Essd-2025-583_Global Surface Mining and Land Reclamation of Time Series from 1985–2022

Dear Reviewer,

We area very grateful for your detailed comments and constructive suggestions on our manuscript (essd-2025-583). The suggestions have led to important improvements in methodology explanation, validation, and documentation. Below we provide point-by-point responses, with our replies in blue text. All revised text in the manuscript is highlighted in yellow. Please note that all line numbers refer to the latest revised version of the manuscript uploaded to the system.

Major Comment 1: Temporal Analysis Clarifications

1a. Baseline Year Consideration

Reviewer's Comment: *"Is the start time of each mine explicitly considered? Please clarify the baseline year used to define mining-induced disturbance for each mining polygon."*

Response: We greatly appreciate the reviewer's inquiry regarding the baseline year for mining disturbance. In our analysis, we adopt a pixel-level disturbance detection approach rather than defining site-specific baseline years for individual mining polygons. This methodological choice is driven by both practical and scientific considerations.

First, comprehensive data on the precise operational start times of mining activities are often unavailable or incomplete for many mining sites globally, particularly for historical operations and artisanal mines that predate systematic record-keeping. Attempting to define baseline years for each of the 74,726 mining polygons in our dataset would introduce substantial uncertainty and potentially erroneous assumptions.

Instead, our approach tracks land cover transitions at the pixel scale (30 m resolution) within delineated mining boundaries from 1985 to 2022. We identify disturbance when a pixel transitions from vegetated land cover (cropland, forest, grassland, shrubland, or sparse vegetation) to mine-related land cover (impervious surfaces, bare areas, water bodies, etc.). This first detected land-cover transition serves as the disturbance indicator for that specific pixel, independent of the overall mining site's operational timeline.

This pixel-based framework offers several advantages: (1) it captures the spatial heterogeneity of disturbance within individual mining polygons, as different areas within a single mine may be disturbed at different times; (2) it avoids the propagation of errors from potentially inaccurate site-level operational dates.

We acknowledge that this approach cannot capture disturbances occurring before 1985, which

represents a limitation of our study, as noted in Section 4.3 (Uncertainty and limitations). We have added clarifying text in Section 2.2 to make this methodological choice more explicit.

Revised Text (Section 2.2, Lines 231–238): "Rather than defining a site-specific baseline year for each mining polygon—which would be constrained by the general unavailability of comprehensive operational records for many global mining sites—we adopt a pixel-level detection approach. Disturbance is identified when a pixel first transitions from vegetated land cover to mine-related land cover within the study period (1985–2022). This first detected transition serves as the disturbance indicator, enabling fine-grained tracking of spatially and temporally heterogeneous mining impacts within individual polygons."

1b. Analysis Window Justification

Reviewer's Comment: *"Do all mines use the same analysis window (2018–2023) for trend assessment? What is the rationale for selecting this period?"*

Response: We thank the reviewer for this important methodological question. Yes, all mining polygons were analyzed using the same temporal window to ensure consistency and comparability across the global dataset. The selection of this period was guided by both statistical and practical considerations.

From a statistical perspective, the Mann-Kendall trend test requires a minimum of approximately four to five data points to achieve adequate statistical power for detecting monotonic trends. Our analysis window provides sufficient observations to satisfy this requirement while maintaining focus on recent mining dynamics.

From a practical perspective, the primary objective of the trend classification was to assess the current development status of mining areas—specifically, whether mines are actively expanding, undergoing reclamation, or in a stable state. A relatively short, recent window better captures the current operational phase of each mine. Extending the analysis window further into the past would potentially conflate historical patterns with present conditions, obscuring the distinction between mines that have recently transitioned to reclamation versus those that underwent reclamation many years ago but have since stabilized.

1c. Static Boundary Assumption

Reviewer's Comment: *"Does the method account for the fact that mining boundaries expand or contract over time? This assumption should be explicitly stated and evaluated."*

Response: We thank the reviewer for this insightful comment, which raises an important methodological consideration. Our current methodology does employ fixed mining boundaries,

and we appreciate the opportunity to clarify and evaluate this assumption explicitly.

