_h = (W_{ij})^T X \tag{7}$$

where the weights between input nodes and hidden nodes = W_{ij} and the weighting sum into hidden node $=S_h$. The probability of each hidden node is determined using

$$Y_h = \frac{1}{1+e^{-s_h}} \tag{8}$$

and

$$Y_j = \frac{1}{1+e^{-s_j}} \tag{9}$$

The weighting sum coming into the output node is calculated as the probability of each hidden node using the equation;

$$S_j = (W_{ij})^T Y_h \tag{10}$$

The error term for each output unit is estimated using

$$\delta_j = Y_j\left(1 - Y_j\right)\left(d_j - Y_j\right) \tag{11}$$

Calculate the mean square error (MSE) of output node $e = \frac{\sum_{i=1}^{j}(d_j - Y_j)^2}{2}$. Calculate the error term of each of the hidden nodes using

$$\delta_h = \left(W_{hj} - \delta_h\right) Y_h \left(1 - Y_h\right) \tag{12}$$

Adjust weights to minimize mean square error using the following equations

$$W_{ih} = W_{ih} + \alpha\, X\delta_h + \beta\left(W_{ih} - W_{ih1}\right) \tag{13}$$

and

$$W_{hj} = W_{hj} + \alpha\, Y_h\delta_j + \beta\left(W_{hj} - W_{hj1}\right) \tag{14}$$

Repeat steps 7–12 till MSE falls within reasonable limits. Find Y_h and Y_j for each pixel using W_{hj} and W_{ih} after training neural networks with training pixels. The pixel is assigned to Y_j class if Y_j have a maximum probability for this pixel. All the other pixels of image are classified per this assertion.

ANN classification in remote sensing has been used for several tasks, including the categorization of land cover, the mapping of vegetation, and the extraction of urban features just to mention a few (Pu et al. 2011). It has the benefit of collecting intricate spatial correlations and patterns in the data, making it possible to classify geographic elements more precisely and automatically (Wang et al. 2016), however, as identified by (Ray 2019)

ANNs have some drawbacks, including extensive training requirements, expensive computing costs, and weight adjustments. A paper by Mahmon and Ya'acob (2014) investigated the capability and accuracy of using ANN in classifying satellite imagery. The paper also compared ANN to the conventional supervised classification technique MLC and the unsupervised classification in ISODATA. The result of the study showed a high overall accuracy score of 89. % and Kappa Coefficient of 0.820. The paper concluded by stating that ANN is a very useful and accurate approach to image classification in remote sensing. Similarly, a paper by Mustapha et al. (2007) compared ANN and MLC in land cover mapping. The result of the accuracy assessment showed that ANN scored higher in comparison to the MLC technique. In terms of overall accuracy, ANN scored 89.3% and Kappa Coefficient of 0.820. On the other hand, MLC scored an overall accuracy of 80.3% and a Kappa Coefficient of 0.672.

Random forest

A flexible and effective classification technique that has been utilized with success in remote sensing for a variety of applications is the random forest classifier (Sheykhmousa et al. 2020). Using a randomly chosen subset of training samples and variables, the RF classifier is an ensemble classifier that generates several decision trees (Shaik & Srinivasan 2019). According to Shaik and Srinivasan (2019), the predictions from an ensemble of decision trees used by the RF classifier produce accurate classifications. Belgiu and Drăguţ (2016) identified some advantages of using random forest classification including high and excellent classification results, fast processing speed, stability, and finally in comparison to other ML techniques, it is less sensitive to the quality of training data samples. Random forest classification also has its shortcomings and limitations. Random forest classification although known for its high classification accuracy, it, however, lacks in the department of interpretability (Biau & Scornet 2016). A recent study by Giri et al. (2024) further demonstrated the adaptability of Random Forest Regression (RFR) when combined with advanced signal processing techniques. In their work analyzing soybean seed composition using hyperspectral data, canopy-level predictions improved dramatically after applying Mexican Hat wavelet transformation - R^2 values increased from 0.07 to 0.44 for protein content and from 0.02 to 0.39 for oil content when using RFR. This significant enhancement highlights Random Forest's ability to leverage transformed spectral features effectively, even when dealing with complex canopy-level data affected by environmental variables and structural noise. The results align with previous RF classification successes while extending its utility to regression tasks requiring sophisticated data preprocessing. In addition to that, the complexity and processing time of random forests increase as the number of trees in the forest rises. Due to lengthier training times, it may be less effective when dealing with huge datasets (Shaik & Srinivasan 2019).

In all as compared to other ML algorithms, RF is a versatile and powerful classifier that balances performance, robustness, and ease of use, making it a practical choice for many machine learning applications.

A study by Tian et al. (2016) employed random forest technique to classify wetland land cover in China. The study suggested that random forest performs preferably well when compared to other techniques particularly support vector machine and artificial neural network. With an overall accuracy of 93% and Kappa coefficient of 0.92, the random forest classifier performed well in classifying the wetland. Rodriguez-Galiano et al. (2012) investigated the performance of random forest classifier for land cover classification in the south of Spain. The results demonstrate that the random forest classifier achieved accurate land cover classifications, with an overall accuracy of 92% and a Kappa coefficient of 0.92.

Application of ML in Galamsey

The use of ML techniques in remote sensing data analysis provides an effective approach to identifying and mapping areas affected by galamsey and assessing their impact on natural resources. Opoku et. al, (2017) employed a combination of unsupervised and supervised classification methods to monitor land use and land cover changes in the mining district of Prestea Huni Valley in Ghana. The study found a significant loss of forest cover as a result of direct mining activities, with an overall forest cover loss rate of about 71.63 square kilometres per annum between 2002 and 2015. Gallwey et al. (2020) in their paper discussed a technique of detecting artisanal small-scale mining (ASM) using a deep convolutional neural network model applied to open-source Sentinel-2 multispectral satellite imagery. The study concluded that this methodology can detect ASM in Ghana with a high degree of accuracy at a minimal cost in terms of financial and human resources. Obodai et al (2019) conducted a study using the Semi-Automatic Classification Plugin (SCP) in QGIS to assess these changes and the possible drivers of change in the Ankobra basin. The study found that closed forest cover in the basin had decreased from 40.4% in 1991 to 22.8% in 2016, with the dominant land use/cover change patterns being from closed forest to open forest, open forest to farmland, settlement/bare land, and mining area. Mining activities, particularly illegal mining, were identified as the dominant driver of deforestation in the basin between 2008 and 2016. A study by Kumi-Boateng and Stemn (2020) used satellite-based data to map out artisanal small-scale mining (ASM) sites in the Tarkwa-Nsuaem municipality. Several image processing techniques were applied to a Landsat 8 satellite image to identify 221 clusters of ASM sites, representing 12.72% of the total size of the study area. By analysing the location and distribution of the ASM sites in relation to environmentally sensitive and critical areas such as forest reserves, rivers, large-scale mines and urban settlements, the study identified areas warranting the urgent attention of society to mitigate the health, safety and ecosystem service effects of ASM.

The studies referenced emphasize the effectiveness of ML and remote sensing in monitoring the environmental impacts of illegal small-scale mining, particularly the loss of forest cover. By demonstrating successful applications of various ML techniques in detecting and quantifying land use changes, these studies provide a foundation for this research. This study builds on their methodologies to conduct a comparative analysis of different ML algorithms, aiming to identify the most effective methods for accurately assessing the impact of galamsey on vegetation in the Wassa Amenfi

East District. This directly supports the study's objective of improving environmental monitoring and management through advanced technology.

