Mapping Mining Areas in the Tropics from 2016–2024

Philipp Sepin[1], **Lukas Vashold**[1], **and Nikolas Kuschnig**[1,*]

[1]Vienna University of Economics and Business
[*]Corresponding author: Nikolas Kuschnig (nikolas.kuschnig@wu.ac.at)

ABSTRACT

Mining provides crucial materials for the global economy and the climate transition, but has potentially severe adverse environmental and social impacts. Currently, the analysis of such impacts is obstructed by the poor availability of data on mining activity — particularly in regions most affected. In this paper, we present a novel panel dataset of mining areas in the tropical belt from 2016 to 2024. We use a transformer-based segmentation model, trained on an extensive dataset of mining polygons from the literature, to automatically delineate mining areas in satellite imagery over time. The resulting dataset features improved accuracy and reduced noise from human errors, and can readily be extended to cover new locations and points in time as they become available. Our comprehensive dataset of mining areas can be used to assess local environmental, social, and economic impacts of mining activity in regions where conventional data is not available or incomplete.

Transition minerals play a crucial role in climate action, necessary for the switch towards cleaner production, storage, and distribution of energy.[1] Mining operations are expanding globally to meet the growing demand for raw materials, often encroaching upon vulnerable regions.[2] Projections suggest that this expansion will accelerate drastically in pursuit of the goals set by the Paris Agreement and subsequent climate conferences.[3] Yearly extraction of critical minerals are projected to increase by 150–450% depending on the mineral, with a cumulative total of material extracted reaching 1.8–3.5 billion tons by 2050.[4] Effective management and an intricate understanding of the impacts of this extraction are crucial, but rely on comprehensive data that is often lacking.

On the one hand, mineral extraction is linked to several adverse environmental and social effects, including deforestation, loss of biodiversity, soil erosion, water pollution, air contamination, corruption, and violent conflicts.[5–13] On the other hand, mining presents an economic opportunity for locals, and has been shown to increase wealth levels, asset ownership, and incomes as well as related socioeconomic indicators.[14–17] If managed successfully, the increasing demand for minerals could positively affect local development, and facilitate the successful delivery of the United Nations sustainable development goals.[18] However, this necessitates information on the location, areal extent, and activity of mines *and* their development over time. Despite previous efforts,[19–21] this information is lacking at larger scales.[22]

In this paper, we introduce a panel (longitudinal) dataset of mining areas in the tropical belt. It covers the years from 2016 to 2024, and mines are automatically delineated from satellite imagery using state-of-the-art machine learning (ML) methods. Our approach employs a transformer-based segmentation model,[23] trained on an extensive dataset of mining polygons from existing literature,[19,21] and applies it to frequently available, high-resolution satellite imagery from Planet,[24] provided under Norway's International Climate and Forest Initiative (NICFI).[25] This way, we provide a comprehensive panel dataset of mine locations and polygons that follows their yearly areal expansion. The nature of our approach allows for high accuracy, reducing noise from human classification errors, and straightforward extension of the dataset — both in terms of locations and temporal range. The resulting data enables large-scale

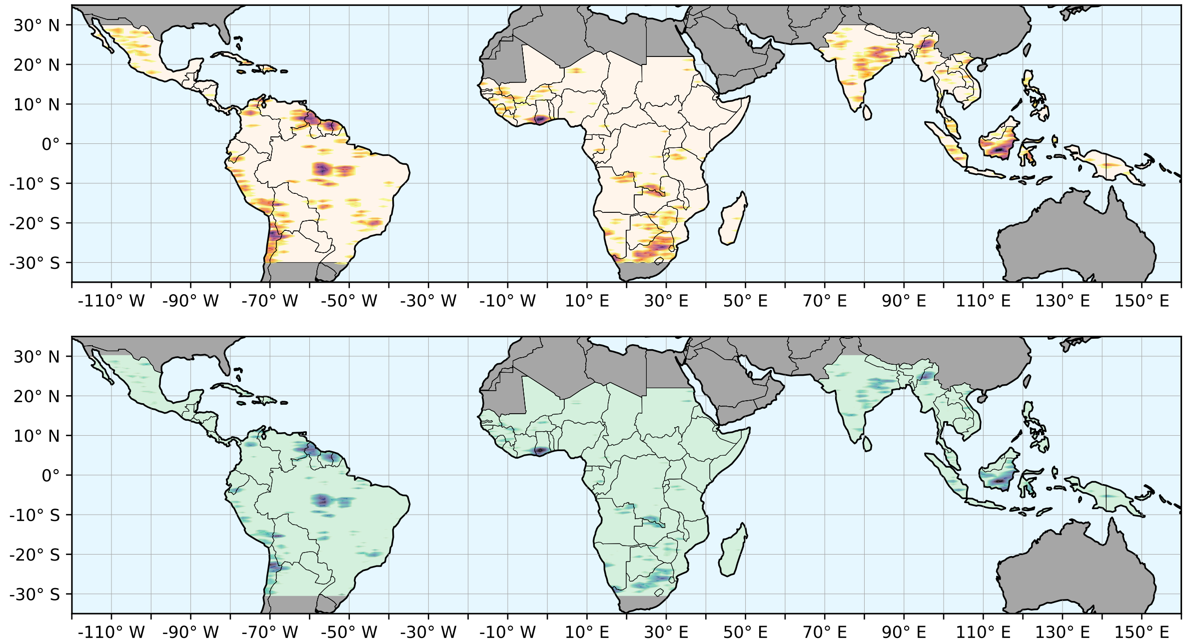

Figure 1. Mine density in 2024 (top panel) and its growth since 2016 (bottom panel). Colored regions indicate the coverage of the dataset (i.e., the tropical belt).

analyses of the various impacts of mineral extraction — particularly in regions where such data has been scarce in the past.

Results and Discussion

Our dataset provides yearly polygons for about 18,000 mining sites covering roughly 64,000 km^2 in the tropical belt from 2016 to 2024. It adds temporal information on the evolution of mining sites to previous delineation efforts,[19,21] providing crucial information needed to evaluate the various impacts of mining operations. An overview of the mine density and coverage is presented in Figure 1. Summary statistics for the most represented countries are provided in Table 1.

Country	2016	2017	2018	2019	2020	2021	2022	2023	2024
Indonesia	11,130.4	10,368.2	11,069.6	11,276.0	10,740.4	12,102.6	14,639.8	12,378.0	13,517.6
Brazil	8,103.0	8,608.2	8,286.2	8,481.2	10,252.6	10,798.2	10,620.0	10,347.1	9,648.3
Ghana	3,040.9	3,490.4	3,202.2	4,295.5	3,004.2	4,539.0	4,094.3	4,983.5	5,824.5
Chile	4,759.9	4,549.8	4,667.9	5,588.6	4,973.0	5,362.4	5,488.2	5,511.5	5,563.1
South Africa	4,435.9	4,620.0	4,536.0	4,717.8	5,151.3	4,894.9	5,014.5	4,919.2	4,884.6

Table 1. Mine area (in km^2) over time in the five countries that are most represented in the dataset.