In this study, we used mining boundaries derived from two existing global datasets (Maus et al., 2022; Tang and Werner, 2023), which were delineated based on satellite imagery from circa 2020. Through our refinement workflow (Section 2.1), we produced what we term the "maximum potential mining disturbance boundary"—the cumulative outer envelope of mining-induced land disturbance captured within the source datasets. These boundaries are indeed static and do not dynamically track spatial expansion or contraction over time.

The implications of this static boundary assumption are twofold:

(1) Mining activities that expanded beyond the delineated boundaries after the source imagery dates (or in areas not captured by the original datasets) would not be detected in our analysis. This represents a potential underestimation of total mining disturbance.

(2) Conversely, areas included within our boundaries that were never actually mined (e.g., buffer zones or future expansion areas delineated conservatively in source datasets) may contribute to overestimation. However, our stable green area exclusion procedure (Section 2.1, Step 2–3) was specifically designed to mitigate this issue by removing persistently vegetated areas from the boundaries.

We have added explicit discussion of this limitation in Discussion section to ensure transparency regarding this assumption.

Revised Text (Section 4.3, Lines 827–835): "A further methodological consideration is the use of static mining boundaries. Our analysis was conducted within the maximum union of two pre-existing global mining datasets, which provides comprehensive spatial coverage but cannot capture mining activities that expanded beyond these boundaries during the study period. It should be noted that these boundaries are static and derived from satellite imagery circa 2020. Newly developed mining areas not represented in the source datasets (Maus et al., 2022; Tang and Werner, 2023) are necessarily excluded from our analysis. Future work integrating time-series boundary delineation methods could address this limitation and provide a more complete picture of mining's evolving spatial footprint."

1d. NDVI Saturation

Reviewer's Comment: "NDVI saturation may lead to underestimation of recovery trends in medium-to-high biomass areas. What would change if alternative indicators such as NIRv were used?"

Response: We thank the reviewer for raising this important methodological consideration regarding vegetation index saturation. It is well established that NDVI tends to asymptotically

saturate in medium-to-high biomass environments (typically NDVI > 0.8), which can reduce sensitivity to vegetation dynamics in dense canopies such as tropical forests.

In this study, NDVI is not used as the sole indicator for disturbance or reclamation detection. The identification of mining disturbance and reclamation status primarily relies on the GLC_FCS30D land-cover classification product, which integrates multiple spectral bands, indices, and temporal features rather than NDVI alone. This multi-feature framework helps mitigate the influence of NDVI saturation on the initial classification of land-cover transitions. NDVI time series derived from MODIS 250-m products are subsequently employed for trend characterization of vegetation recovery. In this context, the reviewer's concern is valid, as NDVI-based trends may underestimate recovery trajectories in high-biomass regions. Alternative indicators such as NIRv, kNDVI, or EVI have been shown to retain greater sensitivity under high leaf area index conditions.

If NIRv were applied for trend assessment, vegetation recovery in densely vegetated reclamation sites—particularly within tropical biomes—would likely be detected more sensitively, potentially resulting in a higher proportion of sites classified as advanced or closed-canopy recovery. While a full NIRv-based reanalysis is beyond the scope of the present study, we acknowledge this limitation and highlight NIRv-based trend analysis as a valuable direction for future research.

Revised Text (Section 4.3, Lines 866–878): "NDVI saturation effects in medium-to-high biomass environments (typically where NDVI > 0.8) may reduce sensitivity to vegetation recovery trends, particularly in densely vegetated reclamation sites within tropical biomes. Although the GLC_FCS30D land-cover classification employed in this study integrates multiple spectral and temporal features that partially alleviate reliance on a single vegetation index, the MODIS-derived NDVI time series used for recovery trend analysis may still underestimate vegetation regrowth under closed-canopy conditions. Alternative indicators such as NIRv (near-infrared reflectance of vegetation), which maintains sensitivity at high leaf area index (Badgley et al., 2017, 2019), or nonlinear variants such as kNDVI (Camps-Valls et al., 2021), could provide more accurate characterization of recovery trajectories. Future regional-scale assessments, especially in tropical forest environments, should consider incorporating NIRv-based trend analysis to better capture advanced stages of ecological reclamation."