Accuracy assessment

Twenty-five (25) accuracy assessment points were generated and randomly distributed across the study area, ensuring that each class (Vegetation, Built-Up, Galamsey Affected Areas, and Water) was equally represented. These points were stratified across all land classes to ensure each category was adequately represented in the accuracy assessment, providing a balanced and comprehensive validation process (Fig. 7). Stratification ensured that no single class was underrepresented or overrepresented, which is essential for obtaining an accurate and reliable evaluation of the classification methods. The selected points were then identified on the classified image, and their real-world features were cross-referenced with the corresponding features in Google Earth for validation. This was done to ensure the accuracy of the classification for each class. To evaluate the performance of the classification techniques (MLC, SVM, RF, and ANN), a confusion matrix was generated for each method. The matrix was used to calculate key accuracy metrics, including producer accuracy, user accuracy, overall accuracy, and the kappa coefficient. These metrics helped assess the quality of the classification output and the agreement between the classified image and real-world observations. The confusion matrix was created using the regions of interest (ROIs) that were defined during the classification process, allowing for a detailed and thorough evaluation of the classification performance across all four techniques and across all land cover classes. This approach provided a clear and quantifiable means of comparing the performance of each machine learning algorithm in the context of galamsey impact assessment.

Results

Normalized Difference Vegetation Index

The NDVI results in Fig. 3 indicated that in 2015, NDVI values ranged from −0.89 to 0.73, reflecting a spectrum from unhealthy to moderately healthy vegetation. This suggests that, at the time, most areas in the district had relatively healthy vegetation cover. However, a noticeable change was observed over the eight-year period. By 2023, NDVI values ranged from −0.92 to 0.58, indicating a significant decline in vegetation health. The drop in maximum NDVI values, along with the occurrence of negative values, particularly −0.92, highlights areas of severe vegetation degradation. This deterioration is particularly concerning in regions affected by galamsey activities, where the vegetation has been extensively damaged, underscoring the environmental impact of these galamsey practices. The decline in NDVI over this period suggests a marked reduction in vegetation quality and coverage, signaling an urgent need for interventions to address the environmental damage.

In addition to the NDVI analysis, a Vegetation Health Index (VHI) analysis was conducted to evaluate the health and stress levels of vegetation cover in the study area affected by galamsey activities. The results, illustrated in Fig. 4, demonstrate the effectiveness of VHI in assessing the temporal impact on vegetation health. The use of

VHI is crucial for identifying areas under stress and understanding the extent of environmental degradation caused by these illegal mining practices.

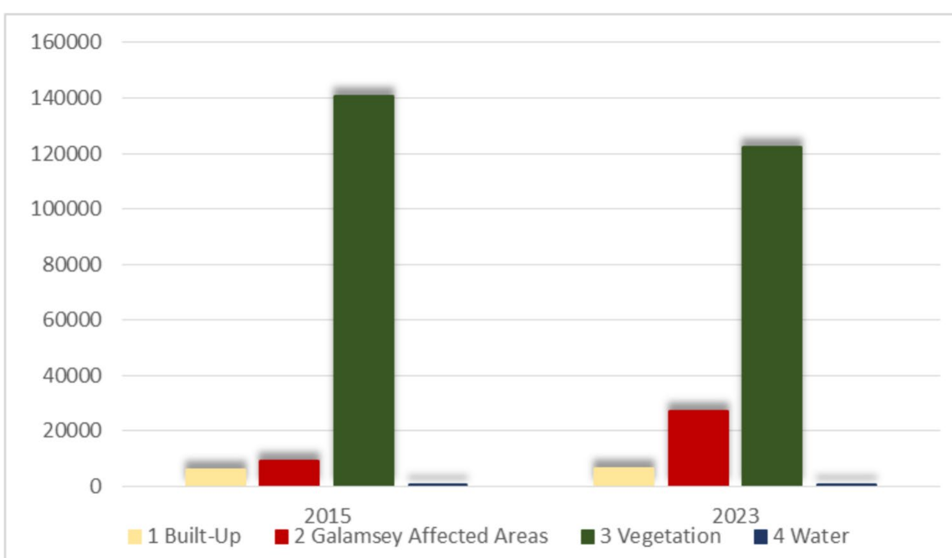

Fig. 7 Change Detection graph for 2015 and 2023 period in hectares

In the galamsey-affected areas within the study area, the VHI value dropped from approximately 14.2 in 2015 to 6.2 in 2023. This decline highlights the negative impact of galamsey activities on vegetation health, increasing stress levels and threatening the ecological balance. If these galamsey operations are not addressed, the continued degradation could lead to severe environmental consequences.

The results of LULC analysis for 2015 and 2023 using the MCL, SVM, ANN and RF algorithms are indicatedin Figs. 5 and 6 respectively. The LULC analysis shows a significant increase in galamsey (illegal mining) affected areas in 2023 compared to 2015, highlighting the growing impact of galamsey activities on vegetation. These findings emphasize the need for effective mitigation strategies to address this environmental challenge.

In 2015, the study area's land cover was predominantly vegetation, which covered 140,532 hectares (ha), accounting for 90% of the total land area. Water bodies occupied a mere 7.99 ha, representing only 0.000051% of the land cover. Galamsey-affected areas covered 9,381.80 ha (6.02%), and built-up areas covered 5,820.24 ha (4%). By 2023, significant changes in land cover were observed. Vegetation cover reduced to 122,343 ha, making up 79% of the land area, while water bodies decreased to 2.88 ha, now representing 0% of the land cover. Meanwhile, galamsey-affected areas expanded dramatically to 27,061.88 ha (17%), and built-up areas increased to 6,333.15 ha (4%).

Between 2015 and 2023, vegetation cover declined by 18,189.15 ha, reflecting a notable reduction in the district's green cover. In contrast, areas impacted by galamsey increased by 17,680.08 ha, and built-up areas grew by 512.91 ha. The change detection graph in Fig. 7 compares changes in vegetation cover, galamsey affected areas, built-up areas, and water bodies between 2015 and 2023. The graph highlights a significant increase in galamsey affected areas and decreased in vegetation cover.

This sharp rise in galamsey-affected land indicates that illegal mining is a significant driver of vegetation degradation in the Wassa Amenfi East district (Ezeji Onyebuchi 2014). The data stresses the urgent need to address the environmental consequences of galamsey, as its expansion poses a severe threat to the region's ecological health.

Assessment of results ML classifiers

Figure 8 shows the spatial distribution of accuracy assessment points used in creating the confusion matrix of and the confusion matrix.

Tables 4 evaluates the performance of various ML algorithms using overall accuracy (OA), Kappa Coefficient (KC), Recall (R), Precision (P), and F-Score. The results indicate that SVM, RF, and ANN outperformed MLC in all parameters. Specifically, RF achieved the highest scores, followed by ANN, for the 2015 classified image across all evaluation metrics, demonstrating a strong agreement with actual class distributions and effectively mapping land cover types.

In 2023, SVM, ANN, and RF continued to outperform MLC, with a slight improvement in the performance of SVM, which achieved the highest recall value of 0.90. Despite this improvement in recall, SVM, ANN, and RF still outperformed MLC in all the other evaluating parameters, maintaining their superior classification capabilities. The superior performance of RF can be attributed to its ensemble learning approach, which aggregates the results of multiple decision trees, thus enhancing its ability to capture complex, non-linear relationships in the data. This also helps mitigate overfitting, a common issue in single-tree models. Furthermore, RF's built-in feature importance mechanism allowed for a more accurate classification of land use and land cover types, especially in distinguishing areas impacted by galamsey activities. Accuracy and Kappa were the primary metrics used in evaluating the models, providing an overview of overall classification performance. While these metrics are valuable, future work could benefit from a broader set of evaluation criteria, such as precision, recall, and F1-score, to provide a more nuanced understanding of model strengths and weaknesses. Throughout both years, RF and ANN consistently showed high performance across all evaluation parameters, with RF achieving the highest scores. This consistency emphasizes the robustness and reliability of RF and ANN classifiers in producing accurate land cover classifications over time.

Using multiple evaluation metrics such as overall accuracy, Kappa coefficient, recall, precision, and F-score provides different perspectives and a balanced way of assessing ML classifiers. This assists in choosing the best and most robust classifier(s) for mapping galamsey activity on vegetation, leading to more efficient decision-making.

Figure 9 illustrates the water pits created by galamsey activities within the vegetated zone of Wassa Amenfi East District. This visual representation highlights the extent of landscape and vegetation cover destruction caused by these activities. It emphasizes the significant environmental impact of galamsey and emphasizes the urgent need for responsible and sustainable mining practices to mitigate these adverse effects.