Delineating a mine over time

Figure 2 provides an example of the delineation of a mining area over time. It shows the Toka Tindung mine, one of the largest gold mines in Southeast Asia, from 2016 to 2024. It is located in the Indonesian province of North Sulawesi, approximately 35 kilometers east of the capital city, Manado. Commercial production began in 2011, with steady increases in production and processing capacity. Multiple upgrades of the processing plant, more than quadrupling its throughput capacity to four million tons per year in the period captured by our predictions,[26] were implemented in conjunction with an increase of the area mined.

This development is clearly visible from Figure 2, and captured by our model. The main pits of the Toka Tindung mine have expanded rapidly, accompanied by additions of necessary infrastructure (such as water storage facilities). The previously rather disconnected Toka pit in the North and the Kopra, the Blambangan and the Araren pits in the South grow closer to each other in 2016–2018 before joining and being segmented as a single mine starting with 2019. Thereafter, development was concentrated in the southern parts of the mining area. Notably, our predictions also capture the development of the Araren South pit (in the bottom right corner of the satellite images) that was not part of the ground truth. Some artifacts, caused by nearby infrastructure, remain in the predictions (cf., the years 2021, 2022 and 2023). However, this exemplary case showcases the potential of the proposed approach to automatically detect changes in mining operations over time.

Mining clusters and expansion in the tropics

Mining is prevalent throughout the tropics but also highly clustered (cf. the upper panel of Figure 1). In Latin America, particular hotspots are observable in the densely forested countries Suriname and Guyana as well as the territory of French Guiana. In Brazil, mining operations are encroaching the Amazon but are also concentrated in the state of Minas Gerais in the South-East of the country. Another notable hotspot is in the Andes, and specifically Chile, the North of which is covered by our dataset. Other countries in the Americas such as Peru, Colombia, or Mexico also show considerable, but more spread out mining activity. In Africa, pronounced pockets of mining activity are in Ghana (mostly gold), South Africa (mostly coal), and Zimbabwe (diversified metal mining). In the Democratic Republic of the Congo, mining operations are concentrated in the southern savanna and temperate biomes, but are expanding rapidly into the Congo rainforest.[27] In Asia, the island of Borneo in Indonesia, eastern India, and the North of Myanmar are particular hotspots of mining activity. Other parts of Indonesia as well as Malaysia, the Philippines, and the North of Vietnam are home to extensive mining operations as well. Growth in mining areas since 2016 was also substantial in these areas (cf. the next section and the lower panel of Figure 1). This suggests that pressures on ecosystems in these hotspots have exacerbated in recent years.

Development of mining area over time

Figure 3 provides an overview of the evolution of total mining area, the number of mines, and the average size of mines over the sample period; overall and on a regional level for the Americas, Africa, and Asia and Oceania. Considering the period from 2017 to 2023,[1] the total area of the covered mining sites has increased by roughly 11% in the tropics globally, extending over about 68,000 km^2 in 2023. In this period, the number of unique mine polygons in our dataset decreased slightly, while the average size of them increased from roughly 3.3 km^2 to 3.8 km^2. These dynamics are similar on the regional level, with the largest absolute increase of mined area in Asia and Oceania of 2,600 km^2, whereas in Africa the relative increase was strongest with total mined area increasing by more than 20% from 2017 to 2023. Compared to the studies that constitute the ground truth in our prediction exercise,[19,21] we report substantially fewer, but larger mines in the year 2019. The main reason for these discrepancies is that our segmentation model tends to merge small, closely situated mines or mining pits into a single, larger mine (cf. Figure 4). The slight decrease in the number of mines over time follows mostly from a similar behavior, where individual mines or pits grow and join to form larger ones.

[1] Recall that the postprocessing steps result in more conservative predictions for the initial and final years (2016 and 2024).