References to Add:

Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, Sci. Adv., 3, e1602244, https://doi.org/10.1126/sciadv.1602244, 2017.

Badgley, G., Anderegg, L. D. L., Berry, J. A., and Field, C. B.: Terrestrial gross primary production: Using NIRV to scale from site to globe, Global Change Biol., 25, 3731–3740, https://doi.org/10.1111/gcb.14729, 2019.

Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., Mahecha, M. D., Muñoz-Marí, J., García-Haro, F. J., Guanter, L., Jung, M., Gamon, J. A., Reichstein, M., and Running, S. W.: A unified vegetation index for quantifying the terrestrial biosphere, Sci. Adv., 7, eabc7447, https://doi.org/10.1126/sciadv.abc7447, 2021.

1e. High Proportion of Undefined Mines

Reviewer's Comment: "*Nearly half of the mining polygons are classified as 'undefined'. This proportion is unexpectedly large and suggests limitations in indicator sensitivity or threshold selection.*"

Response:

We thank the reviewer for raising this important concern. We have carefully considered whether the 48.9% proportion reflects methodological limitations, and our analysis suggests that this proportion actually represents an improvement over existing approaches rather than a deficiency.

We respectfully argue that our approach demonstrates enhanced rather than limited sensitivity. Wang et al. (2025), using NDVI change rates alone for a similar analysis period (2018–2022), classified 64.3% of mining polygons as "Stable." By integrating three complementary indicators (NDVI for vegetation dynamics, BSP for bare soil exposure, and NTL for human activity intensity), our method reduced this proportion to 48.9%—detecting dynamic trends in approximately 15% more mining areas than single-indicator approaches. This improvement confirms that combining multiple indicators captures mining activities that vegetation-based methods alone would miss, particularly in arid environments where bare soil dominates, or in regions where mining operations continue under partial vegetation cover.

The proportion of stable mines also reflects the inherent heterogeneity of the global mining landscape. Our dataset encompasses 74,726 polygons across 155 countries, spanning scales from small artisanal operations (averaging 0.72 km² in Asia) to mega-mines (exceeding 2.0 km² in Oceania). Many operations worldwide genuinely maintain steady-state conditions—neither expanding nor contracting their surface footprint—due to mature production phases, care-and-maintenance status, or underground extraction with minimal surface expression. The 48.9% proportion thus represents a meaningful finding about global mining dynamics: approximately half of the world's mining areas currently operate at equilibrium.

Following this comment and the parallel concern from Reviewer 1, we have renamed the

category from "Undefined Mines" to "Stable Mines" and provided a positive definition specifying the operational states this category represents. Please see our detailed response to Reviewer 1 Comment 1 for the complete revised text, which includes modifications to Sections 2.3, 3.3, and 4.3.

Major Comment 2: Comparisons and Validations

Reviewer's Comment: *"The dataset is derived from merging two existing mining inventories, but the manuscript does not convincingly demonstrate accuracy improvements. The authors should conduct cross-dataset quantitative comparisons using the same validation samples or adding case-based evaluations."*

Response:

We thank the reviewer for this constructive suggestion. We acknowledge that the original manuscript lacked a systematic cross-dataset comparison to demonstrate the accuracy improvements of our refined boundaries. Following the reviewer's recommendation, we have designed and implemented a stratified cross-validation framework that directly compares the three datasets (Tang and Werner 2023, Maus et al. 2022, and our refined dataset) using the same validation samples.