Discussion

Performance of classification algorithms

In this comparative study, four classifiers were employed, demonstrating Random Forest's (RF) superior performance relative to the other models. As detailed in Table 4, RF consistently achieved high performance in both years, with

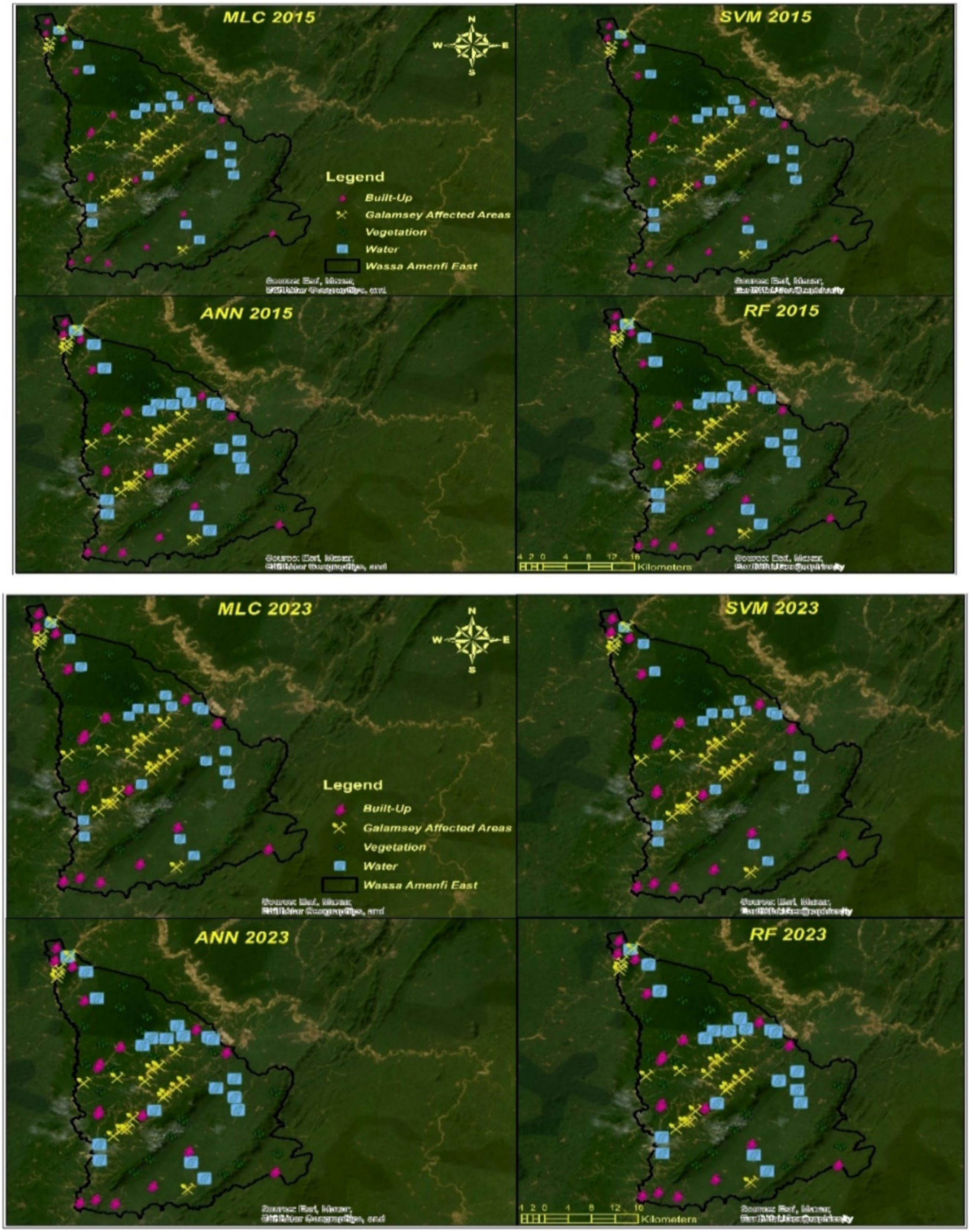

Fig. 8 Sampled points taken for the accuracy assessment over the study area

Table 4 Comparative performance of all the algorithms

ML Algorithm	2015					2023				
	OA	KC	P	R	F-Score	OA	KC	P	R	F-Score
MLC	0.78	0.71	0.78	0.85	0.82	0.83	0.78	0.83	0.79	0.81
SVM	0.84	0.79	0.84	0.85	0.84	0.85	0.8	0.85	0.9	0.87
ANN	0.85	0.8	0.85	0.87	0.86	0.86	0.81	0.87	0.88	0.87
RF	0.88	0.84	0.88	0.89	0.89	0.87	0.82	0.87	0.89	0.88

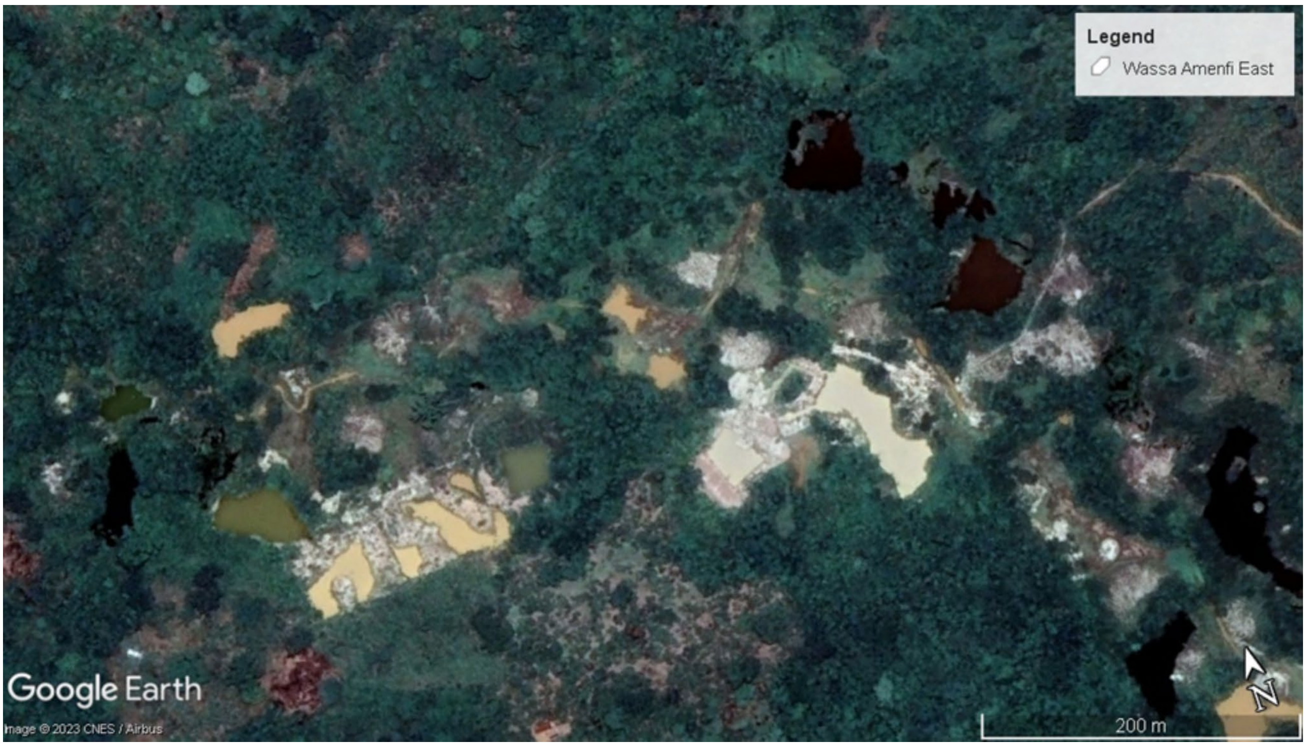

Fig. 9 An imagery in the Wassa Amenfi East District in 2023 showing the water pits created by *galamsey* activities which is a recurring effect in areas plagued with this activity. Source: Google Earth Images

overall accuracy rates of 88% and 87% and Kappa values of 0.84 and 0.82 for 2015 and 2023, respectively. These results corroborate the findings of Barenblitt et al. (2021), Tian et al. (2016), and Shih et al. (2019), which highlight RF's effectiveness in achieving accurate and reliable land cover classifications.

