

# Multi-Sensor Machine Learning and Spatial-Statistical Analysis of Illegal Small- Scale Mining (Galamsey) Dynamics in Wassa Amenfi East District, Ghana, 2023–2025

William Bright Asomani

[williambrightaso@gmail.com](mailto:williambrightaso@gmail.com)

University of Mines and Technology

DOUGLAS OSEI kuffuor

University of Mines and Technology

---

## Research Article

**Keywords:** Galamsey, Sentinel-1, Sentinel-2, Machine learning, Change detection, Getis-Ord Gi\*, Spatial autocorrelation

**Posted Date:** August 13th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10651445/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Multi-Sensor Machine Learning and Spatial-Statistical Analysis of Illegal Small-Scale Mining (Galamsey) Dynamics in Wassa Amenfi East District, Ghana, 2023–2025**

**AUTHORS**

**WILLIAM BRIGHT ASOMANI**

*University of Mines and Technology – Department of Geomatic Engineering*

*williambrightaso@gmail.com*

*Accra, Ghana*

**DOUGLAS OSEI KUFFUOR**

*University of Mines and Technology – Department of Geomatic Engineering*

*Oseidouglas2000@gmail.com*

*Kumasi, Ghana*

## **Statements and Declarations**

### **Data Availability**

The Sentinel-1/Sentinel-2 image stacks, Google Earth Engine and R scripts, the 612-point reference dataset, and the classified/change-detection rasters generated and analysed during this study are available from the corresponding author on reasonable request.

### **Funding**

This research received no external funding.

### **Author Contributions**

*William Bright Asomani*

Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft,

*Douglas Osei Kuffuor*

Writing – review & editing, Visualization, Supervision.

### **Declaration of Competing Interest**

The authors declare no competing financial or personal interests that could have influenced the work reported in this paper.

### **Acknowledgments**

Acknowledge Google Earth Engine, ESA Copernicus Programme for Sentinel-1/Sentinel-2 data, WBA Consults, and my research team who assisted with fieldwork, digitisation and review.

## Abstract

Illegal small-scale gold mining (galamsey) continues to drive rapid environmental degradation in Ghana, yet operational multi-temporal monitoring remains limited. This study developed a multi-sensor machine learning framework integrating Sentinel-1 SAR and Sentinel-2 optical data to map and monitor galamsey in Wassa Amenfi East District between 2023 and 2025. Random Forest, XGBoost and Support Vector Machine classifiers were trained on 13 spectral and SAR features using 612 stratified reference points (70/30 split), achieving high accuracy (Overall Accuracy 0.945–0.962; Kappa 0.827–0.875) with no statistically significant difference between classifiers (McNemar's tests, all  $p > 0.05$ ). The red band (B4), VH backscatter and the green band (B3) were the most influential predictors. An independent, basemap-verified validation of 300 points found comparable Overall Accuracy (0.967) but lower detection completeness than the original test set implied (Producer's Accuracy 0.813 vs 0.939), indicating in-sample accuracies represent an upper bound on real-world performance. Detected mining area increased from 5,443 ha in 2023 to 17,614 ha in 2025 (+224%), driven mainly by newly established sites (13,805 ha). Spatial statistics confirmed very strong clustering (Moran's  $I = 0.567$ ,  $p < 0.001$ ; Geary's  $C = 0.433$ ,  $p < 0.001$ ), with extensive, statistically significant hotspots along river corridors at the 95% and 99% confidence levels and no coldspots detected. The framework offers a practical, low-cost basis for multi-year monitoring that quantifies expansion, separates new mining from persistence, and localises high-intensity hotspots to support targeted enforcement and environmental management in Ghana and similarly affected regions.

**Keywords** Galamsey; Sentinel-1; Sentinel-2; Machine learning; Change detection; Getis-Ord  $G_i^*$ ; Spatial autocorrelation

## 1. Introduction

Illegal small-scale gold mining, locally known as galamsey, remains one of the most pressing environmental and socio-economic challenges in Ghana. Although artisanal and small-scale mining (ASM) provides livelihoods for hundreds of thousands of people and contributes significantly to national gold production (Hilson & Maconachie, 2020), the unregulated component of the sector has caused widespread deforestation, destruction of agricultural land, siltation and chemical pollution of rivers, and long-term damage to ecosystems and community health (World Bank, 2025).

The Western Region, and particularly districts such as Wassa Amenfi East, has emerged as a major hotspot of galamsey. Monitoring the spatial extent and temporal dynamics of these activities is essential for effective enforcement and environmental restoration. However, the dispersed, small-scale and rapidly changing nature of galamsey sites makes conventional field-based monitoring expensive and often impractical.

Remote sensing offers a practical solution for large-area, repeatable monitoring. Previous studies in Ghana have applied Landsat and Sentinel-2 data to map artisanal mining and associated vegetation loss (Owusu-Nimo et al., 2018; Snapir et al., 2017; Baddianaah et al., 2023). Machine learning has further improved detection accuracy, including

Sentinel-1 time-series classification (Forkuor et al., 2020), deep-learning approaches with Sentinel-2 (Gallwey et al., 2020), and recent classifier-based assessments in this same district (Kwang et al., 2025). Despite these advances, several gaps remain: many studies rely solely on optical imagery limited by cloud cover; few systematically combine Sentinel-1 SAR with optical data; and detailed multi-temporal change analysis together with formal spatial clustering statistics are rarely integrated within the same framework.

This study addresses these gaps by developing a multi-sensor machine learning framework to map and monitor galamsey in Wassa Amenfi East District between 2023 and 2025. The specific aims are to: (1) evaluate Random Forest, XGBoost and Support Vector Machine classifiers using combined Sentinel-1 and Sentinel-2 features; (2) quantify the spatial extent and temporal expansion of galamsey; (3) analyse patterns of change (persistence, gain and loss); and (4) assess spatial clustering and identify statistically significant mining hotspots using global and local indicators of spatial association.

The main contribution lies in the integration of three elements rarely combined in existing Ghanaian research: a consistent multi-sensor classification pipeline across three recent years, quantitative change detection that separates new mining from persistence and abandonment, and formal spatial statistical analysis (Moran's I, Geary's C and Getis-Ord  $G_i^*$ ) that demonstrates the highly clustered nature of the expansion. This combination yields both scientifically interpretable results and practical information for targeted enforcement.