We designed a four-zone stratified sampling scheme based on the spatial overlap relationships among the three datasets. The zones are defined as follows:

Zone A (Excluded by Ours): Areas identified as mining by both Tang and Maus datasets but excluded by our refined boundaries (n = 150 samples);

Zone B (Three-way Consensus): Areas identified as mining by all three datasets (n = 300 samples);

Zone C (Maus + Ours, not Tang): Areas identified by Maus and our dataset but not by Tang (n = 150 samples);

Zone D (Tang + Ours, not Maus): Areas identified by Tang and our dataset but not by Maus (n = 150 samples).

A total of 750 validation points were randomly sampled across these four zones. Each point was independently verified through visual interpretation of high-resolution Google Earth imagery in combination with Landsat 8/9 Collection 2 Level 2 surface reflectance data (July 2019 to June 2021). The validation results are summarized in Table R1.

Table R1. Cross-dataset validation results

Zone	Mining	Non-mining	Total	Mining Rate

Zone A (Excluded by Ours)	43	107	150	28.7%
Zone B (Three-way Consensus)	274	26	300	91.3%
Zone C (Maus + Ours, not Tang)	122	28	150	81.3%
Zone D (Tang + Ours, not Maus)	129	21	150	86.0%

These results provide important quantitative evidence for the accuracy improvements achieved by our boundary refinement approach:

(1) The low mining rate in Zone A (28.7%) demonstrates that the areas excluded by our refinement process were predominantly non-mining areas. This means 71.3% of the excluded areas were correctly identified as non-mining, validating the effectiveness of our CCDC-based stable vegetation exclusion approach in reducing commission errors (false positives) from the merged dataset.

(2) The high mining rate in Zone B (91.3%) confirms the reliability of consensus areas where all three datasets agree, providing a solid foundation for subsequent time-series analysis.

(3) Zones C and D show mining rates of 81.3% and 86.0% respectively, indicating that our dataset captures genuine mining areas that were missed by one of the source datasets. This demonstrates improved coverage while maintaining acceptable accuracy.

(4) Compared to directly merging the two source datasets without refinement, our approach reduces the inclusion of non-mining areas (as evidenced by Zone A results) while preserving high-confidence mining areas identified through dataset intersection.

We have revised the manuscript accordingly: the validation methodology is now described in Section 2.4.2, and the validation results along with their interpretation are presented in Section 4.1. This separation follows the standard structure of scientific papers, with methods in the Methods section and results/discussion in the Discussion section.

Revised text:

(Section 2.4.2, Lines 320–360):

2.4.2 Validation

To comprehensively evaluate the performance of the proposed method, we designed a dual-track validation framework addressing both spatial accuracy and temporal accuracy. The spatial validation focuses on assessing the reliability of mining area delineation by comparing

our refined dataset with two existing global mining datasets through stratified cross-validation across four spatially defined zones. The temporal validation targets the accuracy of mining-reclamation transition year detection through multi-index spectral trajectory analysis over a 25-year period (1990–2022). Both validation components employed stratified random sampling and integrated high-resolution Google Earth imagery with Landsat/Sentinel-derived spectral indices for reference label generation.

To quantitatively assess the accuracy improvements of the refined dataset compared to the source datasets, we designed a stratified cross-validation framework based on the spatial overlap relationships among the three datasets: Dataset A (Tang and Werner, 2023), Dataset B (Maus et al., 2022), and our refined Dataset D. Four validation zones were defined based on dataset agreement patterns (Fig. 3):

(1) Zone A: Areas identified by both Tang and Maus but excluded in our refined dataset, representing potentially over-estimated mining extents in the source datasets;
(2) Zone B: Three-way consensus areas identified by all three datasets;
(3) Zone C: Areas identified by Maus and our dataset but not by Tang;
(4) Zone D: Areas identified by Tang and our dataset but not by Maus.

A total of 750 validation points (Fig. A1):were randomly sampled across these zones using stratified random sampling (150 points each for Zones A, C, and D; 300 points for Zone B to ensure adequate representation of consensus areas). Each point was verified through visual interpretation of high-resolution Google Earth imagery, supplemented by Landsat 8/9 Collection 2 Level 2 surface reflectance imagery (July 2019 to June 2021). Cloud-free median composites were generated using true-color visualization (bands B4, B3, B2), and the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Built-up Index (NDBI) were calculated to assist in distinguishing mining areas from vegetated or built-up land covers.