Notably, the conclusive remarks in the investigations conducted by Tariq and Mumtaz (2023) and Taati et al. (2015) indicated SVM as the optimal classifier for land classification. However, these studies did not incorporate RF into their comparative analyses. Tariq and Mumtaz (2023) juxtaposed SVM with Kernel Logistic Regression and Naïve Bayes, while Taati et al. (2015) contrasted SVM with MLC. Yet, introducing RF could have provided a fresh perspective and insight into their study. In a recent study by Pandey et al. (2025), SVM outperformed RF and MLC in delineating land use and vegetation classes, likely due to the use of minimal training data. Their study suggests that SVM can surpass RF in scenarios with minimal training data and high dimensionality, aligning with the earlier assertions made by EL-Omairi & El Garouani (2023). Despite prevailing beliefs that SVM excels as an optimal classifier, the results of this study deviate from that consensus. Specifically, this study's findings reveal that SVM only demonstrated superiority over MLC in both instances, corroborating Taati et al. (2015) conclusions.

Remarkably, all the remaining classifiers outperformed MLC, the conventional technique in classification (Sisodia et al. 2014). Unlike studies solely reliant on MLC for classification, such as Aseidu et al. (2023) and Baddianaah et al. (2022), this research offers valuable insights by assessing multiple techniques' accuracy. Furthermore, when pitted against more robust methods like in this instance, MLC's performance falls short. ANN secured the second position after RF, slightly surpassing SVM and clearly outperforming

MLC, thus validating Mahmon and Ya'acob (2014) findings. Shih et al. (2019) compared random forest (RF), support vector machine (SVM), decision tree (DT), and artificial neural network (ANN); found random forest performed better than the other three algorithms affirm the findings of this study. However, a study by Valero Medina and Alzate Atehortúa (2019) indicates support vector machines as the best, followed by random forest algorithms and maximum likelihood.

In this study, the RF classifier emerged as the superior choice for accurately conducting land cover classification and producing galamsey-affected location maps. Dias et al. (2021) affirmed this assertion when they used SVM, random forest, and maximum likelihood classifiers to classify landslides in Itaóca (São Paulo, Brazil). Majikumna et al. (2024) confirmed that RF outperformed SVM and K-Nearest Neighbor (KNN) in mapping LULC, while Yimer et al. (2024) identified RF as the superior ML classifier in an LULC classification task, surpassing SVM, KNN, MLC, and Classification and Regression Tree (CART). Verma et al. (2025) further reinforced these findings by achieving high classification performance with RF over SVM and ANN using UAV remote sensing imagery. These studies collectively emphasize RF's dominance in classification tasks across diverse geographic and environmental contexts.

Notably, RF's versatility extends beyond classification to regression applications, particularly when paired with advanced preprocessing techniques. Giri et al. (2024) demonstrated this adaptability by combining RFR with Mexican Hat wavelet transformation to analyze soybean seed composition using hyperspectral canopy-level data. This aligns with RF's documented ability to capture non-linear relationships in classification tasks while highlighting its capacity to address regression problems when integrated with feature enhancement methods. The RF classifier's precision in distinguishing galamsey-affected areas in the Wassa Amenfi East district mirrors its broader success in remote sensing. Its robustness against overfitting, coupled with its ability to handle high-dimensional data (even when compounded by preprocessing workflows like wavelet decomposition), positions RF as a critical tool for both classification and regression in geospatial analytics. These findings suggest that RF's utility is not confined to discrete categorical outputs but extends to continuous parameter estimation, particularly when optimized with domain-specific signal processing techniques.

Contribution to global literature on mining and machine learning methods

This study makes a notable contribution to the global literature on galamsey and ML methods, particularly in assessing the environmental impact of mining activities on vegetation. While past studies have generally focused on applying machine learning techniques in environmental monitoring, this research distinguishes itself by conducting a comparative analysis of four prominent machine learning models SVM, RF, ANN, and MLC. By evaluating and comparing these models in the specific context of galamsey-affected areas, the study offers new insights into the relative strengths and limitations of each algorithm, contributing to a deeper understanding of how ML methods can be applied in assessing environmental degradation caused by mining activities. In addition to providing a comparative analysis, this research advances the methodology of using machine learning in environmental monitoring. The study combines high-resolution remote sensing data with ML techniques to accurately assess the effects of galamsey on vegetation in the Wassa Amenfi East District of Ghana. This innovative methodology not only improves our ability to detect changes in land cover but also demonstrates the potential of using ML to map and monitor illegal mining activities in real-time. The integration of satellite imagery, vegetation indices, and other environmental factors provides a holistic approach to assessing land use changes, offering a more reliable and efficient tool for monitoring and managing areas affected by illegal mining.

Moreover, this study has significant global implications for mining-related environmental research. While the focus is on Ghana, the methodology and findings can be applied to other regions facing similar challenges with illegal mining. The comparative analysis of ML models offers a framework adaptable to different environmental contexts, making this approach useful for regions with varying climatic and geographical conditions. Furthermore, the insights gained on the spatial and temporal effects of galamsey on vegetation serve as a model for assessing the broader environmental impacts of mining activities worldwide, contributing to more effective policy-making and sustainable land management practices. This study also offers valuable contributions to sustainable land management strategies in mining-affected regions. By identifying key drivers of environmental degradation and demonstrating the effectiveness of machine learning techniques in detecting and mapping these changes, the research provides actionable insights for policymakers and environmental managers. The study highlights the need for more effective monitoring and management of mining activities, offering a technological solution that can be employed to track and mitigate the environmental consequences of galamsey. As global interest in sustainable development and environmental protection continues to grow, the findings from this research can inform future efforts to implement sustainable land management practices and policies that mitigate the adverse effects of illegal mining on ecosystems and local communities.

Impact of Galamsey on forest resources

In developing countries like Ghana, deforestation and forest degradation continues to be the primary environmental issues as they account for about 18% of global carbon dioxide emission (Boadi et al. 2017). Surface mining in general requires deforestation and clearing of nutritious soil top for the extraction of the precious metal. Galamsey operations are mostly undertaken in forest zones of Ghana (Mantey et al. 2018) and this would lead to flora and fauna destruction, water and air pollution and land degradation (Arthur-holmes et al. 2022).

From Fig. 9, the galamsey sites are left abandoned without any proper reclamation plans to restore forest or vegetation regrowth (Barenblitt et al. 2021). The main reason that accounted for the galamsey site abandonment is that galamsey operations within the forest are hard to detect and monitor. Mantey et al. (2018) highlight the challenging nature of detecting and locating galamsey operations. This difficulty arises from the illegality, scattered nature, and undercover operations associated with galamsey. The clandestine and dispersed characteristics of these activities make them elusive and hard to identify, posing significant challenges for authorities and efforts aimed at monitoring and curbing illegal mining operations. With remote sensing imagery and ML techniques, this present study detected and mapped the possible galamsey sites with high prediction accuracies. From Table 3, as of 2023, galamsey operations have claimed 27,061.9 hectares of forest cover which hosts diverse biodiversity and ecosystems. The loss of the ecosystem and the degraded galamsey environment would have direct and indirect impacts on the livelihood of the surrounding communities (Obodai et al. 2023). The degraded galamsey environments used to host the source of livelihood for the rural poor farmers have been destroyed with or without any proper compensation paid. Protecting forest cover means protecting people's source of livelihood.