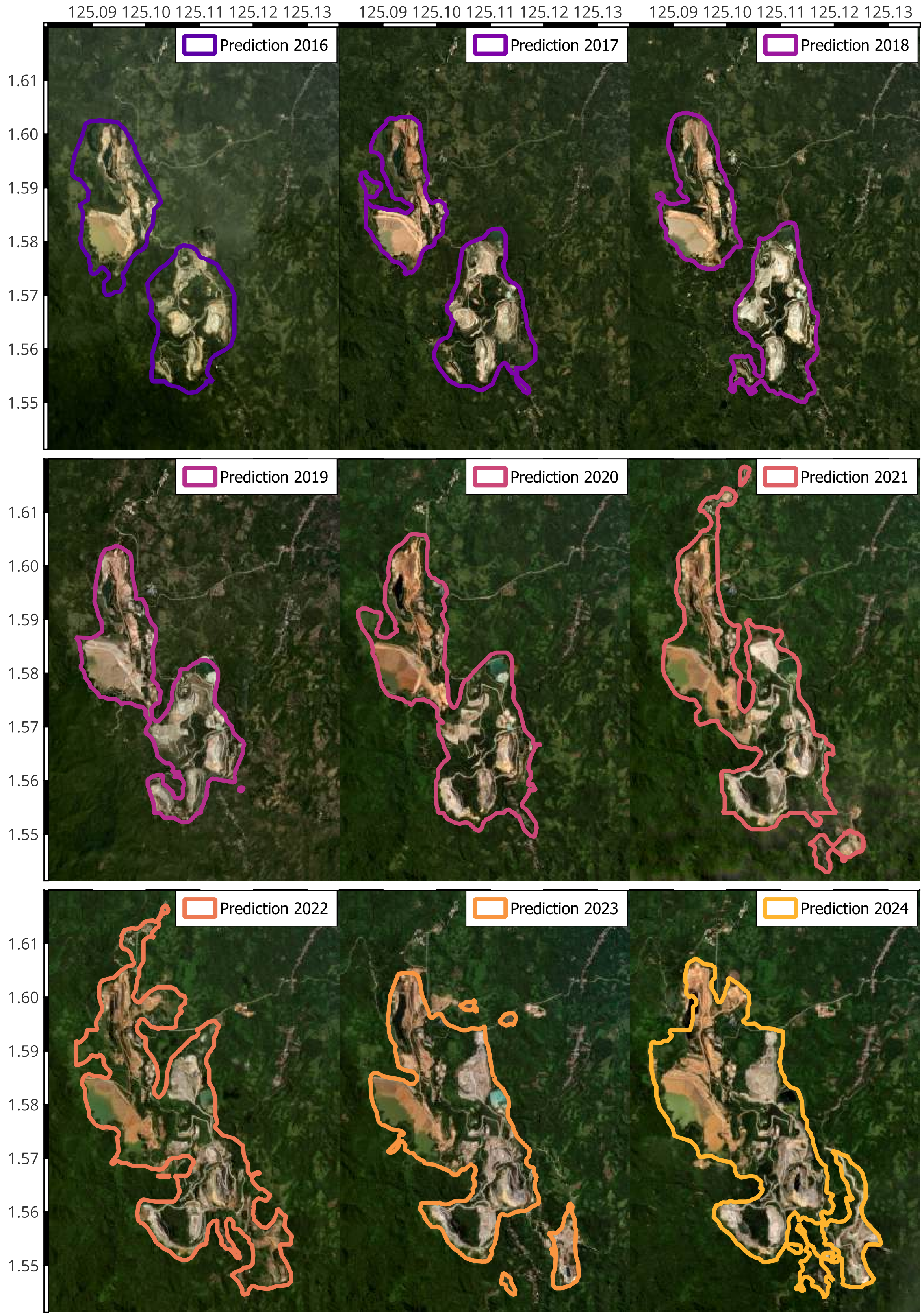

Figure 2. Predictions of the extent of the Toka Tindung gold mine in Indonesia (1°35'N 125°06'E) from 2016 to 2024 over the corresponding satellite image used to predict (Planet/NICFI).

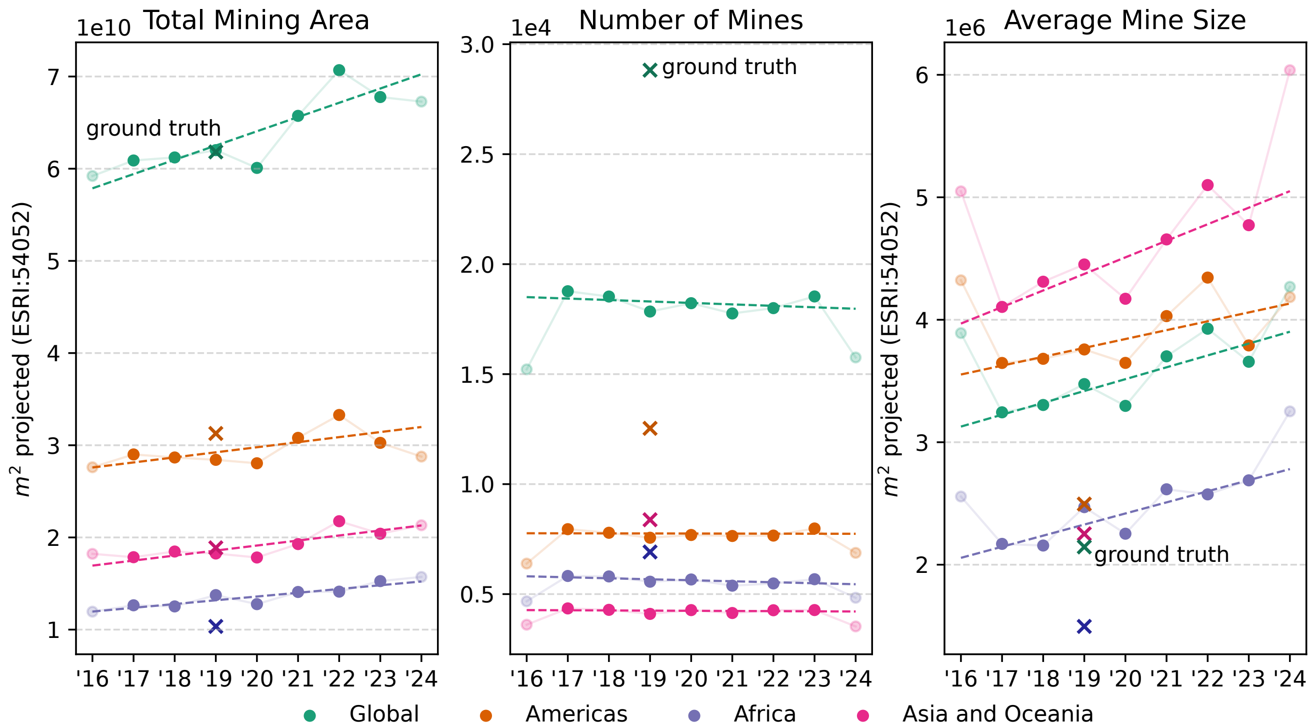

Figure 3. Summary of the predictions over time for the full dataset, and separated into the Americas, Africa, and Asia and Oceania. The left panel shows an increase in mining areas (in 10,000 km^2) over time. The center panel shows the number of individual mine polygons (in thousands), remaining comparatively steady since no new mine locations are considered. The right panel shows the average size of these individual mine polygons (in km^2).

Within regions, mining activities and their expansion are typically concentrated in a few countries. Figure 7 in the Supplementary Information decomposes the regional graphs in Figure 3 to show the three countries within them that have the largest share of total mining area. In the Americas, Brazil is the country with the largest area classified as mining area. Growth therein picked up substantially in the years 2020–2022, during the administration of Jair Bolsonaro, with most of it taking place in the arc of deforestation at the fringe of the Amazon rainforest (cf. Figure 1). This development corroborates early concerns about the environmental damages of legislative changes and increases in anti-environmental rhetoric.[28] In Chile, which hosts mines that are on average larger compared to other countries in the region, the area of mining operations stagnated in the observation period. Next, the Peruvian mining industry saw an investment boom after the onset of the COVID-19 pandemic,[29] and experienced substantial growth in mining areas in the years thereafter. This follows years of stagnating or declining number and area of mines.

Hosting by far the largest share of mining areas in Asia and Oceania, the development of Indonesian mining activities in from 2017–2023 was characterized by two distinct periods. In the years until 2019, total mining area in the country increased slightly, while dipping in 2020, perhaps following declines in commodity prices and reduced demand from China due to the COVID-19 pandemic.[30] Activities picked up substantially again in 2021,[31] reflected by strong increases in mining areas until 2022, before dipping again in 2023. In India, mining operations remained rather stable in terms of their area and number. By contrast, political turmoil in Myanmar may have induced more pronounced dynamics therein. With the ban of exports of certain minerals and the unwinding of Chinese mining operations in 2018,[32] the total mining area and number of mines reduced gradually until 2021. Since the military coup that toppled the civilian government in the same year, mining activities have picked up again, particularly in the Northern region bordering China (cf. Figure 7 and Figure 1).

In South Africa, the country with the most mining area in Africa, the aerial expansion of mines was rather muted. This is aligned with rather flat mining production in terms of physical volume since the early 2010s.[33] Most of the region's overall growth in mining area stems from Ghana, the second mining powerhouse in tropical Africa. Total area mined has expanded steadily in the observation period, coinciding with increases in international prices of gold, which accounts for around 95% of the country's mineral revenues. In Zambia, where the mining sector represents an essential part of the economy, both the number and the area of mining operations were rather stable in the observation period. However, with tax incentives provided to mining companies in a restructuring of the Zambian fiscal regime in 2022, this might be bound to change.[34] In the rest of Africa, mines are of substantially smaller scale and artisanal mining is prevalent.[35] Keeping in mind that the ground truth on which we base our segmentation process potentially misses a significant portion of these operations, we detect no substantial increases in the rest of African mining areas.