## **2. Study Area**

Wassa Amenfi East Municipal District is located in the Western Region of Ghana (Figure 1), approximately between latitudes 5°30' N and 6°15' N and longitudes 1°45' W and 2°11' W, in the north-central part of the region. The district was carved out of the former, larger Wassa Amenfi District in August 2004 and was elevated to Municipal Assembly status on 15 March 2018, with Wassa Akropong as its administrative capital. It shares boundaries with Wassa Amenfi Central District to the west, Upper Denkyira East District to the east, Prestea Huni-Valley Municipal to the south, and Upper Denkyira West District to the north (Wassa Amenfi East Municipal Assembly, 2023). Reported figures for the total land area range from approximately 1,558 km<sup>2</sup> to 1,729 km<sup>2</sup>, depending on the source and the boundary vintage used.

The district is underlain mainly by Birimian and Tarkwaian rocks of the Ghanaian Shield, a geological setting associated with extensive gold mineralisation across south-western Ghana; this endowment underlies both the licensed large- and small-scale mining sector and the widespread illegal galamsey activity that motivates this study. Relief is generally undulating, and the district is drained by a dense network of rivers and streams that form a dendritic pattern clearly visible in the classified maps (Figures 6–8); this drainage network is a key control on where alluvial and near-surface galamsey operations are sited, since operators preferentially work river gravels and floodplain deposits. Four forest reserves cover a combined area of approximately 212.6 km<sup>2</sup>, providing habitat for antelope, deer, elephants, birds and other wildlife and supplying medicinal plants used locally; these reserves are also frequent targets of illegal mining encroachment, making them a priority zone for the monitoring approach developed here.

The district experiences an equatorial climate with bimodal rainfall (approximately 1,300–2,000 mm annually) and lies within the moist semi-deciduous forest zone.

Agriculture (cocoa, oil palm, plantain, cassava and rice) remains the primary livelihood, while both licensed and illegal small-scale mining are widespread and increasingly compete with farmland and forest reserves for land. The population is predominantly Akan, with the Wassa as the indigenous ethnic group alongside substantial migrant farming and mining communities. The 2021 Population and Housing Census recorded a population of 179,696 (Ghana Statistical Service, 2021), more than double the 83,478 recorded in the 2010 census. This growth rate substantially exceeds Ghana's national natural population growth rate over the same period, consistent with significant in-migration into the district, plausibly linked in part to the employment opportunities, both licensed and illegal, created by the district's active gold mining sector. This rapid, migration-driven population growth adds further socio-economic pressure to the land-use change quantified in this study.

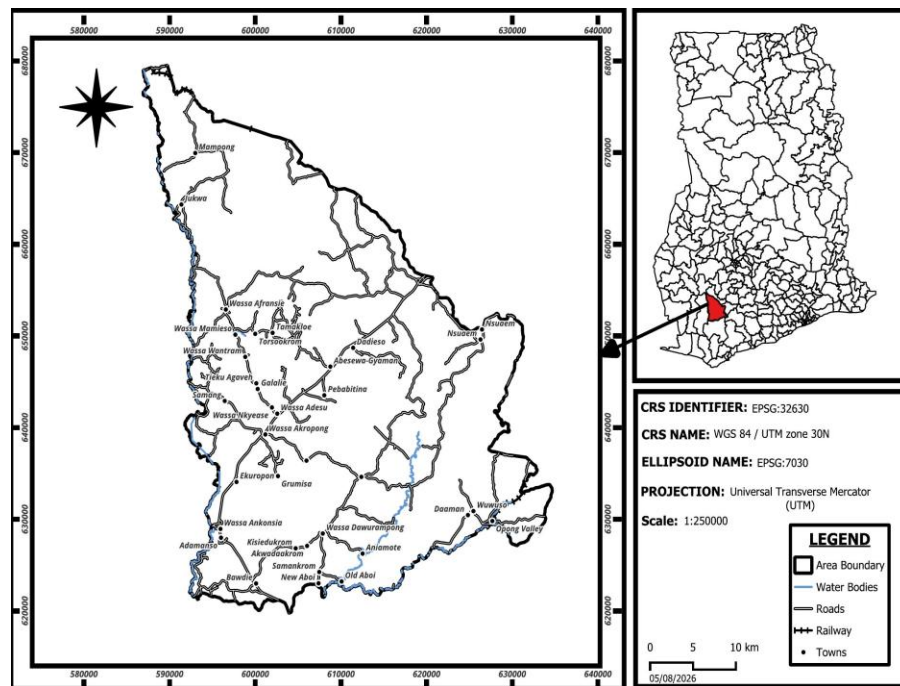


Figure 1. Location of Wassa Amenfi East District in the Western Region of Ghana, showing major towns, roads and water bodies.

### 3. Materials and Methods

#### 3.1 Satellite Data and Pre-processing

Multi-temporal Sentinel-1 SAR and Sentinel-2 MSI data were acquired for three dry-season periods: December 2022–March 2023, December 2023–March 2024, and December 2024–March 2025. Sentinel-2 Level-2A images were filtered to <30% cloud cover, cloud-masked using the QA60 band, and median-composited. Spectral indices (NDVI, NDWI, NBR, BSI) were calculated. Sentinel-1 GRD VV and VH data were median-composited and the VH/VV ratio derived. All layers were clipped to the district boundary and stacked into a 13-band multi-sensor image

for each year. Primary processing and the main Random Forest classification were performed in Google Earth Engine; subsequent analysis was conducted in R.

### **3.2 Reference Data and Sampling**

Reference data were obtained by manual digitisation of mining and non-mining polygons. An enlarged stratified random sample of 612 points (500 non-mining, 112 mining) was extracted from the multi-sensor stack guided by the 2023 classification, improving on the original set of 197 points. The dataset was partitioned into a training set (70%) and an independent test set (30%) using a fixed random seed (42). This district-specific reference set was built independently rather than drawn from existing published benchmarks such as the multi-site GhanaMining-Sat dataset (Danso, 2025), since no publicly available benchmark covered Wassa Amenfi East at the spatial and temporal resolution required for this study.

### **3.3 Machine Learning Classification and Hyperparameters**

Three classifiers were trained on the same 13-feature set: Random Forest (200 trees,  $mtry = 4$ , balanced sampsize), XGBoost ( $\eta = 0.08$ ,  $max\_depth = 6$ ,  $subsample = 0.8$ ,  $colsample\_bytree = 0.8$ ,  $scale\_pos\_weight$  adjusted for class imbalance, 200 rounds with early stopping), and Support Vector Machine (RBF kernel,  $cost = 10$ ,  $gamma = 0.1$ , class weights applied). Performance was evaluated using Overall Accuracy, Cohen's Kappa, Producer's Accuracy (Sensitivity/Recall), User's Accuracy (Precision) and F1-score derived from the confusion matrix. The primary multi-temporal maps were generated with Random Forest in Google Earth Engine.