The area changes in percent for degraded galamsey land for the past seven years from 2015 to 2023 was 188.45% (Table 3). This rate of change was highest among the land use land cover mapped within the study area and if this rate continues it would seriously affect the forest cover. Ghana is losing its accessible rainforests due to mining and other anthropogenic activities such as agriculture and logging as the forest cover of Ghana has reduced from 8.2 million hectares to 1.6 hectares with an annual deforestation rate of 2% (Boadi et al. 2017). This finding of the study affirms Boadi et al. (2017) assertion.

The NDVI and VHI analysis revealed vegetation loss and health decline in Ghana's forest zones, highlighting the urgent need to halt illegal mining (galamsey) activities. Consequently, the Government of Ghana has banned all forms of mining in forested areas, especially within forest reserves.

Ghana has prioritized addressing the impact of galamsey operations in forest and water bodies as the operations of galamsey have a potential impact in achieving SGD 15, especially Target 15.1 “By 2020, ensure the conservation, restoration and sustainable use of terrestrial and inland freshwater ecosystems and their services, in particular forests, wetlands, mountains and drylands, in line with obligations under international agreements”.

Drivers of Galamsey

The factors influencing Ghanaians to engage in galamsey are summarized in Table 4. These drivers of galamsey are numerous but the dominant ones documented in the literature are mentioned in this present study Table 5.

Galamsey in its total is a complex and diverse issue and it has been influenced by diverse factors. Identifying and understanding these drivers would help in proposing solutions to address the menace in the study area and Ghana as a whole. The influence of these drivers varies in urging Ghanaians to engage in galamsey operations. The main driver of this issue in Ghana could be

Table 5 Potential driving factors of galamsey in Ghana

Category	Drivers of Galamsey	Reference
Socio-Cultural Factors	Desire for 'quick money'	Hilson (2017)
	Influence of traditional authorities	Hilson et al. (2014)
	Lifestyle and societal norms	Hilson (2017)
Political Factors	Poor policies and implementation	Eduful et al. (2020)
	Lack of enforcement of environmental laws/justice	Hilson et al. (2014)
	Governance systems	
Technological/Economic Factors	Youth unemployment	Hilson (2017)
	Poverty	Bush (2009)
	Industrialization and influx of private investors	Hilson (2017)
	External factors (e.g., fluctuations in the prices of commodities like gold and timber)	Hilson (2017); Bush (2009)
		Bush (2009)
	Source of livelihood	

attributed to high poverty in rural communities in Ghana (Bush 2009). While galamsey sites are typically owned by wealthy individuals, the majority of the workforce engaged in galamsey operations consists of individuals from impoverished backgrounds who often migrate from socially and economically marginalized communities. This phenomenon stresses the stark contrast between the affluent ownership of these sites and the poverty-driven composition of the labor force involved in galamsey activities (Hilson 2012). Tschakert and Singha, (2007) confirmed this assertion by describing galamsey as poverty-driven activity. Hilson and Hu (2022) also concluded that each of the galamsey sites in Sub-Saharan Africa is populated with heterogeneous groups of poverty-driven people. Even though galamsey has been considered a catalyst to reduce poverty it comes with high risks of illness, accidents and physical hazards that can perpetuate poverty (Eduful et al. 2020).

Another key driver forcing people into galamsey is unemployment among the youth. Galamsey has been an avenue where both the trained skills and unskilled labour are offered a variety of employment opportunities for men, women and even children (Eduful et al. 2020). Galamsey sites also serve as trading grounds where goods and services are exchanged and this in a way helps reduce the high unemployment among the rural communities. *Galamsey* has been a source of livelihood. Globally, it is estimated that about 20 million people depend on illegal mining as their source of livelihood (Obodai et al. 2023). Lydia et al. (2022) indicate that galamsey can improve the livelihood and the living standards of people who engage in its operation because of the money it provides to them. Galamsey has been the source of livelihood for both urban and rural populations who are not offered employment in the formal sector of the Ghanaian economy (Eduful et al. 2020). As agriculture diminishes in prominence as the primary source of livelihood in Sub-Saharan Africa, galamsey has emerged as an alternative means of sustaining the rural population. The decline in the agricultural sector has compelled many individuals in rural areas to turn to galamsey as a substitute for traditional farming activities. This shift underscores the evolving economic landscape in the region, with galamsey becoming a viable option for those seeking alternative sources of income and livelihood in the face of agricultural challenges (Obodai et al. 2023).

There is a perception that some people engage in galamsey, because of their desire to earn 'quick money'(Hilson 2017). The value of gold over other sources of employment is the reason most people engage in galamsey; they see galamsey as an avenue where cheap labour forces are used to achieve high earnings within the shortest possible time. Hilson and Hu, (2022) described galamsey as the platform through which wealth is created for all its players.

Most Ghanaians are of the view that government and state institutions are to be blamed for the negative impacts of galamsey. They argue that government and state institutions have failed to implement sustainable programs and policies to combat the negative impact of galamsey (Afriyie et al. 2016). Eduful et al, (2020) argued that policies and laws implemented by government and state institutions to regularize the operations of galamsey are rigid, centralized and bureaucratic and in the end increase the negative impact of galamsey on the environment especially on the water bodies and forest. For instance, the use of a military approach to stop galamsey did not work to its perfection as it failed to solve the main drivers of galamsey but rather created inequalities and deepened of socio-economic conditions of the galamsey mining communities (Eduful et al. 2020).

The influence of chiefs, traditional authorities and government officials cannot be ignored as drivers of galamsey operations in Ghana. Some chiefs, traditional authorities and government officials are accused of playing active in promoting galamsey activities in their localities. They assist in providing financial support, lands and labour to the operations of galamsey (Hilson et al. 2014). Armah et al, (2013) argued that local government officials, politicians, chiefs, and high-ranking governments are directly or indirectly engaging or supporting galamsey operations. The support from such authorities is making the fights against the negative impacts of galamsey difficult.

Conclusion

The study analysed the impact of galamsey on vegetation/forest by using four different ML algorithms. The result of the analysis showed that RF classification outperformed all other classifications. For overall accuracy, RF scored 88% and 87% with kappa coefficient values of 0.84 and 0.82 in 2015 and 2023 respectively. The high performance is due to RF's non-parametric nature, classification accuracy, and ability to determine variable importance. Using large training samples of about fifty ensured high accuracy and kappa values for classifying land use land cover types in both years. The application of ML in environmental monitoring and assessment in remote sensing has proven effective in gauging the impact of *galamsey* on vegetation. The evaluation of the various classifications provided significant insight into which of the algorithms is much more suitable for mapping galamsey-affected areas. The study indicated that galamsey operations have affected vegetation/forest cover and contributed to the loss of biodiversity. The study identified the key drivers influencing galamsey operations as poverty, youth unemployment, poor policies and implementation, source of livelihood, traditional authorities and chiefs and desire for 'quick money'. The 27061.9 hectares of galamsey-affected areas can be attributed to one of these drivers. Poverty is identified as the main driver influencing people's decision to engage in galamsey activities. Finaly, the application of ML in assessing galamsey impacts on vegetation in Ghana's High Forest Zone offers a robust and scalable approach to mitigating environmental impacts. The methodology can be replicated in other regions in Ghana or globally where galamsey threatens the environment, providing a scalable solution for widespread monitoring and effective intervention.

Limitations and future research perspectives

The application of Landsat 8 imagery, while instrumental in land cover classification, presents notable limitations. The moderate spatial resolution (30 meters) constrains the detection of fine-scale land cover heterogeneity, particularly in fragmented or mosaic landscapes where mixed-pixel effects introduce inherent classification ambiguities. Furthermore, the 16-day revisit cycle impedes the timely capture of rapid land cover transitions, limiting real-time monitoring of transient environmental disturbances. Although atmospheric correction protocols were rigorously applied, residual artifacts from aerosol scattering and cloud contamination may persist and compromise classification outcomes. Future studies could mitigate these constraints by integrating multi-source data fusion approaches for instance, combining Sentinel-2's enhanced spatial-temporal resolution (10–20 meters, 5-day revisit) with commercial platforms like PlanetScope (3-meter resolution) or UAV-derived hyperspectral datasets to refine classification granularity and enable near-real-time detection of subtle ecological shifts.