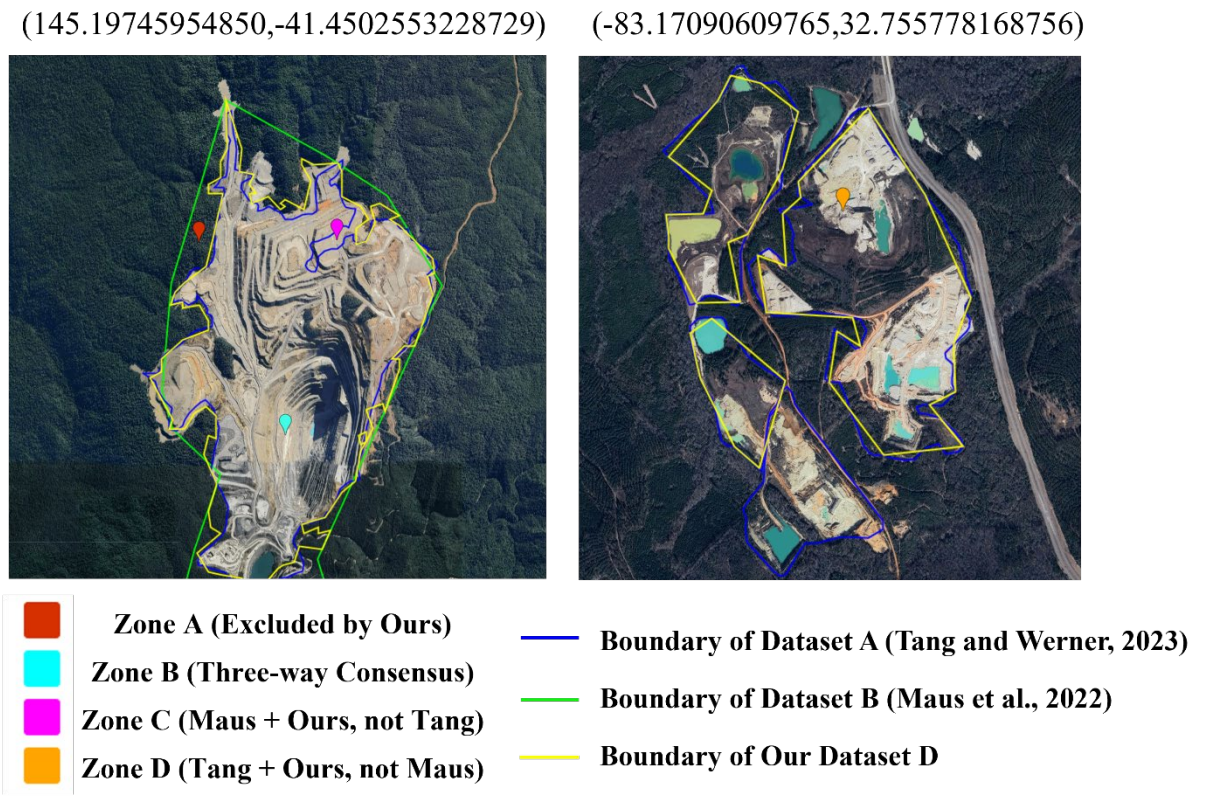
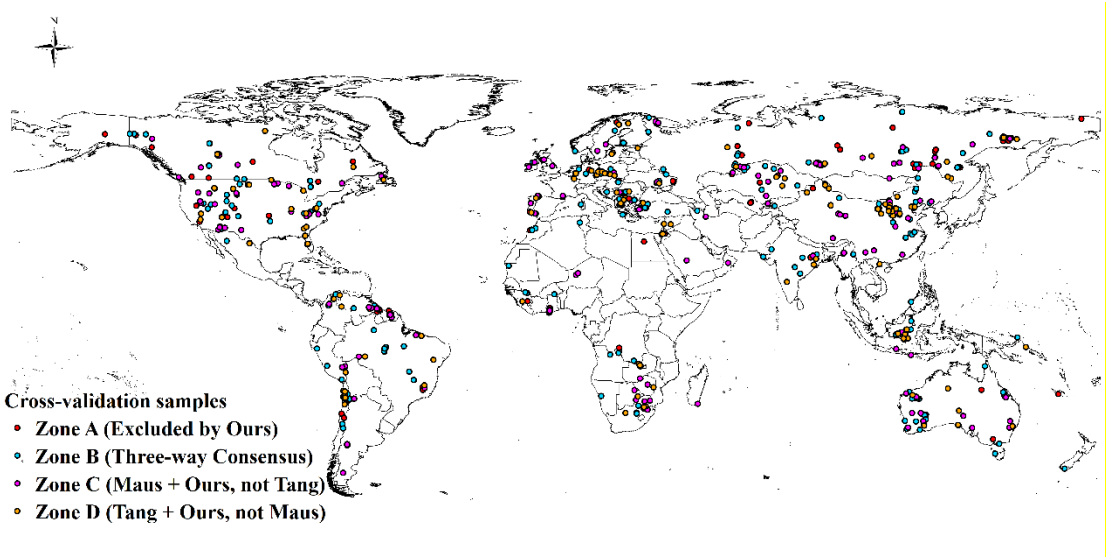