### **3.4 Change Detection**

Annual classified maps (2023–2025) at 10 m resolution were masked to the district boundary. Change detection between 2023 and 2025 was performed by pixel-wise cross-tabulation, defining four transition classes: Persistent Mining, New Mining (Gain), Disappeared Mining (Loss), and Persistent Non-mining. Areas were calculated from pixel counts (100 m<sup>2</sup> per pixel) and converted to hectares.

### **3.5 Spatial Autocorrelation and Getis-Ord Gi\* Hotspot Analysis**

Global spatial autocorrelation of mining occurrence was quantified using Moran's I and Geary's C. Local clustering was assessed with the Getis-Ord Gi\* statistic. The 2025 binary mining map was aggregated to approximately 80–100 m resolution for computational feasibility. Spatial weights were constructed using k-nearest neighbours ( $k = 5–8$ ). The Gi\* statistic and associated z-scores were calculated for each grid cell. Locations with z-scores  $\geq 1.96$  and  $\geq 2.58$  were classified as hotspots at the 95% and 99% confidence levels, respectively; negative z-scores of the same magnitude were classified as coldspots. Kernel density estimation (focal sum) was additionally applied to visualise the intensity of new mining and classified into low, medium and high intensity zones.

## **4. Results**

## 4.1 Model Performance

All three models achieved high classification performance on the independent test set (Table 1). SVM and XGBoost produced the highest Overall Accuracy (0.962) and Kappa coefficients (0.875 and 0.872, respectively). Random Forest also performed strongly (OA = 0.945, Kappa = 0.827). Producer's Accuracy for the mining class exceeded 0.91 for all models, indicating reliable detection of galamsey sites. Pairwise McNemar's tests (continuity-corrected) showed no statistically significant difference in accuracy between any pair of classifiers: Random Forest vs XGBoost ( $\chi^2 = 0.800$ ,  $p = 0.371$ ), Random Forest vs SVM ( $\chi^2 = 0.571$ ,  $p = 0.450$ ), and XGBoost vs SVM ( $\chi^2 = 0.000$ ,  $p = 1.000$ ). The three classifiers should therefore be regarded as performing statistically comparably on this dataset, with XGBoost and SVM offering a small but not significant edge over Random Forest.

**Table 1. Performance of the classification models (n = 612; 70/30 split)**

| Model         | OA    | Kappa | Prod. Acc. | User Acc. | F1    |
|---------------|-------|-------|------------|-----------|-------|
| Random Forest | 0.945 | 0.827 | 0.939      | 0.795     | 0.861 |
| XGBoost       | 0.962 | 0.872 | 0.909      | 0.882     | 0.896 |
| SVM           | 0.962 | 0.875 | 0.939      | 0.861     | 0.899 |

Feature importance analysis (Figure 2) showed that the red band (B4), VH backscatter and the green band (B3) were the three most influential predictors. The strong contribution of B4 and B3 is consistent with the high visible reflectance of bare, excavated mining surfaces relative to vegetated land. VH backscatter ranked second, reflecting increased surface roughness associated with pits and spoil heaps. This confirms the complementary value of optical and SAR data, particularly under cloudy conditions.

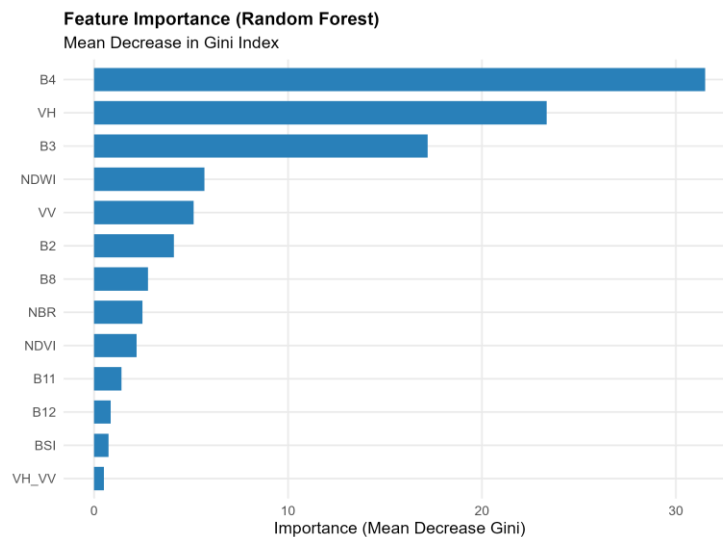


*Figure 2. Feature importance (Mean Decrease Gini) from the Random Forest model. B4, VH and B3 were the most influential predictors.*

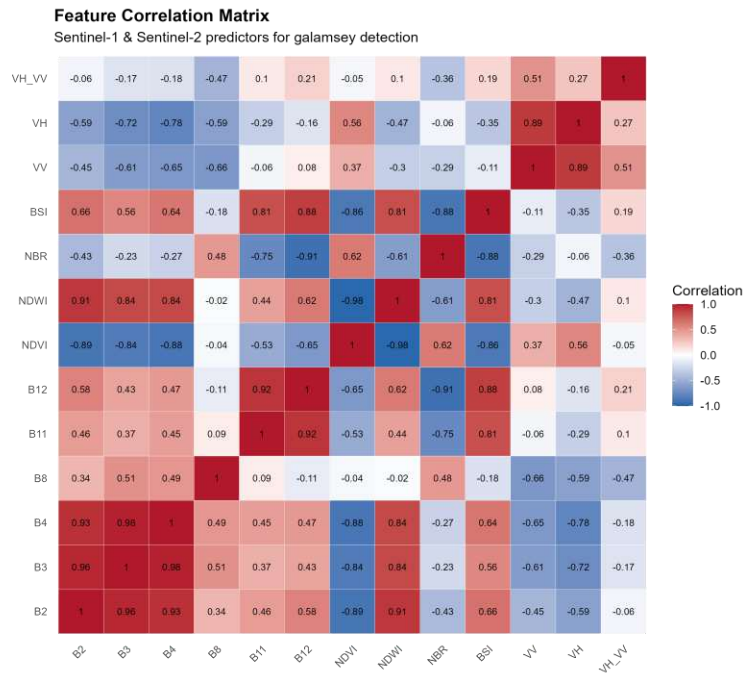


Figure 3. Correlation matrix of the 13 multi-sensor features used for galamsey detection.