Secondly, expanding interdisciplinary research on galamsey's ecological impacts is critical to addressing knowledge gaps. Longitudinal studies spanning decadal scales could systematically elucidate the cascading effects of illicit mining on forest degradation, biodiversity loss, and soil hydrology. Concurrently, spatiotemporal modeling of galamsey hotspots using high-frequency satellite time series (e.g., monthly composites) could reveal operational patterns, enabling predictive analytics for pre-emptive enforcement. Investigating the interplay of socioeconomic drivers such as unemployment, informal gold markets, and weak regulatory frameworks through mixed-methods approaches would strengthen causal inference and inform context-specific policy reforms. Community-engaged participatory research, coupled with geospatial vulnerability mapping, could identify pathways for co-designing restoration initiatives that align ecological resilience with livelihood security.

Lastly, advancing remote sensing innovation is paramount: hyperspectral imaging could discriminate mineralogical signatures of galamsey sites, while UAV-LiDAR fusion might quantify 3D canopy structural damage unreachable by optical sensors. Machine learning architectures (for instance, transformer-based models) trained on multitemporal, multi-sensor datasets could automate anomaly detection at scale, enhancing monitoring efficiency. Lastly, quantifying the cascading loss of ecosystem services carbon sequestration collapse, watershed contamination, and microclimate alteration would galvanize stakeholder urgency for evidence-based conservation frameworks. By synergistically integrating these methodologies, future research can bridge disciplinary silos, foster scalable mitigation strategies, and catalyze transnational cooperation for sustainable land governance.

Acknowledgement The authors will like to thank the handling editor and anonymous reviewers for their insightful comments on the manuscript.

Author Contribution C.K.: Supervision, Project Administration, Provision of Resources, Review and Approval of Final Draft; I.K.A.: Conceptualization, Design, Data curation, Methodology, Formal analysis, Writing-Original and final drafts, Review and Editing; E.Y.: Conceptualization, Design, Data curation, Methodology, Formal analysis, Writing-Original and final drafts, Review and Editing; I.S.: Formal Analysis, Validation, review and editing; M.B.: Review and Editing, A.O: Review and Editing; M.M.S: Review and Editing; W.S.: Review and Editing; D.B: Review and Editing; R.O.L : Review and Editing; J.A.: Review and Editing; C.E.R.: Review and Editing A.H.J.: Review and Editing F.S.: Review and Editing R.A.: Review and Editing O.I.: Review and Editing

Funding This work was not funded by any organization or institution.

Data availability Data will be made available on request.

Declarations

Declaration All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors.

Competing Interests The authors declare no competing interests.

Ethical approval This declaration is not applicable, as no personal or traceable information was required for the analysis.

References

Afele JT, Dawoe E, Abunyewa A, Afari-Sefa V, Asare R (2021) Carbon storage in cocoa growing systems across different agroecological zones in Ghana. Coffee Cocoa Res J 37(1):32–49. https://doi.org/10.22302/iccri.jur.pelitaperkebunan.v37i1.395

Afele JT, Nimo E, Lawal B, Afele IK (2022) Deforestation in Ghana; evidence from selected forest reserves across six ecological zones. Int J For Anim Fish Res 6(1):7–17. https://doi.org/10.22161/ijfaf.6.1.2

Afriyie K, Ganle JK, Adomako JAA (2016) The good in evil: a discourse analysis of the galamsey industry in Ghana. Oxford Dev Stud 44(4):493–508. https://doi.org/10.1080/13600818.2016.1217984

Afriyie K, Abass K, Frempong F, Arthur B, Gyasi RM (2023) The dynamics and livelihood implications of illegal mining in Ghana: a critical assessment. Geogr Res 61(1):32–43. https://doi.org/10.1111/1745-5871.12573

Ahmad M, Jiang P, Majeed A, Umar M, Khan Z, Muhammad S (2020) The dynamic impact of natural resources, technological innovations and economic growth on ecological footprint: an advanced panel data estimation. Resour Policy 69:101817. https://doi.org/10.1016/j.resourpol.2020.101817

Armah FA, Luginaah IN, Taabazuing J, Odoi JO (2013) Artisanal gold mining and surface water pollution in Ghana: have the foreign invaders come to stay? Environ Justice 6(3):94–102. https://doi.org/10.1089/env.2013.0006

Arthur-holmes F, Abrefa K, Yakovleva N, Vazquez- DA (2022) Artisanal and small-scale mining methods and the sustainable development goal 6: perceived implications for clean water supply. Environ Sci Policy 137(September):205–215. https://doi.org/10.1016/j.envsci.2022.08.017

Aryee BN, Ntibery BK, Atorkui E (2003) Trends in the small-scale mining of precious minerals in Ghana: a perspective on its environmental impact. J Clean Product 11(2):131–140. https://doi.org/10.1016/S0959-6526(02)00043-4

Asiedu RA, Amaning KA, Koomson B, Gyamfi C, Niang A (2023) Assessing land degradation in Amansie West District: a surface mined area in Ghana

Baddianaah I, Baatuuwie BN, Adongo R (2022) Local perspectives on the adverse environmental effects and reclamation of illegally mined degraded landscapes in North-western Ghana. Min Econ 36(1):139–155. https://doi.org/10.1007/s13563-022-00336-0

Baddianaah I, Abdulai IA, Dordaa F (2023) Sorting the facts from the lots: contribution of artisanal and smallscale mining (galamsey) to rural livelihood configurations in sub-Saharan Africa. Journal of Degraded & Mining Lands Management, 11(1)

Barenblitt A, Payton A, Lagomasino D, Fatoyinbo L, Aidoo K, Pigott H, Ko C, Smeets L, Seidu O, Wood D (2021) The large footprint of small-scale artisanal gold mining in Ghana. Sci Total Environ 781:. https://doi.org/10.1016/j.scitotenv.2021.146644

Belgiu M, Drăguţ L (2016) Random forest in remote sensing: a review of applications and future directions. ISPRS J Photogramm Remote Sens 114:24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Biau G, Scornet E (2016) A random forest guided tour. Test 25:197–227. https://doi.org/10.1007/s11749-016-0481-7

Biney E, Biney N, Dadzie I, Harris E, Quartey GA, Asare YM, Forkuo EK (2022) Impact of mining on vegetation cover: a case study of Prestea Huni-Valley municipality. Sci Afr 17:e01387. https://doi.org/10.1016/j.sciaf.2022.e01387

Boadi S, Ayine C, Owusu O, Acquah E (2017) An analysis of illegal mining on the Of fi n shelterbelt forest reserve, Ghana : implications on community livelihood. J Sustain Min 15(3):115–119. https://doi.org/10.1016/j.jsm.2016.12.001

Bush R (2009) 'Soon there will be no-one left to take the corpses to the morgue': accumulation and abjection in Ghana's mining communities. Resourc Policy 34(1–2):57–63. https://doi.org/10.1016/j.resourpol.2008.02.002

Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A (2020) A comprehensive survey on support vector machine classification: applications, challenges and trends. Neurocomputing 408:189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Cichy RM, Kaiser D (2019) Deep neural networks as scientific models. Trend Cogn Sci 23(4):305–317 https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(19)30034-8

Dias HC, Sandre LH, Alarcón DAS, Grohmann CH, Quintanilha JA (2021) Landslide recognition using SVM, random forest, and maximum likelihood classifiers on high-resolution satellite images: a case study of Itaóca, southeastern Brazil. Braz J Geol 51(4):. https://doi.org/10.1590/2317-4889202120200105

Domadia SG, Zaveri T (2011) Comparative analysis of unsupervised and supervised imageclassification techniques. In Proceedings of the National Conference on Recent Trends in Engineering & Technology pp 1–5. Available at: https://www.bvmengineering.ac.in/misc/docs/published-20papers/etel/etel/401043.pdf

Eduful M, Alsharif K, Eduful A, Acheampong M, Eduful J, Mazumder L (2020) The illegal artisanal and small-scale mining (Galamsey) 'Menace' in Ghana: is military-style approach the answer? Resourc Policy 68:101732. https://doi.org/10.1016/j.resourpol.2020.101732