Figure 3. Illustration of the four validation zones defined by spatial overlap relationships among three mining datasets at two representative sites. Left: a mining site in Tasmania, Australia (145.197°E, 41.450°S); Right: a mining site in Georgia, USA (83.171°W, 32.756°N). The boundary lines represent mining area delineations from Dataset A (Tang and Werner, 2023; blue), Dataset B (Maus et al., 2022; green), and our refined Dataset D (yellow). The filled zones indicate: Zone A (red) – areas identified by both source datasets but excluded in our refined dataset; Zone B (cyan) – three-way consensus areas identified by all three datasets; Zone C (magenta) – areas identified by Dataset B and our dataset but not by Dataset A; Zone D (orange) – areas identified by Dataset A and our dataset but not by Dataset B. Background imagery: Google Earth.

(Appendix A, Lines 952–953):

Figure. A1 Global spatial distribution of 750 cross-validation samples



(Section 4.1, Lines 752–776):

The cross-dataset validation results provide quantitative evidence for the accuracy improvements achieved by our boundary refinement approach (Table 2). Zone A, representing areas identified by both Tang and Maus but excluded in our refined dataset, exhibited a mining rate of only 28.7% (43 out of 150 samples). This indicates that 71.3% of the areas excluded by our refinement process were correctly identified as non-mining land covers, validating the effectiveness of our CCDC-based stable vegetation exclusion strategy in reducing commission errors inherited from the merged dataset. Zone B, the three-way consensus area, showed a mining rate of 91.3% (274 out of 300 samples), confirming the high reliability of areas where all three datasets agree. This high accuracy in consensus areas provides a solid foundation for the subsequent time-series analysis of mining disturbance and reclamation dynamics. Zones C and D demonstrated mining rates of 81.3% (122/150) and 86.0% (129/150), respectively. These results suggest that our integrated dataset successfully captures additional mining areas that were missed by individual source datasets while maintaining reasonable accuracy. The slightly higher mining rate in Zone D compared to Zone C may reflect differences in the mapping strategies employed by Tang and Werner (2023) and Maus et al. (2022).

Collectively, these validation results demonstrate that our boundary refinement approach achieves two key objectives: (1) effectively reducing commission errors (false positives) from the merged dataset, as evidenced by the substantial proportion of non-mining areas in Zone A (71.3%); and (2) preserving genuine mining areas with high confidence, as indicated by the consistently high mining rates in Zones B, C, and D (81.3%–91.3%). This balance between precision and coverage represents a meaningful improvement over using either source dataset alone or a simple union of both.

