

Overlooked deforestation from global mining activities in the 21st century

Corresponding Author: Professor Bin Chen

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The scientific questions posed by the authors are highly pertinent, and providing satisfactory answers would significantly elevate the quality of the work. However, only the first question—"What is the geographic distribution of different mining activities globally?"—has been addressed sufficiently to merit approval.

In contrast, the second and third questions—"To what extent do mining activities contribute to deforestation and associated carbon emissions?" and "Which types of mining have been the primary drivers of deforestation over the years?"—rely heavily on the Global Forest Canopy Loss (GFCL) product developed by Hansen et al. (2013). This product was designed to monitor deforestation at a global or regional scale, with validation procedures aligned to such scales. However, Zhang et al. apply this product to much smaller, localized areas, phenomena study.

For instance, data from Table S2 reveals that approximately 20% of the mining polygons in their study are smaller than 1 hectare. This translates to just ~11 pixels, a scale at which the quality metrics of Hansen et al. (2013) may not be reliable. The Hansen product was validated using pixels representing forest cover, deforestation, and regrowth (over 5 m height) across five global biomes. The GFCL validation process was not designed to estimate deforestation at scale caused by anthropic activities such as mining.

I recommend that the authors conduct a sampling exercise using appropriate statistical models to estimate the deforested area directly caused by mining activities (see: <https://www.sciencedirect.com/science/article/pii/S0034425714000704>). If the error margins in their deforestation estimates prove unsatisfactory, I further suggest that they consider independently estimating deforestation, bypassing the Hansen et al. (2013) product altogether.

Additionally, there are uncertainties about how the authors account for mining areas located in regions with original natural vegetation less than 5 meters in height. Though not classified as forests in the conventional sense, these areas should still be accounted for, and the authors should consider methods to estimate the impact on natural vegetation in such regions.

A few minor points also require attention. On line 89, I believe the authors intended to write "1,215.18 thousand hectares" rather than "million hectares." To maintain consistency with the other measurements in the article, I suggest using km² instead of hectares. Additionally, please correct the unit of measurement in Table S2, where the title specifies km² but the data appears to use hectares, which is somewhat confusing.

In conclusion, while the manuscript is well-written and tackles important questions, the lack of robust error estimates in the deforestation data linked to mining activities weakens the validity of the results. Without addressing these concerns, the study is not yet ready for publication.

Reviewer #2

(Remarks to the Author)

The work of Zhang et al. (2024) provides a new global dataset for mining activities, identifying documented and undocumented sites and quantifying the deforestation and CO₂ emissions associated with these mining polygons. Overall, the manuscript is well written and has good research questions. I have a few suggestions below to improve its understanding and outreach:

The dataset considers only mining where there is land cover change, that is, artisanal small-scale and mining on shafts in unaccounted. This is partially discussed in the "Discussion" section but must be mentioned in the main text.

We have the impression during the text that the authors created the dataset by classifying images and then combined this

classification with existing datasets, but this is not the case. Please clarify this throughout the text.

The so-called validation of the dataset consists of the visual interpretation of 2,000 polygons (as shown in Fig. S2). My concern is that this represents less than 1% of the total polygons, so it is insufficient to corroborate the more than 87% accuracy of the dataset. Moreover, how many of these polygons were documented and undocumented? Authors must provide a more robust validation of the dataset.

Most of the figures are confusing, sometimes combining 8 panels, and need to be improved. Please see specific comments below.

Authors confuse key terms related to deforestation, as they consider deforestation, forest loss, and forest cover the same throughout the text. For example, forest loss encompasses other causes than deforestation such as hurricanes. You must define and revise these terms.

I do not agree with the Global North x Global South comparisons made by the authors. This reproduces colonialist practices that omit the fact that the majority of the mining in the Global South is consumed in the Global North. So, I suggest removing this comparison.

Specific Comments:

Line 17: Please remove "believed" and adjust the sentence.

Lines 37-38: This is an outdated reference. Could you please check if it still stands?

Lines 38-40: Now you have returned to the global scale. This is confusing.

Lines 42-43: Which decades?

Lines 37-45: Text flow is not good in here. It keeps shifting from global to regional scale.