Methods

In this section we describe the methods employed to produce the dataset.

Satellite imagery

For our study, we use high-resolution ($<$ 5m per pixel) satellite imagery from Planet,[24] provided by Norway's International Climate and Forest Initiative (NICFI).[25] This imagery covers the tropical belt, ± 30 degrees of latitude, for a geographical area of around 45 million square kilometers. As shown in Figure 1, this includes important mineral extracting countries such as Indonesia, the Democratic Republic of the Congo, and Brazil, but excludes some other notable countries such as Australia, the US, Russia, or China, and only partially covers others such as Chile, Argentina, or South Africa. The imagery is readily accessible for non-commercial use, and is provided biannually from December 2015 to June 2020 and monthly thereafter. The biannual data from 2016 to 2020 has been pre-processed to be relatively cloud-free, while the monthly data from 2021 onward has not been cleared of clouds. For our delineation, we use data from the second half of the year for the biannual data and the least cloudy month between May and September for the monthly data, resulting in a yearly dataset that covers the years from 2016 up to and including 2024.

Mining locations

The ground truth of our training data, as well as the rough mining locations for prediction, are based on mining polygons from two datasets. The first one is first presented in Maus et al. (2020)[36], and updated by Maus et al. (2022)[19]. This dataset covers industrial and artisanal mines, including the ones present in the 'SNL Metals and Mining' database. The updated data is based on Sentinel-2 imagery from 2019, while the original dataset combined imagery from Google Satellite, Microsoft Bing Imagery, and Sentinel-2 from various years. It contains $44,929$ polygons covering roughly $102,000$ km^2. The second dataset is an extended version of Tang et al. (2021)[37], which uses imagery mainly from the period 2018–2020, updated by Tang and Werner (2023)[21], for which no clear information on the baseline satellite imagery was given. It contains $74,548$ polygons, covering roughly $66,000$ km^2.

In terms of mining locations, there is a large overlap between these datasets. However, the delineation approaches and polygon's resolution adopted by the authors differ. The polygons of Tang and Werner[21] are delineated precisely, distinguishing between features of singular mines, such as tailing dams and administrative buildings. By contrast, the polygons of Maus et al.[19] often contain surrounding areas that may be used by mining operations, but can hardly be described as mining areas themselves.[2]

For our approach, we set the polygon ground truth to be the union of these two datasets, allowing us to take advantage of both sources of information.[3] The sometimes imprecise nature, and in particular the lack of exact timestamps, of the ground truth provided by these two datasets presents one major limitation of our approach. We have to assume that mines were delineated from satellite imagery as of 2019, although many polygons might have been delineated using on earlier imagery. Moreover, the sensors used to delineate the ground surface are heterogeneous, and human error and the different delineation approaches introduce further noise. As we will demonstrate below, these issues limit the performance of our modeling approach in some regards, but also present an opportunity for the model to distill a common 'truth' from two noisily measured datasets.

Segmentation

For the delineation of mining areas over time, we use a segmentation model based on state-of-the-art transformer architectures from the ML literature.[38,39] The process of obtaining polygons of mining operations over time consists of four steps: (1) preprocessing the annual satellite imagery and generating images with which the segmentation model can be trained, (2) training the model on ground truth data with satellite images from 2019, (3) predicting mine delineations for the years 2016–2024 and tracing them back to their geospatial location, and (4) postprocessing the resulting predictions to reduce noise.

Preprocessing Two data sources are required to train the segmentation model — an image and its corresponding segmentation mask, which indicates the mining areas on the image. As the model is trained on satellite imagery from 2019, the masks only need to be generated for this year. The images, however, are used for predictions and need to be generated for every year. They are generated as square images that are twice the maximum side length of the polygon and, if necessary, down-scaled to a resolution of 512×512 pixels. The resulting data is split into train/test/validation samples using an $80/10/10$ ratio, resulting in approximately 16,600 training samples and 2,000 samples for testing and validation.

Segmentation model For the segmentation task, we use the SegFormer model.[23] Its strong performance has been demonstrated for benchmark data sets,[40,41] and it has been shown to outperform older architectures, such as U-Net[42,43] which has been used to delineate mining areas in Sub-Saharan Africa before.[20] Unlike older segmentation models, SegFormer is not based on convolutional neural networks (CNNs), but on a transformer architecture.[38] In computer vision, transformers have been shown to possess a larger effective receptive field,[44] whilst having lower complexity than comparable CNNs,[23] and therefore outperform them at segmentation tasks.[43] This enables transformer-based models to analyze images with more global context, whilst offering computational benefits due to their parallelizable training.[38]

Specifically, we work with SegFormer-B5,[23] which is the largest model variant. The encoder is pre-trained on the Imagenet-1k data set[45] before training it on our global mining area training data set for 160,000 iterations.[4] Our main measure to evaluate training performance is the *mean Intersection over*

[2]An illustration of the training data is provided in the Supplementary Information as Figure 6, where polygons of four locations from these datasets are plotted over Planet/NICFI imagery from 2019.

[3]We also experimented with using either of the datasets, but achieved worse results. For the dataset by Tang and Werner[21], we find that the precise delineation cannot be accurately recovered. When considering only the dataset by Maus et al.[19], we also find slightly worse segmentation results than when using the union of both datasets.

[4]We train the model on two NVIDIA A30 cards and use class balanced loss to address the great class imbalance of mining

Union (mIoU), which is commonly used in the ML domain. Our model reached a mIoU of 56.97 and a mean accuracy of 73.08 on the test set. For comparison, a U-Net model[42] trained with comparable training parameters reached a mIoU of 54.25 and a mean accuracy of 57.11 on the same test set.

Prediction The prediction process is illustrated in Figure 4. Following the training, we use the model to predict mine delineations on all images of mining locations (visualized in the left column) for the years 2016–2024. The output of the model represents a 'probability of mine' for a given pixel (center column), indicating the model's confidence for the presence of a mine at that specific position. To classify a pixel as part of a mine, we use a fixed probability threshold of 0.5. Notably, this parameter can be adjusted to tweak the false positive and false negative rates.[47] Predictions are then available in the same binary array format as the segmentation masks. We transform them back to polygons using a contour finding algorithm,[48] and trace the predicted polygons back to their geospatial positions using their previously calculated bounding boxes. The accuracy of our predictions is assessed in the *Technical Validation* section.