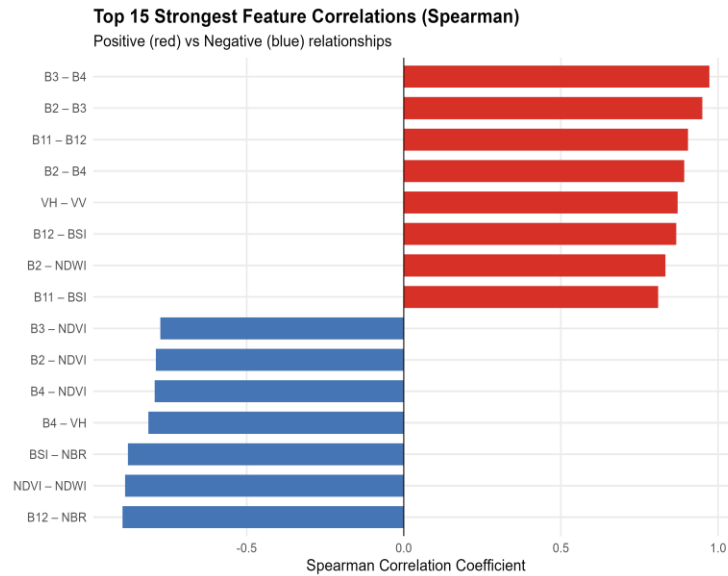


Figure 4. Top 15 strongest Spearman rank correlations among predictor variables.

The correlation structure among predictors (Figure 3; Figure 4) shows expected redundancy within each sensor: the optical bands B2–B4 are strongly inter-correlated, as are the vegetation indices derived from them (NDVI, NBR), while VV and VH backscatter form a separate, largely independent cluster. This confirms that the optical and SAR feature groups carry complementary rather than duplicate information, supporting the multi-sensor design, while also indicating some redundancy within each group that tree-based feature importance (Figure 2) implicitly accounts for by favouring a small subset of bands (B4, VH, B3).

## 4.2 Temporal Expansion of Galamsey

Detected mining area increased from 5,443 ha in 2023 to 17,614 ha in 2025, a net gain of 12,171 ha (224%) (Table 2; Figure 5). The largest annual increase occurred between 2023 and 2024 (+175%), followed by a slower but still positive rise in 2025 (+18%). This pattern suggests a pulse of rapid expansion followed by a partial consolidation phase.

Table 2. Detected galamsey area in Wassa Amenfi East District

| Year | Area (ha) | Change from previous year |
|------|-----------|---------------------------|
| 2023 | 5,443     | —                         |
| 2024 | 14,973    | +9,530 ha (+175.1%)       |
| 2025 | 17,614    | +2,641 ha (+17.6%)        |

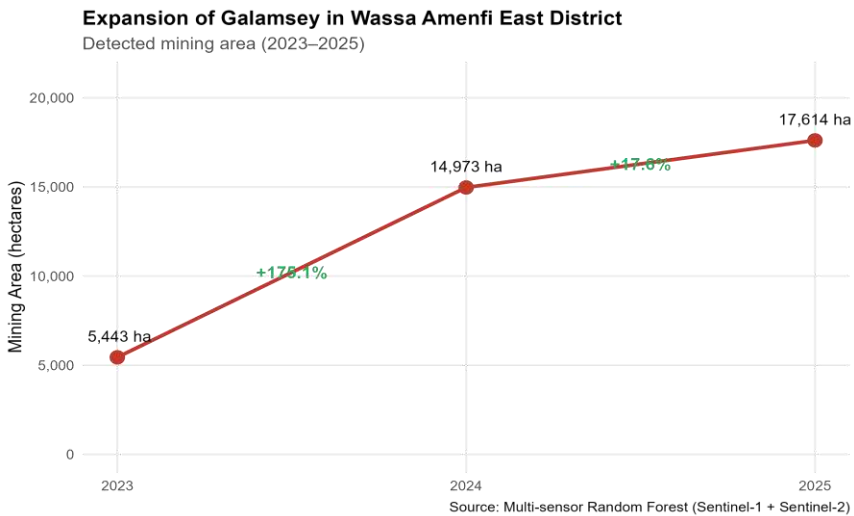


Figure 5. Expansion of detected galamsey area (2023–2025) with percentage increases between consecutive years.

4.3 Multi-temporal Classification Maps

The classification maps (Figures 6–8) reveal progressive expansion of mining (shown in red). In 2023, mining was relatively limited and fragmented, concentrated mainly along a few major drainage lines in the central and northern parts of the district. By 2024 the red signature had expanded markedly, filling previously non-mining corridors and forming more continuous linear belts. The 2025 map shows further densification and lateral spread, particularly in the north-west and along secondary tributaries. The progressive filling of riverine corridors is consistent with the expansion of alluvial and near-surface galamsey operations that follow gold-bearing sediments.

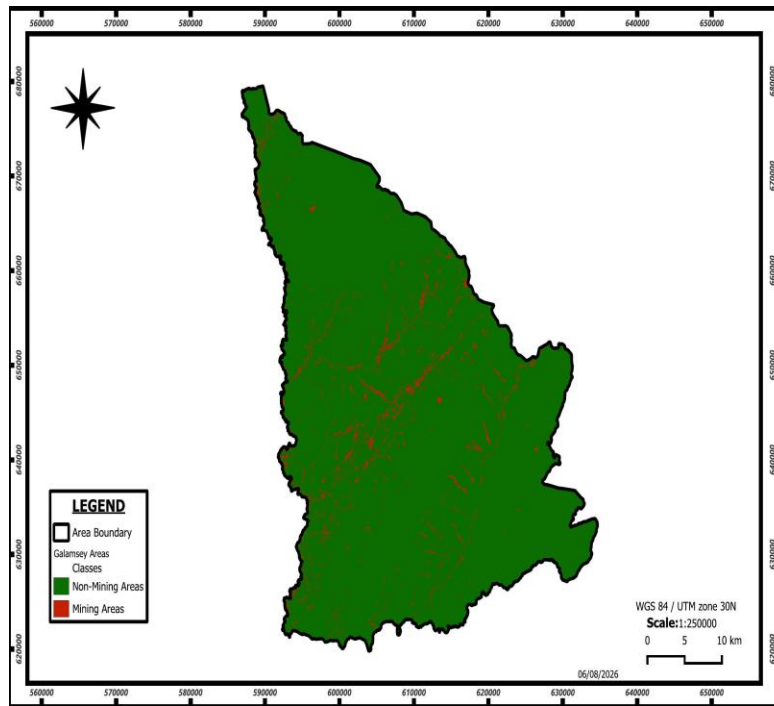


Figure 6. Classified galamsey map for 2023. Mining (red) is relatively limited and fragmented along major drainage lines.