Lines 45-46: Which negative impacts?

Fig 1. It is impossible to evaluate the heatmap results due to the overlap with the points related to the primary commodities.

Numbers are hard to read in a. What is the meaning of the numbers in d?

Table 1. What did you mean by count? Number of mining polygons?

Fig. S2 appears before Fig. S1 in the text.

Fig 2. There is a non-described graph inside the map. What does it mean? The figure is related to deforestation, so why did you add CO₂ and budget information on it? Why to divide in Global South and Global North in e? Meaning of the black line in c? In b, divide the emissions in documented and undocumented too?

Line 109: You have CO₂ information in Fig. 2 too. Why? Is it Fig. 2 or Fig. 3?

Line 121: Deforestation from mining.

Lines 125-128: I do not see the point of this North x South analysis. This might lead to conclusions that reflect the colonialism process.

Lines 129-130: Did you perform a statistical analysis to corroborate this so-called trend?

Fig. 3. In a, missing names and values. C appears after d in the caption. Please, clarify in the graphs that you are considering only gold mines. Not sure what f means.

Lines 151-153: Also needs a statistical analysis to affirm this.

Lines 163-137: How did you aggregate the minerals in these classes?

Line 178: Forest cover x forest loss x deforestation.

Line 204: Missing statistical analysis to support this.

Line 213: How did you derive the mining land use from OSM?

Lines 245-248: Which images were used? Planet? How did you classify them?

Line 255: 1 or 0.1 hectare (Line 88)?

Line 266-277: Please discuss the uncertainties on the SNL database.

Line 342: Some references are outdated (e.g., 7, 13, and 31).

Version 1:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

Dear authors,

Thank you for your paper titled "Overlooked deforestation from global mining activities in the 21st century". After the first review round, the paper is written well and makes an excellent contribution to the mining and responsible sourcing of minerals debate, i.e. with regards to the consideration of small scale mining.

My background is in mining and sustainability, hence a number of my comments below are based on this background, i.e. concerning your use of mining related terms, which I think can be improved. I also used the opportunity of Springer to involve an early career researcher, one of my PhD students, in this review. The comments are also below.

In general, you incorporated the comments of the first-round reviewers well. I think, however that your paper needs another revision. See especially the comments concerning lines 94 and 96 below.

Overall: there are still a number of grammatical and spelling mistakes in the text that need correction, e.g. Line 45: "A plenty of case studies have documented". Delete "A". Line 47: "ecosystems", etc.

Line 31: I suggest you replace "mining commodities" with the term "mineral products".

Line 32: "..., supports billions of people in mining-related activities, ... This is the first time, that I hear of billions of people IN mining. Should this be millions (see line 204)? Or do you consider indirect impacts from mining activities– if so, I suggest you replace "in" with "through".

Line 94/95: "...therefore, mining activities in shafts are not included here". There is no such thing as mining activities in shafts! Shafts (as well as portals and ramps) are only ways on how to enter an underground mine. So this raises a question: Do you mean that underground mines are not included? Or just their access points? This is important, because whilst underground mines have generally smaller land footprints (also depending on the mining method, i.e. use of backfill (little to no impact on the surface) vs. caving (large subsidence areas with high impact, see e.g. Kiruna in Sweden)) than surface mines, they still have a footprint, i.e. they also have processing and waste / tailings storage facilities, etc. This is also relevant for line 224 onwards and 303, and is not clear to me from the description there as well. Please clarify.

Line 96: "We identified a total of 230,068 undocumented mines...". I suggest you delete "undocumented" in this sentence- as you then explain that these mines are just not in the S&P database. After this explanation, using the term is fine for me. Otherwise, "undocumented" has this unwritten link to "informal" and/ or "illegal" right away, which is certainly not the case for Europe, where the high percentage of "undocumented mines" are simply small mines and quarries that are too small for S&P but certainly they are permitted and hence not illegal or informal. In line 184 you make exactly this link yourself and talk of "undocumented informal mining operations". "Undocumented" in S&P and "informal" are not the same thing! And then, in line 191, you only use the word "informal" anymore. Did you generally analyze the "undocumented" mines in such a way that you are able to say that they are "informal"? In lines 199-204 making this link is ok for me, as there you do so specifically and supported by references.