Table 2. Cross-dataset validation results for boundary refinement assessment

Zone	Definition	Mining	Non-mining	Total	Mining Rate (%)
Zone A	Tang + Maus, excluded by Ours	43	107	150	28.7
Zone B	Three-way consensus	274	26	300	91.3
Zone C	Maus + Ours, not Tang	122	28	150	81.3
Zone D	Tang + Ours, not Maus	129	21	150	86.0

Minor Comments

Reviewer's Comment:

1. *Line 306-312： These sentences should be moved to the relevant figure caption. Please check for similar cases throughout the manuscript.*
2. *Figure 3(a): The lower chart duplicates information already shown in Figure 3(b). Please remove it to reduce visual clutter.*
3. *Figure 3(b), upper plot: The purpose for scaling the y-axis by 1,000 is unclear and seems unnecessary.*
4. *Figure 3(b), lower images: The images are too blurry. Please regenerate them using downloaded satellite imagery rather than screenshots from Google Earth.*
5. *Percentages in Sections 3.2 and 3.3 are inconsistently formatted (some integers, some with one decimal place). Please harmonize formatting across the manuscript.*
6. *Please provide a dataset user guide / metadata documentation, explaining the meaning, units, and calculation logic of each field in the dataset. This will significantly improve usability.*

Response:

We sincerely thank the reviewer for these constructive and detailed minor comments. We have carefully revised the manuscript accordingly. Our responses are provided point by point below.

1. **Lines 306–312:** We have relocated the methodological details to the caption of Figure 3 and carefully reviewed the entire manuscript for similar instances, making corresponding revisions where appropriate.
2. **Figure 3(a) redundancy:** The redundant lower chart has been removed from Figure 3(a) to improve visual clarity.

3. **Figure 3(b) y-axis scaling:** We have revised the figure to display the original NDVI values on the y-axis without artificial scaling.
4. **Figure 3(b) image quality:**

We sincerely thank the reviewer for this constructive suggestion. We have completely revised Figure 4(b) by replacing the original screenshots with satellite imagery downloaded directly from Google Earth Engine (GEE) using the Landsat 8 Collection 2 Level-2 Surface Reflectance product (LANDSAT/LC08/C02/T1_L2).

For Point 688 (coordinates: 3.552883°N, 117.169372°E, located in Borneo, Indonesia), we generated annual median composites for 2014, 2015, and 2016 using cloud-masked imagery with less than 50% cloud cover. The images were processed with optimized visualization parameters (true color composite with min: 7000, max: 16000, gamma: 1.2) and exported at 10 m spatial resolution through 3× supersampling of the native 30 m Landsat data. This supersampling approach enhances visual presentation quality while preserving the spectral characteristics of the original observations.

We acknowledge that the imagery may still appear less sharp compared to very high-resolution commercial satellites. However, this reflects the inherent spatial resolution limitation of Landsat 8 OLI (30 m), which represents the finest freely available optical imagery with consistent global coverage for the 2014–2016 study period. Since our validation focuses on 30 m resolution land-cover products (GLC_FCS30D), the use of Landsat imagery at its native resolution ensures methodological consistency between the validation imagery and the dataset being validated.

Higher-resolution commercial imagery (e.g., WorldView, Pleiades, SPOT) for this specific remote tropical location and time period is either unavailable or requires costly licensing beyond the scope of this study. The Google Earth historical high-resolution imagery coverage for this region during 2014–2016 is also limited and inconsistent.

The revised Figure 4(b) now includes GeoTIFF imagery exported directly from GEE with the following specifications: (1) Landsat 8 Collection 2 Level-2 Surface Reflectance data; (2) Annual median composites with cloud masking; (3) True color visualization (RGB: Bands 4, 3, 2); (4) 10 m export resolution for enhanced visual quality; (5) 1 km × 1 km region of interest centered on Point 688; and (6) A yellow marker clearly indicating the 30 m validation pixel location. These improvements ensure that the figure accurately represents the satellite observations while maintaining appropriate resolution for validating our 30 m land-cover dataset.