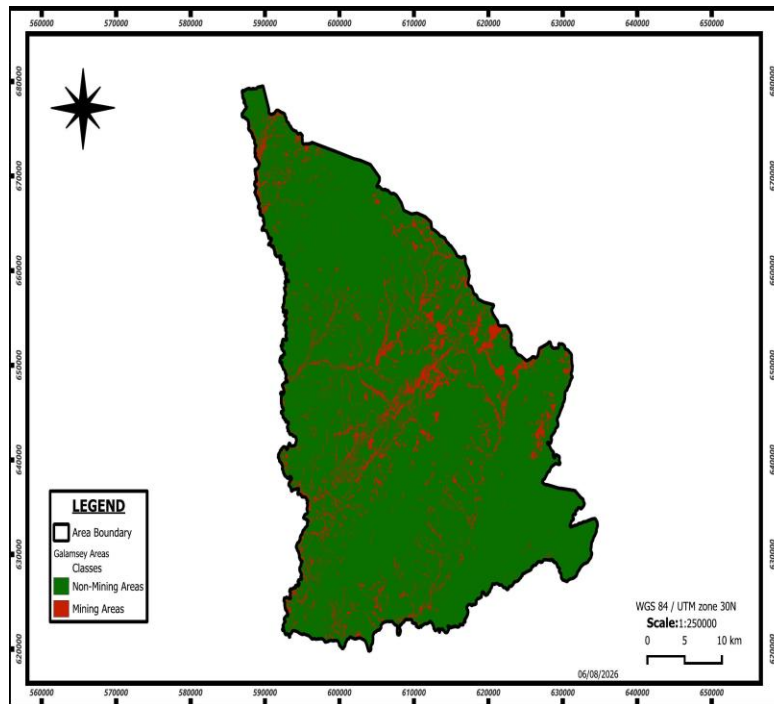


Figure 7. Classified galamsey map for 2024. Mining has expanded into continuous linear belts along river corridors.

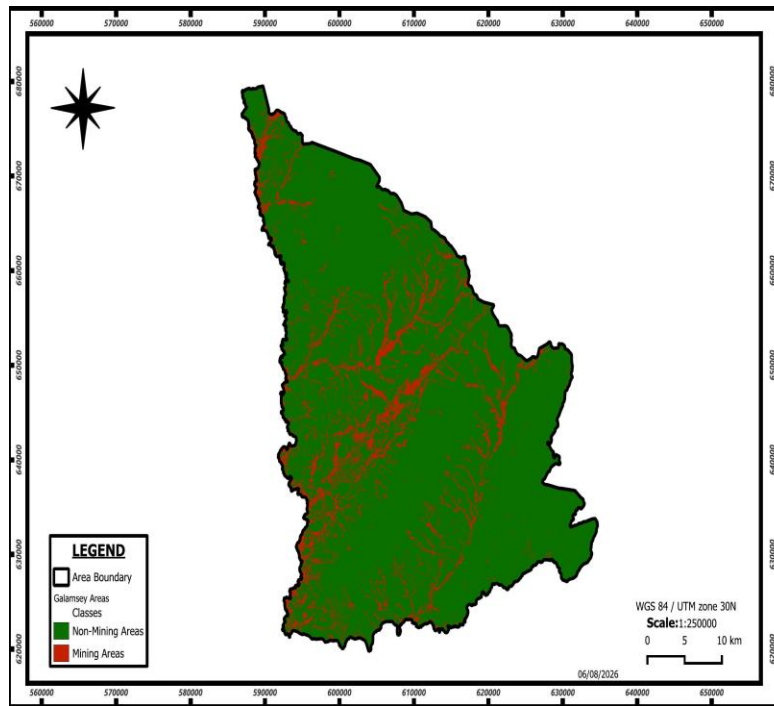


Figure 8. Classified galamsey map for 2025. Further densification and lateral spread are evident, especially in the north-west.

#### 4.4 Change Detection

Change detection between 2023 and 2025 (Table 3; Figure 9) shows that most expansion consisted of newly established mining sites (13,805 ha). Persistent mining occupied 4,072 ha, while 1,460 ha of former mining disappeared. The pixel-wise cross-tabulation gives a net change of 12,345 ha (New Mining minus Disappeared Mining), which is close to but not identical to the 12,171 ha difference between the independently derived 2023 and 2025 total-area estimates (Table 2); the 174 ha (1.4%) discrepancy reflects the different computational bases of the two methods (whole-map area totals versus pixel-level 2023-to-2025 transition counts) rather than a data error, and is well within the uncertainty expected at 10 m resolution. On the change map, new mining (dark red) forms clear dendritic patterns that closely follow the river network. Persistent mining occurs in compact clusters, mainly in the central sector. Disappeared mining (green) appears as small, scattered patches, often at the edges of earlier workings, suggesting limited abandonment or partial recovery rather than large-scale site closure. Overall, expansion is driven by the opening of new sites rather than the intensification of a few large existing operations.

Table 3. Change detection summary (2023–2025)

| Change Class              | Area (ha) |
|---------------------------|-----------|
| Persistent Mining         | 4,072     |
| New Mining (Gain)         | 13,805    |
| Disappeared Mining (Loss) | 1,460     |
| Persistent Non-mining     | 138,495   |

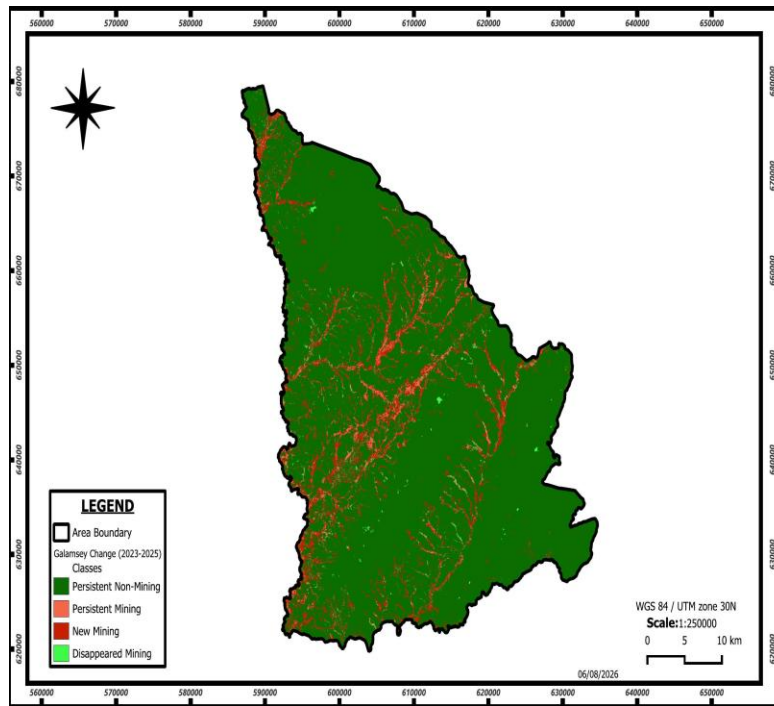


Figure 9. Change detection map (2023–2025). New mining (dark red) follows dendritic river patterns; persistent mining forms compact central clusters; disappeared mining (green) occurs as small marginal patches.