Line 143 onwards: Do all four of the datasets used for your analysis provide you with the same quality of commodity data or how did you align them? E.g. S&P gives you main products and by-products. Gold can be a by-product of copper mining – considered for copper only? Did you then use allocation and if so which method? How is that in the other databases? Lines 335 onwards in the methods description does currently not explain these aspects – this should be added.

Comments from the early-stage researcher:

Lines 96-102 & 113-124 & 300-301: It is recommended that the results section be reviewed, since there are inconsistencies between the values reported in the text and the data presented in the corresponding tables.

Lines 306-307 & 362-364: It would be useful to include a justification on the sample size used for the accuracy assessment, both in the selection of the 11,878 mining polygons and in the 2,200 deforestation pixels analyzed. On what criteria is the determination of these values based and why was a larger sample size not chosen?

Line 320/365: It is recommended to correct the typographical error on line 320, where "23,6028" appears to be misspelled.

Also, in line 365, it would be helpful to clarify the use of the term 'Medoid'. Does it refer to a median-based composite?

Lines 364-378: It would be useful to clarify the methodology used to define deforestation: was a specific NDVI/NBR/EVI threshold applied, and was its impact on the classification evaluated to avoid errors? In addition, it is recommended to indicate if the Landsat images were atmospherically corrected and if topographic effects in mountainous regions were considered.

Best regards,
Michael Tost
Professor for Sustainable Mining Technology
Chair of Mining Engineering and Mineral Economics
Montanuniversität Leoben, Austria

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)

PREAMBLE

Firstly, I would like to apologise for the additional time it has taken me to render a review. I had been unexpectedly indisposed for health reasons for a few weeks, and I hope you will forgive the delay on the spaces.

All accounts indicate that I have reviewed and manuscript that has already been revised. All the same, and perhaps annoyingly for the authors, my review encompass is all elements of the manuscript, not just the revised portions. As the manuscript is new to me, I address all its facets without discrimination, as they are all new to me.

MAIN COMMENT

The manuscript is, in effects, a simple or, better terms, straightforward overlay analysis. Two datasets are gathered up, and overlaid to see how they intersect. I refer here specifically to a 'new' composite data on mining, produced by combining existing datasets, and on the other hand's a well-worn global dataset on gross forest loss over the last 2 decades. There's nothing wrong with this approach, nor with the data brought to bear here. My main preoccupation when reviewing the study was to ensure that's messages were clear and findings were sufficiently rigorous and detailed.

I observed no fatal flaws in the data or analysis per se - nothing that would unquestionably prevent publication, at least. That said, there are many points that are unclear, sometimes seemingly in part due to English being perhaps a second language for some authors, I presume, and in part for other unknown reasons, speculatively and unfamiliarity with certain datasets or reporting standards. These points of lack of clarity, sometimes bordering on under-reporting. This will have to be addressed before anything is actually publishable.

So I am recommending 'major revision' to clarify and tighten up reporting and a few key analytical details of the work already done. I don't recommend any new analysis per se, notwithstanding one elaboration of an analysis already realised, in relation to protected areas.

In the following texts I raise points as I observe them in the manuscript, in order of reading them, also organized by their headings and subheadings. My comments do not appear in priority order. I do however indicate in my text which issues are relatively important, by saying just so in my individual comments, below.

MAIN RESULTS

> A GLOBAL INVENTORY...

1. In the first few paragraphs of the results, I note that it is reported that's over 80% of mining areas in the authors composite mining data sets are supposedly undocumented. This means that the mines were not recorded in the S&P mining dataset the authors adopted to differentiate between documented and undocumented mines. It is impossible for me to believe that 81% of mining areas in Europe are unknown or unregulated by local governments. This, of course, is not what 'undocumented' means, in the sense that the authors use the term, though it still may be fairly the reader's initial impression. So, on this basis, I become sceptical of the ultimate worth of the S&P mining data set, even if it is, as the authors allude, the best available global data set on mining activity.