EL Omairi MA, El Garouani A (2023) A review on advancements in lithological mapping utilizing machine learning algorithms and remote sensing data. Heliyon 9(9):e20168. https://doi.org/10.1016/j.heliyon.2023.e20168

Ericsson M, Löf O (2019) Mining's contribution to national economies between 1996 and 2016. Mineral Economics 32(2):223–250

Ezeji Onyebuchi E (2014) Artisanal and Small-Scale Mining in Wassa Amenfi East District, (Ghana). https://urn.fi/URN:NBN:fi:amk-2014090413740

Florian, R, Sarasua C, Bernstein A (2022) Bam: Benchmarking argument mining on scientific documents. CEUR Workshop Proceedings

Gallwey, J., Robiati, C., Coggan, J., Vogt, D., & Eyre, M. (2020). A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining. Remote Sensing of Environment, 248, 111970

Ghana Statistical Service. (2021). Ghana 2021 Population and Housing Census General Report. https://census2021.statsghana.gov.gh/

Giri A, Sagan V, Alifu H, Maiwulanjiang A, Sarkar S, Roy B, Fritschi FB (2024) A Wavelet Decomposition Method for Estimating Soybean Seed Composition with Hyperspectral Data. Remote Sens 16(23):. https://doi.org/10.3390/rs16234594

Ha NT, Manley-Harris M, Pham TD, Hawes I (2020) A comparative assessment of ensemble-based machine learning and maximum likelihood methods for mapping seagrass using sentinel-2 imagery in Tauranga Harbor, New Zealand. Remote Sens 12(3):355. https://doi.org/10.3390/rs12030355

Hilson G (2012) Poverty traps in small-scale mining communities: the case of sub-Saharan Africa. Cana J Dev Stud/Revue Canadienne d'études Du Développement 33(2):180–197. https://doi.org/10.1080/02255189.2012.687352

Hilson G (2017) Shootings and burning excavators: some rapid reflections on the Government of Ghana's handling of the informal Galamsey mining 'menace.' Resourc Policy 54:109–116. https://doi.org/10.1016/j.resourpol.2017.09.009

Hilson G, Hu Y (2022) Changing priorities, shifting narratives: remapping rural livelihoods in Africa's artisanal and small-scale mining sector. J Rural Stud 92:93–108. https://doi.org/10.1016/j.jrurstud.2022.03.010

Hilson G, Hilson A, Adu-Darko E (2014) Chinese participation in Ghana's informal gold mining economy: drivers, implications and clarifications. J Rural Stud 34:292–303. https://doi.org/10.1016/j.jrurstud.2014.03.001

Isinkaralar O, Yeboah E, Isinkaralar K, Sarfo I, Öztürk S, Yilmaz D, Bojago E (2025) Investigating the effects of local climate zones on land surface temperature using spectral indices via linear regression model: a seasonal study of Sapanca Lake. Environ Monit Assess 197(3):252. https://doi.org/10.1007/s10661-025-13705-3

Jabbar FA, Rithanya KS, Miruthubashini M, Naveena KS (2021) Risks of chronic kidney disease prediction using various data mining algorithms. International Journal of Informatics, Information System and Computer Engineering (INJIISCOM) 2(2):165–177

Jabbar M, Yusoff MM (2022) Assessing the spatiotemporal urban green cover changes and their impact on land surface temperature and urban heat island in Lahore (Pakistan). Geogr Environ Sustain 15(1):130–140. https://doi.org/10.24057/2071-9388-2021-005

Kumi-Boateng B, Mireku-Gyimah D, Duker AA (2012) A spatio-temporal based estimation of vegetation changes in the Tarkwa mining area of Ghana. Rese J Environ Earth Sci 4(3):215–229 (https://www.airitilibrary.com/Article/Detail/20410492-201203-201507220025-201507220025-215-229)

Kumi-Boateng B, Stemn E (2020) Spatial analysis of artisanal and small-scale mining in the Tarkwa-Nsuaem Municipality of Ghana. Ghana Mining Journal 20(1):66–74

Kwang C, Blagogie L (2024) The Impact of artisanal mining on agricultural production and food security in the Birim North District of Ghana. Acta Sci Pol Formatio Circumiectus 23(4):47–60. https://doi.org/10.15576/ASP.FC/196261

Kyere-Boateng R, Marek MV (2021) Analysis of the social-ecological causes of deforestation and forest degradation in Ghana: Application of the DPSIR Framework. Forests 12:409

Lydia O, Godwin A, Isaac L (2022) the extractive industries and society " We have done nothing wrong " : youth miners ' perceptions of the environmental consequences of artisanal and small-scale mining (ASM) in Ghana. Extr Ind Soc 12(July):101179. https://doi.org/10.1016/j.exis.2022.101179

Mahmon NA, Ya'acob N (2014) A review on classification of satellite image using Artificial Neural Network (ANN). 2014 IEEE 5th Control and system graduate research colloquium. IEEE, Shah Alam, pp 153–157. https://doi.org/10.1109/ICSGRC.2014.6908713

Mahmoudi A, Takerkart S, Regragui F, Boussaoud D, Brovelli A (2012) Multivoxel pattern analysis for fMRI data: a review. Comput Math Methods Med 2012:1–14. https://doi.org/10.1155/2012/961257

Majikumna KU, Zineddine M, Alaoui AEH (2024) Assessing drought stress impact on olive trees using AI and Sentinel-2 satellite imagery. World Water Policyhttps://doi.org/10.1002/wwp2.12243

Mantey J, Nyarko KB, Aubynn A (2018) Spatial distribution patterns of illegal artisanal small scale gold mining (Galamsey) operations in Ghana: a focus on the Western Region. Heliyon 2018(4):e00534. https://doi.org/10.1016/j.heliyon.2018.e00534

Maxwell AE, Warner TA, Fang F (2018) Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens 39(9):2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mohan L, Pant J, Suyal P, Kumar A (2020) Support vector machine accuracy improvement with classification. 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, Bhimtal, pp 477–481. https://doi.org/10.1109/CICN49253.2020.9242572

Mustapha AO, Mbuzukongira P, Mangala MJ (2007) Occupational radiation exposures of artisans mining columbite–tantalite in the easternDemocratic Republic of Congo. J Radiol Prot 27(2):187

Obeng EA, Oduro KA, Obiri BD, Abukari H, Guuroh RT, Djagbletey GD, ... Appiah M (2019) Impact of illegal mining activities on forest ecosystem services: local communities' attitudes and willingness to participate in restoration activities in Ghana. Heliyon, 5(10)

Obodai J, Adjei KA, Odai SN, Lumor M (2019) Land use/land cover dynamics using landsat data in a gold mining basin-the Ankobra, Ghana. Remote Sensing Applications: Society and Environment 13:247–256

Obodai J, Mohan G, Bhagwat S (2023) The extractive industries and society beyond legislation : unpacking land access capability in small-scale mining and its intersections with the agriculture sector in sub-Saharan Africa. Extr Ind Soc 16(April):101357. https://doi.org/10.1016/j.exis.2023.101357

Obodai J, Bhagwat S, Mohan G (2024) The interface of environment and human wellbeing: Exploring the impacts of gold mining on food security in Ghana. Resourc Policy, 91:. https://doi.org/10.1016/j.resourpol.2024.104863

Ocansey I (2013) Mining impacts on agricultural lands and food security: case study of towns in and around Kyebi in the Eastern Region of Ghana. https://www.theseus.fi/handle/10024/53720

Onumah J (2010) Implementing warehouse receipt systems in Africa: Potential and challenges. In Fourth African Agricultural Markets Program Policy Symposium, organized by the Alliance for Commodity Trade in Eastern and Southern Africa (ACTESA) of the Common Market for Eastern and Southern Africa (COMESA) pp 6–7. Available at: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=fa8fc05ae5cafbea996280e357de06074782f46b