#### 4.5 Spatial Clustering and Getis-Ord $G_i^*$ Hotspots

Global spatial autocorrelation statistics confirmed that galamsey is strongly clustered. Moran's  $I$  was 0.567 ( $z = 459.6$ ,  $p < 0.001$ ) and Geary's  $C$  was 0.433 ( $z = 418.4$ ,  $p < 0.001$ ), both indicating highly significant positive spatial autocorrelation.

Getis-Ord  $G_i^*$  analysis identified extensive statistically significant hotspots: 19,917 grid cells at the 99% confidence level and 13,092 cells at the 95% confidence level. No cells reached the threshold for coldspots at any confidence level. High-intensity zones on the kernel density map (Figure 10) are concentrated in the north-western and central parts of the district, coinciding with areas of dense river branching. Medium-intensity zones form elongated corridors linking these cores. Low-intensity areas occupy most of the remaining landscape. The absence of widespread low-level scattering supports the statistical finding of strong spatial clustering: galamsey is organised into discrete high-density hotspots rather than being diffusely distributed across the district.

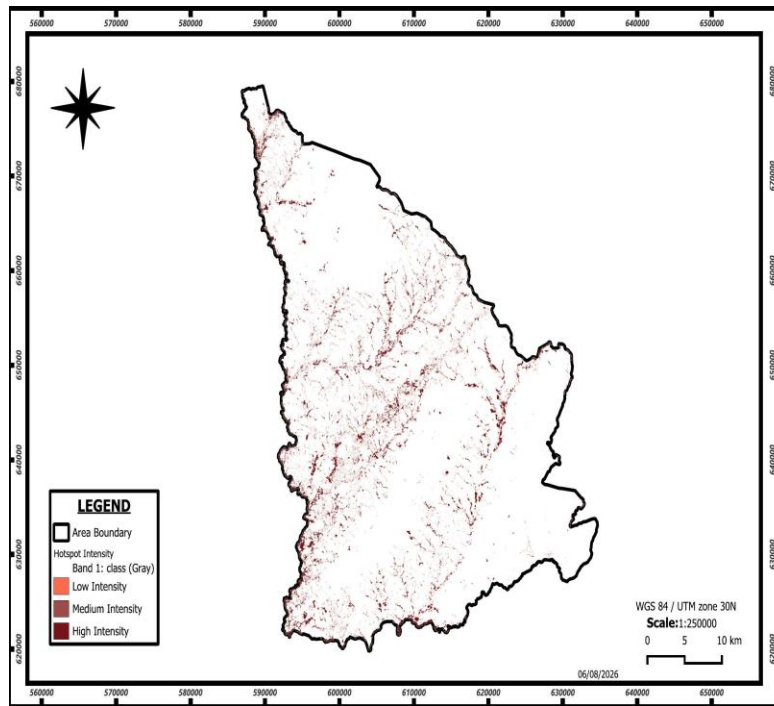


Figure 10. Hotspot intensity map of new mining. High-intensity zones (dark) concentrate in the north-west and centre along river corridors; no coldspots were detected.

#### 4.6 Independent Accuracy Validation

To address the risk that the original reference sample was partly guided by the classified maps, an independent validation was carried out. A fresh set of 300 points was drawn by simple random sampling over the district boundary (seed = 2025, distinct from the training/test split), with no reference to the classification output. Each point was labelled blind, as mining or non-mining, using high-resolution basemap imagery (Google/Bing/Esri) rather than the Sentinel-1/Sentinel-2 composites used to train the classifier, and compared against the predicted class from the primary Random Forest map (Table 4).

**Table 4. Comparison of the original classifier-guided test set and the independent, basemap-verified validation set (Random Forest map).**

| Validation Set                                            | N   | OA    | Kappa | Prod. Acc. | User Acc. | F1    |
|-----------------------------------------------------------|-----|-------|-------|------------|-----------|-------|
| Original test set (n = 184, classifier-guided reference)  | 184 | 0.945 | 0.827 | 0.939      | 0.795     | 0.861 |
| Independent validation (n = 300, basemap-verified, blind) | 300 | 0.967 | 0.867 | 0.813      | 0.975     | 0.886 |

Overall Accuracy on the independent set (0.967, 95% CI 0.940–0.984) was comparable to, and marginally higher than, the original test-set estimate (0.945), and Kappa was similarly high (0.867). However, Producer’s Accuracy (Sensitivity) for the mining class fell from 0.939 on the original test set to 0.813 on the independent set, while User’s Accuracy (Precision) rose from 0.795 to 0.975. In other words, when evaluated against labels that could not have been influenced by the classifier, the map under-detects a larger share of true mining sites (omission error) than the

original, partly circular validation suggested, while rarely flagging non-mining areas as mining (very few false positives). This pattern is consistent with the possibility that the original reference sample, which was partly guided by the classified maps, inflated the apparent completeness of mining detection. The independent estimates in Table 4 should therefore be treated as the more conservative and credible measure of real-world map performance, and are used in preference to the original test-set figures when discussing operational reliability in Section 5.

## 5. Discussion

The results of this study document a rapid and spatially structured expansion of illegal small-scale mining in Wassa Amenfi East District between 2023 and 2025. Detected mining area increased from 5,443 ha to 17,614 ha, a net gain of 12,171 ha or approximately 224%. Change detection reveals that the large majority of this increase (13,805 ha) consisted of newly established mining sites, while only 1,460 ha of former mining was no longer classified as active by 2025. Persistent mining occupied 4,072 ha. These figures indicate that the dominant process has been the opening of new sites rather than the continuous intensification of a limited number of existing operations.

The spatial pattern of expansion is strongly non-random. Mining signatures on the multi-temporal classification maps and the change detection map follow dendritic, river-aligned corridors. This geometry is consistent with the prevalence of alluvial and near-surface gold deposits that are exploited by galamsey operators along drainage networks. The 2023 map shows relatively limited and fragmented mining; by 2024 continuous linear belts have formed; and by 2025 further densification and lateral spread are evident, especially in the north-west and along secondary tributaries.