2. The authors reports that 139 protected areas, plus an additional simply sent from the wilderness areas, are intersected or otherwise overlaid by mining areas as they observe them in their composite dataset on mining areas. The authors report that these protected areas are of "the highest category and protected area", but do not give any criteria by which category of protected area is defined, or even what it means. Protected areas are commonly categorised by six or so categories the IUCN (IUCN protected area categories - Wikipedia), but these categories are often not listed for many protected areas in the world database unprotected areas the author is used. All the a fraction of all protected areas in Africa, for example, have attribute data on the IUCN management category, so it is doubtful to me that this categorical attribute was used to describe protected areas as reported, and I am left with no other guests as to how protected areas were described in terms of their management category. The authors seem equally unclear, and the reporting on their selection and categorization of protected areas in the methods section effectively said nothing on the topic or otherwise. This must be corrected.

3. A related point of interest to me and my field, regarding protected areas impacted by mining, is that there does already exist a global database on this phenomenon - PADDDtracker.org. This online database is compiled on the basis of any and all reports and datasets on phenomenon like mining or logging that effect protected areas, leading to their degradation, downgrading in status, or de-gazettment. Reports and listings in the online data set are vetted by an expert panel, not just random notes by citizen scientists. The author is really should, even must, compare their list of protected areas impact by mining against that compiled by this online database of the same.

>DEFOR FROM MINING...

4. Perhaps the biggest source of uncertainty for me is that what qualifies as a so-called "mining area" is not so transparent. Seemingly, no single standard or criterion to describe a spatial area as dedicated to mining, rather than simply being in proximity of mines, has been applied. This of course reflects the fact that the authors' composite data on mining areas is derived from multiple independent datasets, so at the very least consistency in this respect cannot be expected. But even the authors' description of these data in the relevant section of the methods text does not detail well enough for my liking the criteria by which a given area surrounding mining activity but itself not dedicated to mining activity may be included as a "mining area". In effect, the criteria and/or distance threshold that defines lands around mining activity as part of the "mining area" of observation here remains an open question. I am, of course, reasonably presuming that the entirety of a nominal "mining area" is not, in fact, dedicated to mining (hole digging) itself.

a. The reason the spatial definition of a mining area (i.e., mining polygon) matters for this analysis is that the authors quantified deforestation within the spatial bounds of a so-called mining area (polygons). Hence, if the bounds of these mining areas are liberally defined, Relative to the actual extent of explicit mining (hole digging), then the estimates of deforestation attributed to mining can likewise be liberal. In many areas of mining I know in central Africa, mines attract new settlers, and so settlement expansion, then also informal / artisanal mining around the corporate mines, new agriculture and

related forest conversion, and so on - basically the whole gamut of human-driven forest loss. Not all of this forest loss is related to the mining itself, except perhaps indirectly. So, the more liberal the normal mining areas of this study were defined spatially, the more of such secondary deforestation may be captured and attributed to the mining itself, again somewhat liberally. About 1/3 and 1/9 of the mining areas defined by the authors' composite dataset are greater than 1 km² and 5 km², respectively, as shown in Table S2, while a select few can be as large as 50 km², suggesting that liberal definitions of mining areas may apply here, even if to a minority of the mines observed. To assuage all preoccupation on the part of your readers, I would ask you to comment and clarify, for each individual input dataset, the criteria, means, scales etc. by which the nominal mining areas of your study were delineated spatially around the actual mining activity of interest (hole digging).

>GOLD AND COAL

5. As a suggestion for future research, I would urge the authors to explore whether the increasing then decreasing trends they observe for some mining related deforestation track the global prices for the mined commodities in question.

6. You write "Our results indicate that between 2001 and 2012, 66.20% of deforestation in mining areas was attributed to mines extracting minerals for renewable energy production (see Methods, Fig. 4b)." The language here is too presumptive that the minerals extracted will be used for renewable energy production. The language does not give due recognition to the fact that these minerals might, or might not, be used for renewable energy production. In reality, one would imagine that the bulk of most minerals are not used for this purpose, even when appropriate for that purpose. Please revise accordingly.

METHODS

7. There is no such thing as the "Global Forest Watch dataset". The data you mean to refer to, notably uncited at the beginning of your methods text, are typically referred to as global forest change data from the university of Maryland, or similarly global forest loss data from the university of Maryland. Global forest watch is merely a website that disseminates these data.