Revised Text (Section 2.4, Lines 382–392):

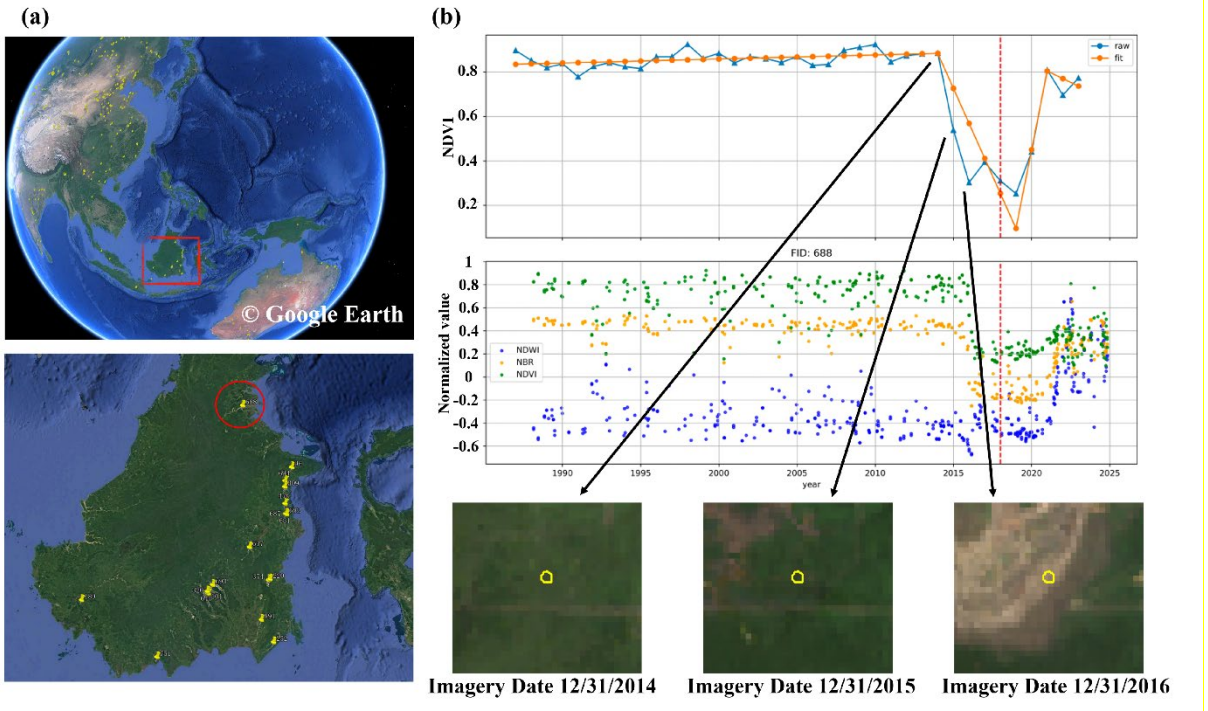


Figure 4. (a) Spatial location of a sample mine (WGS84: 3.552883°N, 117.169372°E) located on the island of Borneo, Indonesia. (b) Spectral validation results for the selected sample point (Point 688). The upper panel displays the NDVI time series (blue), LandTrendr segmentation results (orange), and the mining year inferred from the land cover time series (LCTS) indicated by a red vertical dashed line. The lower panel shows high-resolution Landsat 8 OLI imagery (Collection 2 Level-2 Surface Reflectance) downloaded directly from Google Earth Engine for 2014, 2015, and 2016. The yellow marker indicates the location of the 30 m validation pixel. Imagery was exported at 10 m resolution (3× supersampling) to enhance visual clarity while maintaining spectral fidelity of the original 30 m Landsat data.

5. **Percentage formatting:** All percentages in Sections 3.2 and 3.3 have been standardized to one decimal place for consistency throughout the manuscript.
6. **Dataset documentation:** We have prepared comprehensive metadata documentation that details the meaning, units, and calculation logic for each field in the dataset.