The pulse of expansion, dominated by the +175% area increase between 2023 and 2024 followed by a much smaller +18% rise to 2025, plausibly reflects the combined influence of gold-price dynamics and shifting policy enforcement, although these variables were not formally incorporated into the analysis and the following interpretation should be read as contextual rather than causally tested. Global gold prices rose from an annual average of approximately US\$1,941/oz in 2023 to US\$2,386/oz in 2024 (+23%) and further to US\$3,445/oz in 2025 (+44%) (MetalCharts.org, 2024; GoldTickers.com, 2025), directly increasing the profitability of galamsey and the incentive for new entrants. In Ghana specifically, small-scale miners produced more gold in the first seven months of 2024 alone than in the whole of 2023, and gold export earnings from the sector approached US\$20 billion in 2025 (Ghana Citizenship, 2026, citing Ghana Gold Board and Semafor data). This price signal alone would predict continuous acceleration through 2025, yet the observed area expansion instead decelerated sharply after 2024. This decoupling is consistent with a parallel tightening of enforcement: Ghana's Environmental Protection (Mining in Forest Reserves) Regulations, 2023 (L.I. 2462) had permitted mining in previously protected forest reserves, plausibly contributing to the 2023–2024 surge, before mounting public pressure through 2024 led the government to order its revocation, deploy additional military and police personnel to mining sites and water bodies, revoke small-scale mining licences issued after 7 December 2024, and retake control of several forest reserves during 2025 (Graphic Online, 2024a, 2024b; CitiNewsroom, 2024; AllAfrica, 2025). The timing of this enforcement escalation, concentrated in the final quarter

of 2024, corresponds closely with the point at which area expansion in this study slowed markedly, suggesting that regulatory and enforcement interventions may have partially offset an even stronger underlying economic incentive to expand. Formally testing this relationship, for example by regressing monthly or quarterly change-detection area against gold-price series and a categorical indicator of enforcement intensity, was beyond the scope of the present remote-sensing framework but is a natural extension for future work.

Getis-Ord  $G_i^*$  analysis confirmed that galamsey activity is highly localised. A large number of grid cells were classified as hotspots at the 95% and 99% confidence levels, while no cells reached the threshold for coldspots. The hotspots are predominantly aligned with the major river networks and their tributaries, particularly in the north-western and central portions of the district. This pattern is consistent with the operational characteristics of alluvial and near-surface gold extraction. The absence of coldspots indicates that mining is not merely absent from large parts of the district by chance; rather, activity is actively concentrated into a limited number of high-density corridors. Global statistics reinforce this conclusion: Moran's  $I$  of 0.567 ( $z = 459.6$ ,  $p < 0.001$ ) and Geary's  $C$  of 0.433 ( $z = 418.4$ ,  $p < 0.001$ ) both indicate strong positive clustering.

The strong spatial clustering of galamsey revealed by Moran's  $I$ , Geary's  $C$  and Getis-Ord  $G_i^*$  has direct and actionable implications for Ghanaian policy and operational practice. Because mining is concentrated in statistically significant hotspot corridors rather than dispersed evenly across the district, enforcement, monitoring and restoration resources can be prioritised with greater efficiency. Targeting the 95% and 99%  $G_i^*$  hotspot zones—particularly the north-western and central river-aligned clusters—is likely to capture a substantial share of current and emerging illegal activity while reducing the cost and logistical burden of uniform district-wide patrols.

The multi-temporal classification maps and annual area statistics (5,443 ha in 2023 to 17,614 ha in 2025) provide a quantitative baseline against which the effectiveness of ongoing interventions can be measured. These include the National Mine Surveillance Centre established in 2023, the expanded deployment of drones for aerial monitoring, the registration and geo-fencing of heavy mining equipment, and successive taskforce operations aimed at removing excavators and disrupting illegal sites. Regular updating of the Sentinel-1/Sentinel-2 classification workflow offers a low-cost, repeatable means of detecting new illegal sites early—often before they expand into continuous corridors—and of assessing whether interventions are succeeding in slowing or reversing net area growth.

The change detection results further refine operational priorities. The dominance of new mining (13,805 ha) over persistent mining (4,072 ha) and the relatively small area of disappearance (1,460 ha) indicate that the principal challenge is the continuous opening of fresh sites rather than the re-occupation of already known locations. Monitoring systems should therefore emphasise early detection of new clearings and pits along river networks, complementary to the surveillance of established hotspots. The limited extent of disappeared mining also suggests that spontaneous recovery is currently insufficient; active restoration and land-reclamation programmes will be required if abandoned sites are to return to productive or ecological use.

At the institutional level, the framework can support coordination among the Minerals Commission, the Environmental Protection Agency, district assemblies, the security agencies and traditional authorities. Spatially explicit hotspot maps can be shared as common operational layers, reducing duplication of effort and enabling joint prioritisation of high-risk corridors. Integration of the remote-sensing outputs with existing equipment-tracking and geo-fencing systems would further strengthen the ability to distinguish licensed from unlicensed activity in near real time.

In the longer term, sustained multi-year monitoring of the type demonstrated here can inform higher-level policy evaluation. Trends in net mining area, the balance between gain and loss, and the stability or shift of hotspot locations provide objective indicators of whether regulatory reforms, formalisation initiatives and community-level interventions are producing measurable spatial outcomes. In this respect, the multi-sensor machine-learning and spatial-statistical approach moves beyond one-off mapping exercises toward a practical decision-support system for the ongoing management of galamsey in Wassa Amenfi East and, by extension, in other affected districts of Ghana.

The multi-sensor machine learning approach proved effective. On an independent test set drawn from a stratified sample of 612 points, Random Forest, XGBoost and Support Vector Machine all achieved high overall accuracies (0.945–0.962) and substantial Kappa coefficients (0.827–0.875). Producer's accuracy for the mining class exceeded 0.91 in all cases, indicating reliable detection of galamsey sites. Feature importance analysis showed that the red band (B4), VH backscatter and the green band (B3) were the most influential predictors. The prominence of B4 and B3 is physically consistent with the high visible reflectance of bare excavated surfaces relative to surrounding vegetation. The high ranking of VH backscatter reflects the increased surface roughness associated with pits, spoil heaps and disturbed ground, and underscores the complementary value of Sentinel-1 data under the persistent cloud cover typical of southern Ghana.

Comparison with previous research supports the observed patterns while highlighting the added value of the present framework. Snapir et al. (2017) and Forkuor et al. (2020) documented progressive conversion of forest and agricultural land to mining-related surfaces in Ghana using optical and SAR data, respectively. Gallwey et al. (2020) demonstrated the potential of deep learning for ASM detection with Sentinel-2. Kwang et al. (2025), working in the same district, reported substantial vegetation loss attributable to galamsey using machine learning classifiers. The present study extends these efforts by providing consistent three-year area statistics, an explicit separation of gain, persistence and loss, and formal spatial clustering metrics that quantify the degree of concentration of mining activity.