>A COMPREHENSIVE DATASET...

8. The large paragraph at the outset of this subsection is written somewhat awkwardly, and not very clearly. Here, but generally everywhere to a lesser degree, a thorough re-checking by a native English speaker would be of benefit. Some of the revised sentences is blue do not make any sense at all here. One feature lacking clarity for me here was the listing of each of the four input datasets. I strongly suggest that you list them here listwise, as by numbered bullet points, rather than within the current large block of text.

9. In regards to your accuracy assessments of your composite dataset on mining areas, you mention a SPX data set. But this SPX data was not previously defined or introduced. Did you mean to write 'SNL' and/or 'S&P'?

>DEFINITION OF UNDOCUMENTED...

10. The authors seemingly refer to the 'SNL' and 'S&P' dataset interchangeably, causing some confusion. Please check that you used a consistent term when referring to the data set in question.

11. This statement is unclear: "However, mining activities located outside project boundaries are generally not covered by the database^{24,45}" What are project boundaries in the data, and how do they matter to the present analysis?

12. This sentence is unclear: "Similarly, scan reporting mining and informal mining activities such as artisanal and small-scale mining areas typically missed in S&P database. ". So is this one: "The mining areas interacting with SNL mining sites"... Are you referring to spatial intersection by the term 'interacting with'.

13. Finally, this sentence is also unclear, but the lack of clarity here seems much more weighty than above: "We categorized the mining polygons based on the nearest-neighbor mining activities as reported in the SNL mining database.". Nearest neighbour mining activities reported in the SNL database seems to be a key term, wholly undefined notably, and one that determines the "categorization" of the mining polygons in your main composite dataset, This categorization is also unclear to me. What are we categorising? That is what are the categories? Perhaps documented versus undocumented?. But returning to the term "nearest neighbour", having a definite spatial meaning, the sense I get from your text is that your composite dataset was intersected spatially with the SNL database, but some spatial liberties were taken to allow for intersections between elements that were not, in fact, strictly spatially coincident per se. For example, a mine in your composite that datasets could thus intersect with a mine in the SNL database, and be "categorised", Even if the two minds were some distance apart from each other. Perhaps a few 100 metres, or perhaps a few kilometres, your text does not say, but it really should specify this, and clarify what the upper distance threshold was in this approach to spatial pairing of mines.

> MINING DEFOR

14. You write: "The result demonstrates that the deforestation in mining areas, as mapped by TMF (10,403.42 km²), closed aligns with GFW mapped deforestation (11,880.64 km²), showing an uncertainty of 12.43% in tropical regions (Supplementary Fig. S10)." . It is unclear what the uncertainty pertains to. ...Upon re reading, I now understand that the 12%

reported here is a differential. The area deforested according to the TMF data differs from the area deforested according to the GFW data by 12% of the ladder area. Please clarify your reporting key.

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

Dear authors,

thank you for considering and incorporating our comments. We think that the manuscript now reads really well, we offer no further comments and think that the manuscript can be published.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)

----SUMMARY OPINION----

The authors' responses have largely satisfied my earlier queries. I recommend publication, following some final housekeeping, namely the revision of a few terms, the inclusion of a few caveats, and the reporting of some simple data already held by the authors.

My comments below are in no particular order, save for 'major' and 'minor'. Here, 'major' comments are not those requiring 'major revision', just the most important of comments I offer here.

----MAJOR COMMENTS----

1. Documented vs Undocumented Mines

Comment no. 1 of my earlier review raises the issue of the use of the terms 'documented' and 'undocumented' in reference to the mining areas (polygons) compiled by the present study. A mining area is said to be 'undocumented' where it does not spatially match a mining point from the separate SNL global mining database, taken here as the global 'gold standard' spatial database on mining activity. The authors have revised to explain this definition, and while clearer, I still take issue with the term 'undocumented'. The term 'undocumented' implies that such a mining area is 'unknown' or 'unregulated' by local authorities, not simply omitted from the SNL database. Again, the authors have clarified this explicitly – an undocumented mining area is one merely unrecognised by the SNL database, and may still be subject to local regulation – yet the 'undocumented' term simply says, or rather implies, that opposite, inadvertently. I strongly recommend the use of more appropriate terminology, instead of 'documented' and 'undocumented'. For example, 'recognized' vs 'unrecognized', 'observed' vs 'unobserved', or 'recorded' vs 'unrecorded', would be better. Doubtless the authors can think of even better terms.