Opoku ME, Salkushu W, Frank H (2017) A hybrid image classification approach to monitoring LULC changes in the mining district of Prestea-Huni Valley, Ghana

Osman N, Afele JT, Nimo E, Gorleku D, Ofori L, Abunyewa A (2022) Assessing the impact of illegal small-scale mining (Galamsey) on cocoa farming and farmer livelihood: a case study in the Amansie West District of Ghana. Pelita Perkebunan (Coffee Cocoa Res J) 38(1):70–82

Owusu-Nimo F, Mantey J, Nyarko KB, Appiah-Effah E, Aubynn A (2018) Spatial distribution patterns of illegal artisanal small scale gold mining (Galamsey) operations in Ghana: a focus on the Western Region. Heliyon 4(2):e00534

Pandey M, Mishra A, Swamy SL, Anderson JT, Thakur TK (2025) Machine learning-based monitoring of land cover and reclamation plantations on coal-mined landscape using Sentinel 2 data. Environ Sustain Indic 25:100585. https://doi.org/10.1016/j.indic.2025.100585

Patil, U. M., & Patil, J. B. (2012, August). Web data mining trends and techniques. In Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 961–965

Pu R, Landry S, Yu Q (2011) Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. Int J Remote Sens 32(12):3285–3308

Ray S (2019) A quick review of ML algorithms. 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon). IEEE, Faridabad, pp 35–39

Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP (2012) An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS journal of photogrammetry and remote sensing 67:93–104

Sachin DK (2022) Emotion Recognition from DEAP Dataset Using SVM Classifier. Available at https://ijsrem.com/download/emotion-recognition-from-deap-dataset-using-svm-classifier/

Salcedo-Sanz S, Ghamisi P, Piles M, Werner M, Cuadra L, Moreno-Martínez A, ..., Camps-Valls G (2020). ML information fusion in Earth observation: A comprehensive review of methods, applications and data sources. Inform Fus 63:256-272

Shaik AB, Srinivasan S (2019) A brief survey on random forest ensembles in classification model. International Conference on Innovative Computing and Communications: Proceedings of ICICC 2018, vol. 2. Springer Singapore, Singapore, pp 253–260

Sheykhmousa M, Mahdianpari M, Ghanbari H, Mohammadimanesh F, Ghamisi P, Homayouni S (2020) Support vector machine versus random forest for remote sensing image classification: a meta-analysis and systematic review. IEEE J Select Top Appl Earth Obs Remote Sens 13:6308–6325

Shih HC, Stow DA, Tsai YH (2019) Guidance on and comparison of ML classifiers for Landsat-based land cover and land use mapping. Int J Remote Sens 40(4):1248–1274

Singh RP, Roy S, Kogan F (2003) Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. Int J Remote Sens 24(22):4393–4402. https://doi.org/10.1080/0143116031000084323

Sisodia PS, Tiwari V, Kumar A (2014) Analysis of supervised maximum likelihood classification for remote sensing image. International conference on recent advances and innovations in engineering. IEEE, Jaipur, pp 1–4

Suthaharan S (2016) ML Models and Algorithms for Big Data Classification, vol. 36. Springer US, Boston. https://doi.org/10.1007/978-1-4899-7641-3

Taati A, Sarmadian F, Mousavi A, Pour CTH, Shahir AHE (2015) Land use classification using support vector machine and maximum likelihood algorithms by Landsat 5 TM images. Walailak J Sci Technol (WJST) 12(8):681–687

Tariq A, Mumtaz F (2023) Modeling spatio-temporal assessment of land use land cover of Lahore and its impact on land surface temperature using multi-spectral remote sensing data. Environ Sci Pollut Res 30(9):23908–23924

Taylor IS (2024) Illegal Small-Scale Mining and Its Effects on the Economic Development of Ghana: A Dissertation. Available at: https://uniselinus.education/sites/default/files/2025-01/issah%20salifu%20taylor.pdf

Tian S, Zhang X, Tian J, Sun Q (2016) Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China. Remote Sens 8(11):954

Tschakert P, Singha K (2007) Contaminated identities: Mercury and marginalization in Ghana's artisanal mining sector. Geoforum 38(6):1304–1321. https://doi.org/10.1016/j.geoforum.2007.05.002

Valero Medina JA, Alzate Atehortúa BE (2019) Comparison of maximum likelihood, support vector machines, and random forest techniques in satellite images classification. Tecnura 23(59):3–10

Verma B, Singh P, Khamrai S, Prajapati B, Prasad R, Srivastava PK (2025) Drones in high resolution land use assessment using artificial intelligence. Earth Observation for Monitoring and Modeling Land Use. Elsevier, Varanasi, pp 83–104. https://doi.org/10.1016/B978-0-323-95193-7.00014-2

Wang W, Yang N, Zhang Y, Wang F, Cao T, Eklund P (2016) A review of road extraction from remote sensing images. J Traff Trans Eng (English Edition) 3(3):271–282

Wang J, Bretz M, Dewan MAA, Delavar MA (2022) ML in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. Science of the Total Environment 822:153559

Watuwaya BK, Syamsu JA, Budiman B, Useng D (2023) The role of remote sensing and GIS to support grassland identification. case study: East Sumba Regency, East Nusa Tenggara Province, Indonesia. In: AIP Conference Proceedings, vol. 2628. AIP Publishing, East Nusa Tenggara, p 1

Yimer SM, Bouanani A, Kumar N, Tischbein B, Borgemeister C (2024) Comparison of different machine-learning algorithms for land use land cover mapping in a heterogenous landscape over the Eastern Nile river basin, Ethiopia. Adv Space Res 74(5):2180–2199. https://doi.org/10.1016/j.asr.2024.06.010

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Clement Kwang[1] · Ian Kofi Afele[1] · Emmanuel Yeboah[2] · Isaac Sarfo[3,12] · Michael Batame[4] · Abraham Okrah[5] · Myint Myint Shwe[6] · Williams Siaw[7] · Dinah Boyetey[8] · Richard Odoi Larbi[9] · Augustine O. K. N. Mensah[10] · Charafa El Rhadiouini[2] · Ali Hasan Jaffry[2] · Fareeha Siddique[2] · Rukhshinda Aftab[2] · Oznur Isinkaralar[11]

✉ Emmanuel Yeboah
emmanuelyeboah@nuist.edu.cn

Clement Kwang
ckwang@ug.edu.gh

Ian Kofi Afele
iankofiafele@gmail.com

Isaac Sarfo
sarfo.power@gmail.com

Michael Batame
Michael.Batame@uga.edu

Abraham Okrah
okrahabraham2014@gmail.com

Myint Myint Shwe
myintmyintshwe66@gmail.com

Williams Siaw
202352250006@nuist.edu.cn

Dinah Boyetey
dinahboyetey@nuist.edu.cn

Richard Odoi Larbi
richardodoi29@gmail.com

Augustine O. K. N. Mensah
202451290002@nuist.edu.cn

Charafa El Rhadiouini
charafa.elrhadiouini@gmail.com

Ali Hasan Jaffry
alijafrey1122@gmail.com

Fareeha Siddique
fareehasiddique01@gmail.com

Rukhshinda Aftab
rukhshinda.aftab@gmail.com

Oznur Isinkaralar
obulan@kastamonu.edu.tr

1 Department of Geography and Resource Development, University of Ghana, Legon-Accra, Ghana

[2] School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[3] College of Geography and Environmental Science, Henan University, Kaifeng, 475004, Henan Province, China

[4] Warnell School of Forestry and Natural Resources, University of Georgia, Athens, USA

[5] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disaster, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[6] School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[7] School of Business, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[8] Department of Management Science Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[9] Departments of Mechanical Engineering, University of Mines and Technology, Tarkwa, Ghana

[10] Research Institute of History of Science and Technology, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China

[11] Department of Landscape Architecture, Faculty of Engineering and Architecture, Kastamonu University, 37150 Kastamonu, Türkiye

[12] Organization of African Academic Doctors (OAAD), Off Kamiti Road, P.O. Box 25305000100, Nairobi, Kenya