Several sources of uncertainty should be acknowledged. Residual cloud contamination and cloud-shadow effects may introduce local errors despite masking. Mixed pixels at the 10 m resolution of Sentinel-2 are common along the edges of small pits and narrow riverine corridors and can produce both omission and commission errors. Seasonal differences in vegetation phenology and soil moisture between the three dry-season composites may affect spectral signatures. SAR data remain subject to residual speckle, although median compositing reduces this effect. In addition, because part of the enlarged reference sample was guided by the classified maps, an independent, basemap-verified

validation (Section 4.6) was carried out to check for optimistic bias; it confirmed that Producer's Accuracy for the mining class was overestimated by the original test set (0.939 vs 0.813 independently), while User's Accuracy was, if anything, underestimated (0.795 vs 0.975). Reported test-set accuracies should therefore be read as an upper bound on detection completeness, with the independent estimates in Table 4 offering a more conservative and operationally realistic benchmark. The Gi\* results are also sensitive to the choice of neighbourhood size and the level of spatial aggregation. Pairwise McNemar's tests confirmed that the differences in Overall Accuracy and Kappa between SVM, XGBoost and Random Forest are not statistically significant (all  $p > 0.05$ ; Section 4.1), so the models are best interpreted as performing comparably rather than one being definitively superior. Finally, while Section 5 discusses gold-price trends and documented enforcement actions as plausible contextual drivers of the 2023–2024 expansion pulse and its subsequent slowdown, these socio-economic and policy variables were not formally incorporated into the spatial or statistical analysis; a quantitative time-series model linking classified area change to gold-price series and enforcement records would provide a more rigorous test of these relationships.

Despite these limitations, the study contributes a coherent operational monitoring framework that integrates multi-sensor classification, multi-temporal change detection and spatial statistical analysis. The resulting maps and statistics can support more targeted enforcement, restoration planning and policy evaluation by the Minerals Commission, the Environmental Protection Agency, district assemblies and security agencies. Regular updating of the Sentinel-1/Sentinel-2 classification workflow offers a practical, low-cost means of tracking both the spread of new illegal mining and the effectiveness of subsequent interventions.

Future research should prioritise the collection of independent field validation data, the testing of deep-learning architectures on larger and more diverse training sets, and the coupling of remote-sensing outputs with socio-economic and governance variables. Extending the temporal window and expanding the analysis to neighbouring districts would further improve understanding of regional-scale galamsey dynamics.

## 6. Conclusions

This study applied a multi-sensor machine learning framework (Sentinel-1 + Sentinel-2) to map and monitor galamsey in Wassa Amenfi East District from 2023 to 2025. Random Forest, XGBoost and SVM all achieved high performance (OA 0.945–0.962) on a stratified sample of 612 points, with no statistically significant difference between classifiers (McNemar's tests, all  $p > 0.05$ ). An independent, basemap-verified validation of 300 points confirmed a comparable Overall Accuracy (0.967) but revealed lower completeness of mining detection than the original test set suggested (Producer's Accuracy 0.813 vs 0.939), underlining the value of independent validation for operational deployment. Mining area expanded by 224% (5,443 ha to 17,614 ha), driven mainly by new sites. Spatial analyses confirmed very strong clustering (Moran's  $I = 0.567$ ,  $p < 0.001$ ; Geary's  $C = 0.433$ ,  $p < 0.001$ ) and extensive statistically significant hotspots along river networks, with no coldspots detected. The framework provides timely, spatially explicit information that can support enforcement prioritisation and environmental management. Regular updating offers a practical means of tracking future changes in the extent and distribution of galamsey.

## References

- AllAfrica (2025). War on galamsey intensifies: 60 excavators seized; small-scale mining licences issued after December 7, 2024 revoked. AllAfrica.
- Baddianaah, I., Baatuuwie, B. N., & Adongo, R. (2023). Geospatial characterisation and distribution of illegal gold mining (galamsey) operations in the Upper West Region of Ghana. *GeoJournal*, 88, 789–810. <https://doi.org/10.1007/s10708-022-10642-3>
- CitiNewsroom (2024). Galamsey: Govt deploys over 100 soldiers to crack down on illegal mining. CitiNewsroom.com.
- Danso, E. H. (2025). GhanaMining-Sat: A multi-site benchmark dataset for artisanal mining detection in Ghana. Zenodo. <https://doi.org/10.5281/zenodo.17379308>
- Forkuor, G., Ullmann, T., & Griesbeck, M. (2020). Mapping and monitoring small-scale mining activities in Ghana using Sentinel-1 time series (2015–2019). *Remote Sensing*, 12(6), 911. <https://doi.org/10.3390/rs12060911>
- 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. <https://doi.org/10.1016/j.rse.2020.111970>
- Ghana Citizenship (2026). Galamsey in Ghana: History, dangers, and environmental damage. [ghanacitizenship.com](http://ghanacitizenship.com).
- Ghana Statistical Service (2021). 2021 Population and Housing Census: General Report. Accra: Ghana Statistical Service.
- GoldTickers.com (2025). Gold price in 2025: Historical chart & market data.
- Graphic Online (2024a). Galamsey crackdown: Govt rolls out radical measures; orders military deployment. Graphic Online.
- Graphic Online (2024b). We'll deal with galamsey ruthlessly: Ban, state of emergency, other options on table – Lands Minister. Graphic Online.
- Hilson, G., & Maconachie, R. (2020). Artisanal and small-scale mining and the Sustainable Development Goals: Opportunities and new directions for sub-Saharan Africa. *Geoforum*, 111, 125–141. <https://doi.org/10.1016/j.geoforum.2019.09.006>
- Kwang, C., Afele, I. K., Yeboah, E., Sarfo, I., Batame, M., Okrah, A., et al. (2025). Machine learning assessment of illegal mining (Galamsey) impacts on forest vegetation: a case study of Wassa Amenfi East District, Ghana. *Earth Science Informatics*, 18(2), 348. <https://doi.org/10.1007/s12145-025-01860-7>
- MetalCharts.org (2024). Gold price in 2024: \$2,386 average. High, low & chart.

Owusu-Nimo, F., Mantey, J., Nyarko, K. B., 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. <https://doi.org/10.1016/j.heliyon.2018.e00534>

Snapir, B., Simms, D. M., & Waine, T. W. (2017). Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 58, 225–233. <https://doi.org/10.1016/j.jag.2017.02.009>

Wassa Amenfi East Municipal Assembly (2023). District profile: location, geology, forest resources and demographics. Western Regional Co-ordinating Council.

World Bank (2025). Strategic Environmental and Social Assessment of Artisanal and Small-Scale Mining in Ghana. World Bank Group.