2. --In Page 4 Lines 106-118--

To write to greater clarity, please use the term "IUCN Management Class" before referring to a given protected-area designation, as thusly: 'Protected areas with IUCN Management Class of Ia (strict nature reserves) or Class Ib (wilderness areas) are something something...'.

In this same text, you wrote:

"Noted, 5550 other protected areas have also observed mining activities; however, these areas are not reported or assigned management categories on WDPA, demonstrating the potential extensive impacts of mining operations on protected areas worldwide."

These 5550 other protected areas (PAs) are half of all PAs intersected by your mining areas. It would be worthwhile to further report on the nature of these PAs. While these PAs will not have IUCN Management classes recorded in the WDPa, you may still at least identify those that are designated as 'national parks' - typically, PAs with such a designation have a high level of protection, consistent with IUCN Class II. In the WDPa, one of its attribute fields, possibly [Desig] (i.e., designation), or [Desig_Eng] (i.e., designation in English), will specify whether a given PA is a national park or some other designation (e.g., forest reserve, game reserve, etc.).

In acknowledging that many PAs are not described by any IUCN Management Class in the WDPa, you may cite Sloan et al. (<https://doi.org/10.1111/aje.12377>), who quantify how the WDPa omits IUCN Management Class information for something like 80% of all African PAs. This fact raises the strong possibility that the omissions of IUCN Management Class data in the WDPa is greatest in 'developing countries', where environmental regulations for sustainable mining are probably weakest. This is a valid point for your discussion.

In this same text, you wrote:

"Furthermore, 2,163 protected habitat or species management areas overlap with mining areas, and 2,184 protected nature landscape and seascape areas have observed mining activities"

Please cite the corresponding IUCN Management Classes for these two classes of protected areas, as above.

3. --Methods--

-On the use of near-neighbour algorithm to match mining-area polygons compiled by the authors to mining commodities recorded for mines in the SNL database.

Comment 13 of my earlier review addressed your use of the nearest-neighbour (NN) algorithm. I understand your response. Your use of the NN algorithm is not unreasonable per se. I however would certainly have specified an upper limit on the allowable distance between a given mining area and a given SNL mining commodity, when establishing linkages between them via the NN algorithm. You did not. As a result, it is possible that some mining areas (polygons) in your study were assigned mineral commodities according to SNL spatial data that were, in fact, quite distant, even if also 'nearest'.

Your run of the NN routine should however at least have recorded the Euclidean distance between each mining area (polygon) and its nearest SNL data point. Given this data, I ask you to increase transparency by reporting the average, median, and standard deviation distance between each mining areas (polygons) and each nearest SNL data point, for those mining areas submitted to the NN algorithm.

----MINOR COMMENTS----

4. --Methods--

The term "suage" appears. It is unfamiliar to me, as it will be to 99% of readers. It is seemingly misused in the text. Please replace it.

5. --Methods--

-On the double-counting of PAs intersected by mines-

In your Methods text, it will be worth noting that counts of protected areas (PAs) intersected by mines will be slightly inflated, due to double counting, in some cases. The reason for this double counting is that, in the WDPa, each discrete designation of a PA is mapped by its own spatial polygon. If a given protected area has two separate designations, it will have two separate and highly congruent polygonal records in the WDPa. For example, where a National Park is also separately designated as a UNESCO World Heritage Site, the same area will effectively be mapped by two polygonal records in the WDPa - one for the National Park, and another for the World Heritage Site. Given this fact, your GIS overlay analysis would have identified two PAs intersected by the same mine, not just one, in this example above; hence, the double counting of PAs I mention.