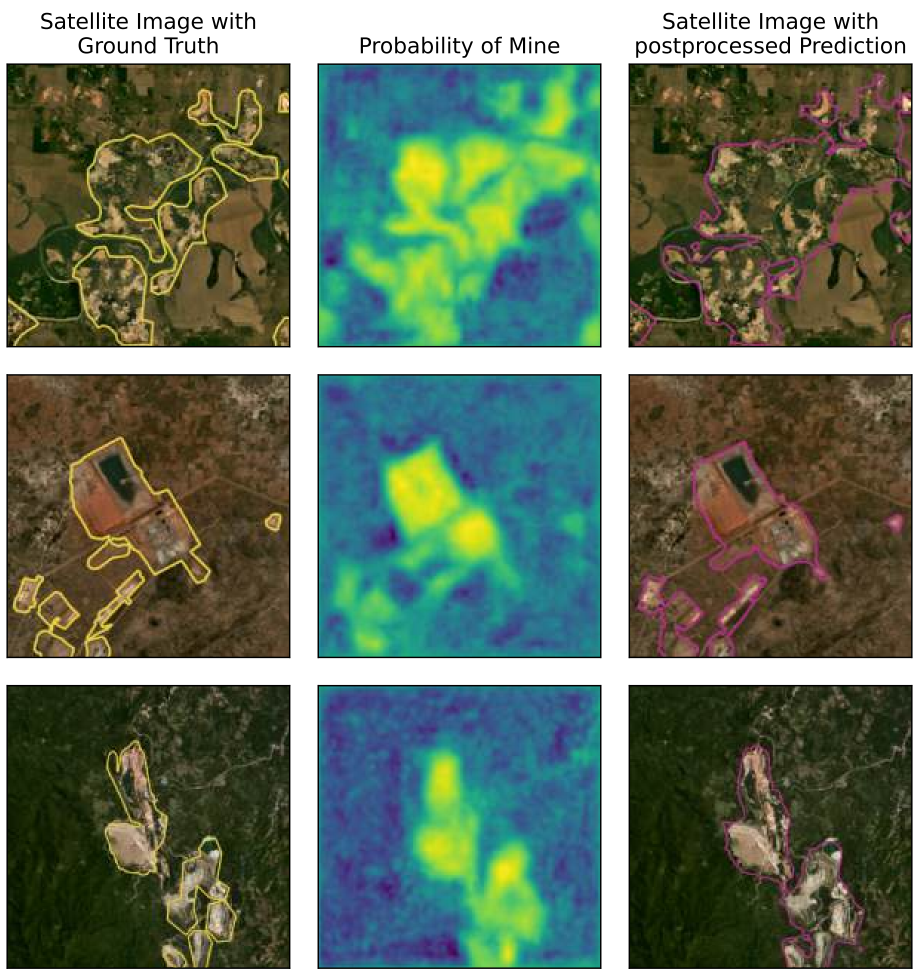

Figure 4. Prediction process for three mines that are located in Brazil, Mozambique, and Indonesia. The left column shows polygons of the ground truth, the center column shows the model's probabilistic prediction, and the right column shows the postprocessed predictions.

---

and non-mining areas in the images. We also apply data augmentation techniques, such as random horizontal and vertical flipping, random cropping and random resizing, as well as online hard example mining (OHEM),[46] which focuses the training on difficult examples.

Unlike manual methods, predictions with our automated approach allow for straightforward scalability in both spatial and temporal dimensions. The temporal scope of the dataset can be easily expanded by generating new predictions as more recent satellite images become available. Additionally, the spatial scope of the dataset can be broadened by incorporating new mining locations for the model to predict on. For instance, our method allows us to use point locations with a single timestamp to generate polygons of mine areas with their temporal evolution over several years at these locations.

Postprocessing To ensure temporal continuity of the predictions and reduce the number of false positives, we use several postprocessing steps. The central one, which allows us to borrow information across years, entails the assumption that mines do not emerge and disappear within a single year. Consequently, we remove any polygons that do not have an intersecting polygon in the previous *or* subsequent year. This means that the years 2016 and 2024, which only have a single 'neighboring' year, which we can compare the polygons to, feature a lower number of polygons. As a result, they are to be seen as rather conservative predictions. We also employed morphological operations, such as hole-filling, binary opening, and binary erosion during postprocessing to remove fragments. Since some secondary mines reach into the bounding boxes of other ones, they have multiple, intersecting predictions. We deal with those intersecting polygons by taking their union. Lastly, we removed predicted polygons that were too small to be considered mines and could be regarded as noise.

Technical Validation

To validate our data, check for errors, and ensure comparability with the ground truth, we followed the best practices for assessing map accuracy and selected a representative subset of the validation points generated provided by Maus et al.[19] These points were randomly generated within the 'region of interest', which they defined as a 10 km buffer around the mines. Of the 1,220 original validation points, 450 were covered by Planet/NICFI data. From these, a random subset of 200 points was manually inspected and labelled as either mine or non-mine for each of the nine years, and compared to our predictions for these years.

With an average accuracy of 0.88 and an average F1 score of 0.81, our predictions demonstrate performance comparable to both data sources used for the ground truth. The temporal variation may in part be due to varying cloud cover and inconsistencies in the satellite imagery. All of our average scores lie between the two ground truths, demonstrating how our method combines the liberal delineation style of Maus et al.[19] with the more precise delineation style of Tang and Werner.[21] This suggests that our model effectively integrates information from both ground truths, identifying a shared notion of 'truth' and extending it to nine different timestamps.

Usage Notes

The dataset presented here, including validation points, is available at this link and will be made available openly on Zenodo and Kaggle (the training data is posted at kaggle.com). It is licensed under the Open Database License (available at opendatacommons.org). Any rights in individual contents of the database are licensed under the Database Contents License (available at opendatacommons.org). The dataset is provided in the standard Geopackage (GPKG) format and contains information on the polygons themselves, the countries where they are located, the years in which they are segmented, as well as their areas. The validation points are also provided in the standard Geopackage (GPKG) format and include their locations, the countries where they are located, and the labels assigned to them for each year (1 representing the presence of a mine, 0 representing no presence).