In my experience, this double counting tends to be most common in the WDPa where national parks are also designated as some other, typically high-level protected area, as in the example above. However, the possibility of double counting is not restricted to PAs of high levels of conservation. Polygonal record in the WDPa can be messy and often partially overlap each other. Sloan et al., cited above, confronted this (inevitable) issue of double counting in Africa. At the global scale of your study, resolving this issue is probably not feasible. The only workable approach would be to spatially 'dissolve' all PAs in the WDPa into spatially discrete polygons. Doing so would however remove all attribute data from such polygons, thus providing a revised number of intersected PAs, now free of double counting, but absent any information on their IUCN Management Classes or similar. Again, Sloan et al. probably adopted this approach of spatial dissolving to ensure accurate counts of intersected PAs - it's been a while since I wrote that paper, but I have at least implemented that approach somewhere.

6. --Methods--

-On the Global Forest Cover data used-

These data are cited by reference no. 12. This citation is the correct citation of the study that first created these data, but it is not an adequate citation of your source of the data themselves. Since the 2013 study by Hansen et al. (reference no. 12), the data have been updated and revised. Please therefore also cite the online source (url) of the updated data you used, including the observation years of the data, its version (e.g., version 2, version 3...), url, etc., for transparency.

7. --TYPO--

Line 266-267, reports "TME-based deforestation". 'TME' should be 'TMF'.

=====

Sean Sloan
Canada Research Chair (T2)
Human Dimensions of Sustainability
Dept. of Geography
Vancouver Island University
Office: Bld 180 Rm 220
Tel: 250-740-6128
sean.sloan@viu.ca

(Remarks on code availability)

NIL

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response letter

> We appreciate the editor and reviewers for their time spent in assessing this manuscript and for their thoughtful and valuable comments and suggestions, which have significantly enhanced the robustness of our study and improved its clarity and implication outreach.

> Please find our point-by-point response to the comments below. To make it easy to track our revisions, all modified text is marked in blue in the revised manuscript.

Reviewer #1 (Remarks to the Author):

The scientific questions posed by the authors are highly pertinent, and providing satisfactory answers would significantly elevate the quality of the work. However, only the first question—"What is the geographic distribution of different mining activities globally?"—has been addressed sufficiently to merit approval. In contrast, the second and third questions—"To what extent do mining activities contribute to deforestation and associated carbon emissions?" and "Which types of mining have been the primary drivers of deforestation over the years?"—rely heavily on the Global Forest Canopy Loss (GFCL) product developed by Hansen et al. (2013). This product was designed to monitor deforestation at a global or regional scale, with validation procedures aligned to such scales. However, Zhang et al. apply this product to much smaller, localized areas, phenomena study.

For instance, data from Table S2 reveals that approximately 20% of the mining polygons in their study are smaller than 1 hectare. This translates to just ~11 pixels, a scale at which the quality metrics of Hansen et al. (2013) may not be reliable. The Hansen product was validated using pixels representing forest cover, deforestation, and regrowth (over 5 m height) across five global biomes. The GFCL validation process was not designed to estimate deforestation at scale caused by anthropic activities such as mining. I recommend that the authors conduct a sampling exercise using appropriate statistical models to estimate the deforested area directly caused by mining activities (see: <https://www.sciencedirect.com/science/article/pii/S0034425714000704>). If the error margins in their deforestation estimates prove unsatisfactory, I further suggest that they consider independently estimating deforestation, bypassing the Hansen et al. (2013) product altogether.

> Thank you for your insightful comments and constructive suggestions. Yes, we fully agree with your raised point regarding the accuracy and uncertainty of the Global Forest Watch (GFW) dataset, which will be critical to the findings of

our second and third questions, particularly in small-scale mining areas.

To address this issue, we now have conducted extensive analyses to validate the overall accuracy and robustness of deforestation in mining activities our estimated, specifically:

1. Following your suggestion, we have added an independent assessment of GFW-based deforestation in mining areas from visually interpretations using Landsat time series. The result indicates that the overall accuracy of GFW-based deforestation in these areas exceeds 82%.
2. We also analyzed deforestation in mining areas within tropical regions using Tropical Moist Forest (TMF) dataset for the same period 2001-2023. These two independent datasets assessed the uncertainty of deforestation in mining areas, revealing that the uncertainty of deforestation in tropical mining areas is approximately 12%.
3. Furthermore, we compared deforestation across various groups categorized by mining sizes, verifying that small-scale mining areas cover only a small proportion and contribute minimally to the overall assessment.