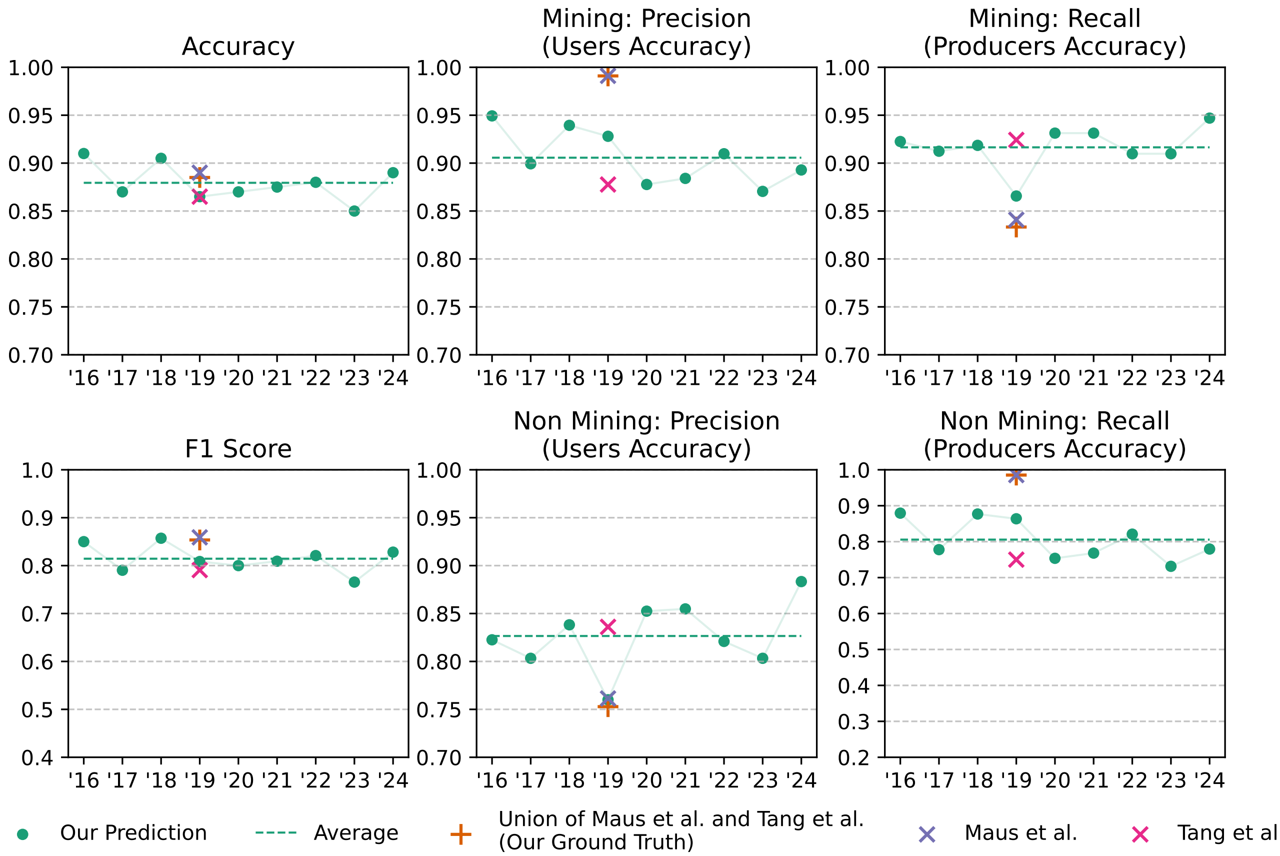

Figure 5. Validation results for our prediction and the equivalent for the ground truth and its components.

Code availability

All code used to produce the dataset and results of the paper is provided under an open GNU General Public License v3.0 (GPL-v3) from the repository at GitHub. All scripts were written in Python, with geospatial processing heavily utilizing the GeoPandas[49] and Shapely[50] packages. The model was implemented using PyTorch[51] and MMSegmentation.[52]

References

1. Shukla, P. R. *et al.* *Climate Change 2022: Mitigation of Climate Change* (Cambridge University Press, Cambridge, 2022). [Online; accessed 15. Apr. 2023].

2. Luckeneder, S., Giljum, S., Schaffartzik, A., Maus, V. & Tost, M. Surge in global metal mining threatens vulnerable ecosystems. *Glob. Environ. Chang.* **69**, 102303, 10.1016/j.gloenvcha.2021.102303 (2021).

3. IEA. Energy technology perspectives 2023. Tech. Rep., International Energy Agency (2023).

4. Hund, K., La Porta, D., Fabregas, T. P., Laing, T. & Drexhage, J. Minerals for climate action: The mineral intensity of the clean energy transition. Tech. Rep., World Bank Group (2023).

5. Kossoff, D. *et al.* Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. *Appl. Geochem.* **51**, 229–245, 10.1016/j.apgeochem.2014.09.010 (2014).

6. Berman, N., Couttenier, M., Rohner, D. & Thoenig, M. This mine is mine! how minerals fuel conflicts in Africa. *Am. Econ. Rev.* **107**, 1564–1610, 10.1257/aer.20150774 (2017).
7. Sonter, L. J., Dade, M. C., Watson, J. E. M. & Valenta, R. K. Renewable energy production will exacerbate mining threats to biodiversity. *Nat. Commun.* **11**, 1–6, 10.1038/s41467-020-17928-5 (2020).
8. Liu, Y. *et al.* A review of water pollution arising from agriculture and mining activities in central asia: Facts, causes and effects. *Environ. Pollut.* **291**, 118209, 10.1016/j.envpol.2021.118209 (2021).
9. Giljum, S. *et al.* A pantropical assessment of deforestation caused by industrial mining. *Proc. Natl. Acad. Sci.* **119**, e2118273119, 10.1073/pnas.2118273119 (2022).
10. Villén-Pérez, S., Anaya-Valenzuela, L., Conrado da Cruz, D. & Fearnside, P. M. Mining threatens isolated indigenous peoples in the Brazilian Amazon. *Glob. Environ. Chang.* **72**, 102398, 10.1016/j.gloenvcha.2021.102398 (2022).
11. Berman, N., Couttenier, M. & Girard, V. Mineral resources and the salience of ethnic identities. *Econ. J.* **133**, 1705–1737, 10.1093/ej/uead018 (2023).
12. Macklin, M. G. *et al.* Impacts of metal mining on river systems: A global assessment. *Science* **381**, 1345–1350, 10.1126/science.adg6704 (2023).
13. Mwelwa, S., Chungu, D., Tailoka, F., Beesigamukama, D. & Tanga, C. Biotransfer of heavy metals along the soil-plant-edible insect-human food chain in Africa. *Sci. The Total. Environ.* **881**, 163150, 10.1016/j.scitotenv.2023.163150 (2023).
14. Aragón, F. M. & Rud, J. P. Polluting industries and agricultural productivity: Evidence from mining in Ghana. *The Econ. J.* **126**, 1980–2011, 10.1111/ecoj.12244 (2015).
15. von der Goltz, J. & Barnwal, P. Mines: The local wealth and health effects of mineral mining in developing countries. *J. Dev. Econ.* **139**, 1–16, 10.1016/j.jdeveco.2018.05.005 (2019).
16. Bazillier, R. & Girard, V. The gold digger and the machine. evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso. *J. Dev. Econ.* **143**, 102411, 10.1016/j.jdeveco.2019.102411 (2020).
17. Ofosu, G., Dittmann, A., Sarpong, D. & Botchie, D. Socio-economic and environmental implications of artisanal and small-scale mining (asm) on agriculture and livelihoods. *Environ. Sci. & Policy* **106**, 210–220, 10.1016/j.envsci.2020.02.005 (2020).
18. Ali, S. H. *et al.* Mineral supply for sustainable development requires resource governance. *Nature* **543**, 367–372, 10.1038/nature21359 (2017).
19. Maus, V. *et al.* An update on global mining land use. *Sci. data* **9**, 1–11 (2022).
20. Provenzano, S. & Bull, H. The local economic impact of mineral mining in Africa: Evidence from four decades of satellite imagery. *arXiv preprint arXiv:2111.05783* (2021).
21. Tang, L. & Werner, T. T. Global mining footprint mapped from high-resolution satellite imagery. *Commun. Earth & Environ.* **4**, 134 (2023).
22. Maus, V. & Werner, T. T. Impacts for half of the world's mining areas are undocumented. *Nature* **625**, 26–29, 10.1038/d41586-023-04090-3 (2024).
23. Xie, E. *et al.* Segformer: Simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inf. Process. Syst.* **34**, 12077–12090 (2021).

24. PBC, P. L. Planet application program interface: In space for life on Earth (2016–2024).

25. NICFI. Norway's international climate and forest initiative (NICFI) (2023).

26. Indonesia, A. Pt archi indonesia tbk (idx: Arci) public expose 2024 (2024). [Online; accessed 05. Aug. 2024].

27. Verweijen, J., Schouten, P., Simpson, F. O. & Pascal, C. Z. Conservation, conflict and semi-industrial mining: the case of eastern DRC. *Analysis Policy Briefs* (2022).

28. Ferrante, L. & Fearnside, P. M. Brazil threatens indigenous lands. *Science* **368**, 481–482, 10.1126/science.abb6327 (2020).

29. Belapatiño, V., Crispin, Y., Grippa, F. & Vega, H. Peru — mining sector outlook (2023).

30. EITI. Indonesia 2019–2020 eiti report (2022). [Online; accessed 8. Jul. 2024].

31. EITI. Indonesia 2021 eiti report (2022). [Online; accessed 8. Jul. 2024].

32. GW. Fuelling the future, poisoning the present: Myanmar's rare earth boom (2024). [Online; accessed 8. Jul. 2024].

33. STATS-SA. P2401 – mining: Production and sales (2024). [Online; accessed 8. Jul. 2024].

34. EITI. Zambia 2022 eiti report (2023). [Online; accessed 8. Jul. 2024].

35. Hilson, G. & McQuilken, J. Four decades of support for artisanal and small-scale mining in sub-Saharan Africa: A critical review. *Extr. Ind. Soc.* **1**, 104–118, 10.1016/j.exis.2014.01.002 (2014).

36. Maus, V. *et al.* A global-scale data set of mining areas. *Sci. Data* **7**, 1–13, 10.1038/s41597-020-00624-w (2020).

37. Tang, L., Werner, T. T., Heping, X., Jingsong, Y. & Zeming, S. A global-scale spatial assessment and geodatabase of mine areas. *Glob. Planet. Chang.* **204**, 103578, 10.1016/j.gloplacha.2021.103578 (2021).

38. Vaswani, A. *et al.* Attention is all you need. *Adv. neural information processing systems* **30** (2017).

39. Li, X. *et al.* Transformer-based visual segmentation: A survey. *ArXiv e-prints* 10.48550/arXiv.2304.09854 (2023). 2304.09854.

40. PapersWithCode. Semantic segmentation on ade20k (2024).

41. PapersWithCode. Semantic segmentation on cityscapes (2024).

42. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, 234–241 (Springer, 2015).

43. Sourget, T., Hasany, S. N., Mériaudeau, F. & Petitjean, C. Can SegFormer be a true competitor to U-net for medical image segmentation? In *27th Conference on Medical Image Understanding and Analysis 2023* (2023).

44. Luo, W., Li, Y., Urtasun, R. & Zemel, R. Understanding the effective receptive field in deep convolutional neural networks. *Adv. neural information processing systems* **29** (2016).

45. Russakovsky, O. *et al.* ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis. (IJCV)* **115**, 211–252, 10.1007/s11263-015-0816-y (2015).

46. Shrivastava, A., Gupta, A. & Girshick, R. Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 761–769 (2016).

47. Boss, K., Groeger, A., Heidland, T., Krueger, F. & Zheng, C. Forecasting bilateral refugee flows with high-dimensional data and Machine Learning techniques (2023).

48. Suzuki, S. *et al.* Topological structural analysis of digitized binary images by border following. *Comput. vision, graphics, image processing* **30**, 32–46 (1985).

49. Jordahl, K. *et al.* Geopandas 1.0.1, 10.5281/zenodo.2585848 (2024).

50. Shapely. Shapely: A python package for manipulation and analysis of planar geometric objects. https://shapely.readthedocs.io/en/stable/index.html (2024).

51. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Adv. neural information processing systems* **32** (2019).

52. MMSegmentation. Mmsegmentation: Openmmlab semantic segmentation toolbox and benchmark. https://github.com/open-mmlab/mmsegmentation (2020).

Acknowledgements

The authors gratefully acknowledge financial support from the Austrian National Bank (OeNB anniversary fund, project No. 18799).

Author contributions statement

N.K and P.S conceived the study. P.S implemented the method. All authors wrote the paper.

Competing interests

The authors declare no competing interests.

Supplementary Information

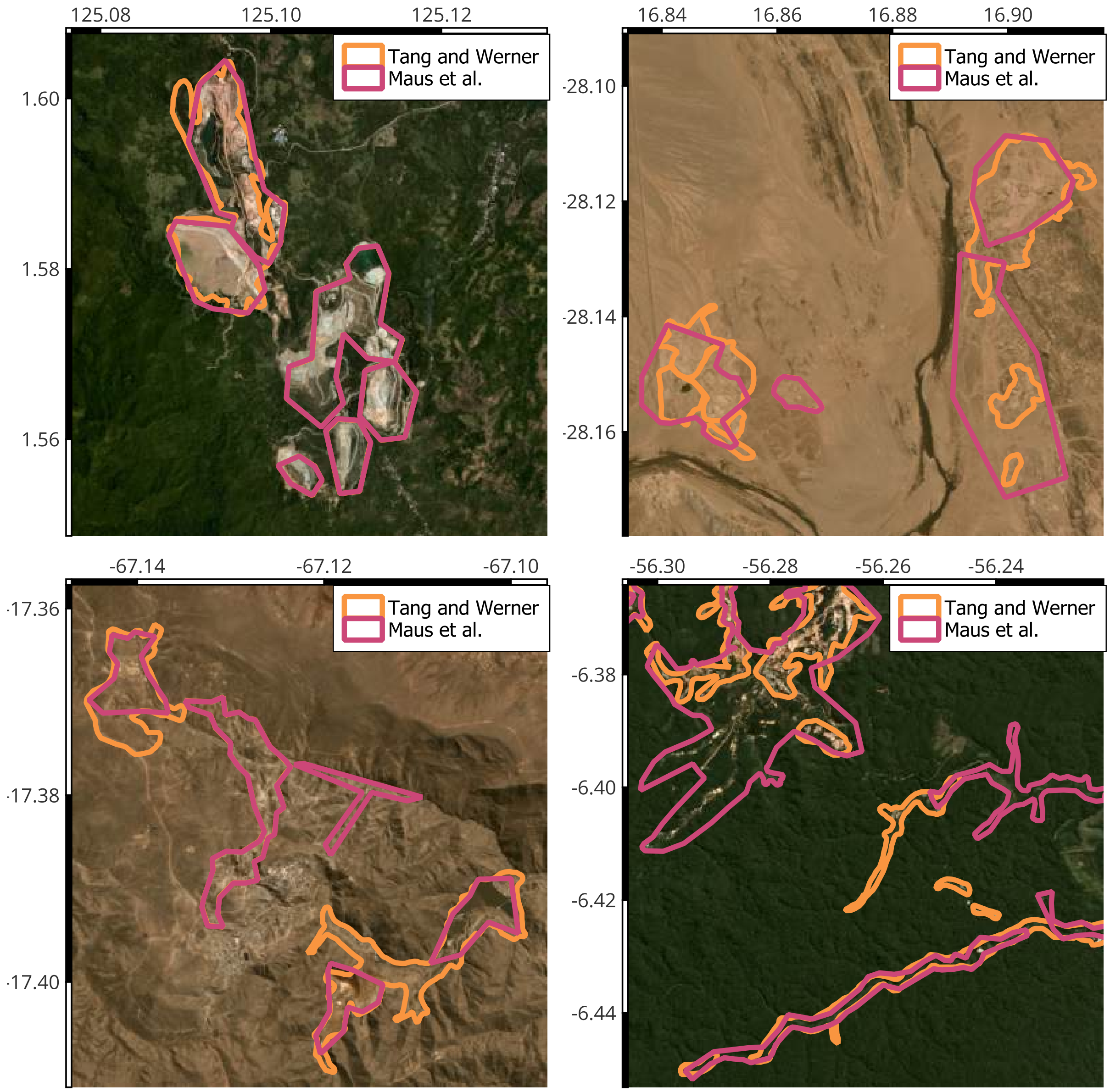

Figure 6. Four examples of the ground truth[19,21] for mines located in Indonesia, South Africa, Bolivia, and Brazil. Note the discrepancies in coverage (left panels) and segmentation approach (right panel).

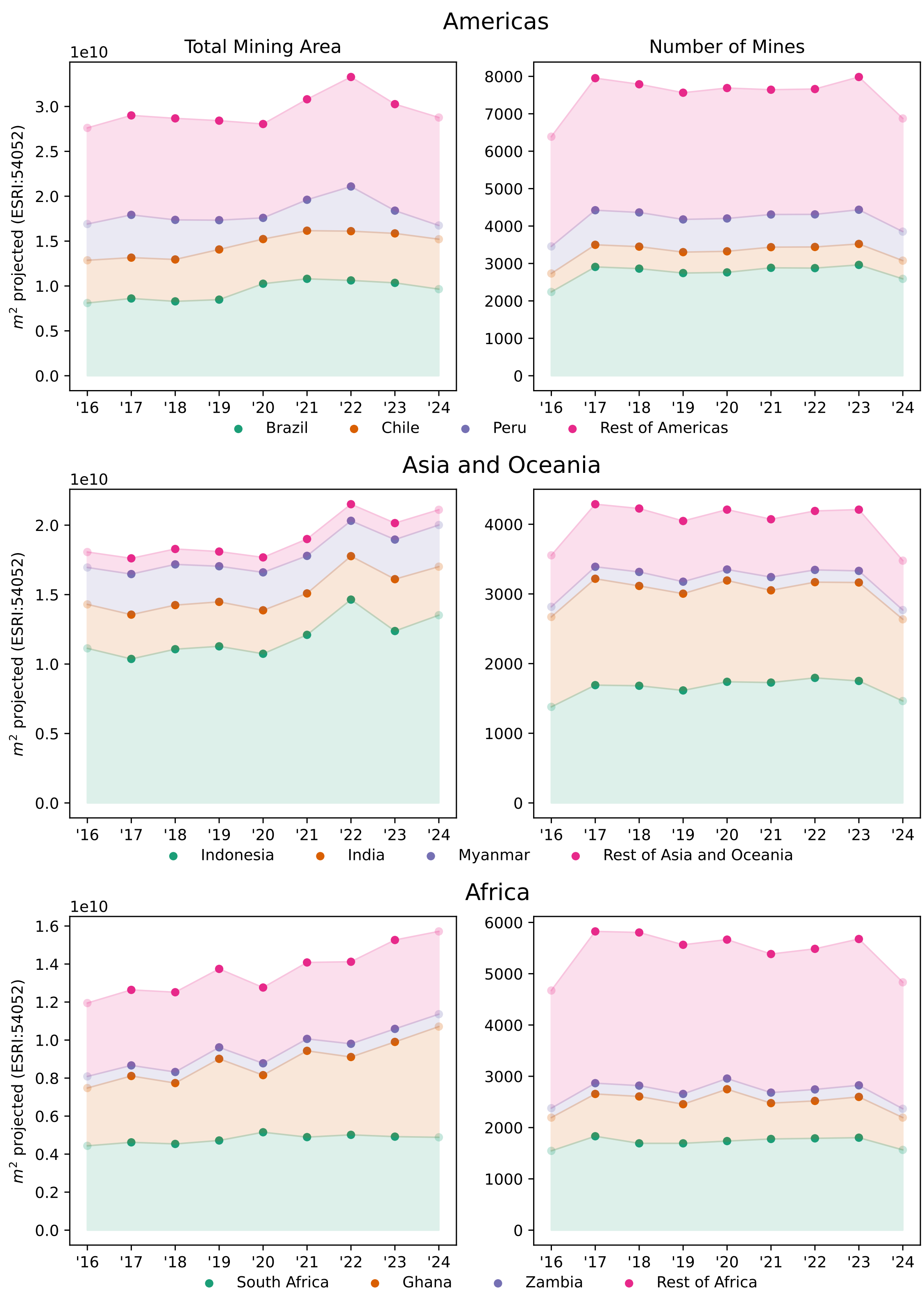

Figure 7. Regional summaries of the predictions over time; the left column shows total mining areas, and the right the number of mine polygons. Note, in particular, the slight increase in area for Brazil in 2020 (top row), the sharp increase for Indonesia in 2021 and 2022, and the steady expansion for Ghana.