The detailed revisions are as follows:

➤ **Independent estimation of accuracy for GFW deforestation in mining areas**

We performed an independent assessment of GFW-based deforestation in mining areas from visually interpretations using Landsat time series. Specifically, we randomly selected 2000 deforested pixels and 200 non-deforested pixels in mining areas across various regions using stratified random sample methods (**Supplementary Table S5**), which is a much larger sample size than the original assessment of global product (v.1.0, $n = 1500^1$). We then visually interpreted with Landsat time series to evaluate the accuracy of GFW-based deforestation in mining areas. The overall accuracy in various regions exceeds 82% (**Supplementary Table S6**). We believe the results from GFW forest loss can support the accuracy of deforestation directly caused by mining activities. This information has been added to the Discussion (Page 7, Lines 230-241) and the Methods (Pages 10, Lines 356-377), which are duplicated as follows.

Page 7, Lines 230-241 (Discussion):

“Second, the primary dataset we used to assess deforestation is the high-resolution global forest cover change maps that delineate tree cover change at a spatial resolution of 30 m provided by Global Forest Watch (GFW)¹¹. The accuracy and uncertainty inherent in this dataset are sensitive to the reliability

of associated findings. To ensure robustness, we have performed extensive validations. On the one hand, we conducted an independent assessment of GFW-based deforestation in mining areas through visual interpretations using Landsat time series. Specifically, we randomly selected 2,000 deforested pixels and 200 non-deforested pixels in mining areas across various regions using stratified random sampling methods (**Supplementary Table S7**). This represents a significantly larger sample size compared to the original global product assessment (v.1.0, $n = 15,001$). We then visually interpreted these samples using Landsat time series to evaluate the accuracy of GFW-based deforestation data in mining areas. The overall accuracy across various regions exceeds 82% (**Supplementary Table S8**)."

Page 10, Lines 357-377 (Methods):

"To assess the accuracy of GFW forest loss in mining areas, we perform an independent assessment of deforestation in mining areas using stratified random sample methods (**Supplementary Data**). We randomly sample 2000 deforested pixels and 200 non-deforested pixels in our mining areas across six continents (**Supplementary Fig. S7**), which represents a much larger sample population than the original assessment of global product (v.1.0, $n = 1500$)¹¹. Specifically, we randomly selected 2000 deforested pixels in mining areas across six continents, based on the percentage of deforestation areas (**Supplementary Table S7, Supplementary Fig. S7**). We then aggregated the annual medoid values for Landsat time series (masked for cloud cover) and calculated the time series for the Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) from 2000-2023 using Google Earth Engines (**Supplementary Fig. S8**). We visually interpreted these pixels to validate whether deforestation occurred and when deforestation took place using Landsat imagery with true color visualization (RGB) alongside the annual time series of NBR, NDVI and EVI (**Supplementary Fig. S8**), which are widely used indicators for monitoring forest cover⁴⁹. We subsequently calculated the error matrix, which includes overall accuracy (OA), user's accuracy (UA) and producer's accuracy (PA). From our assessment, we estimate the OA, UA and PA for GFW deforestation in mining areas. In general, the OA are > 82%, UA are > 98% and PAs are > 81% in each continent, underscoring the robustness of GFW forest loss in mining areas (**Supplementary Table S8**). The uncertainty of GFW loss year is 0.80, Root Mean Square error (RMSE) of the loss year is 2.60 years (**Supplementary Fig. S9**), confirming that there is no significant bias in the observed deforestation changes in mining areas over time."

Supplementary Table S7. The number of samples used for validation in various continents, employing a stratified random sampling method.

Continent	Percentage of deforestation (%)	Number of deforested samples	Number of non-deforested samples	Total samples of validation
<i>Africa</i>	8.10	162	16	178
<i>Asia</i>	40.01	800	80	880
<i>Australia</i>	5.05	101	10	111
<i>Europe</i>	6.26	125	13	138
<i>North America</i>	16.63	333	33	366
<i>South America</i>	23.96	479	48	527
Global	100	2000	200	2200

[Figure Redacted]

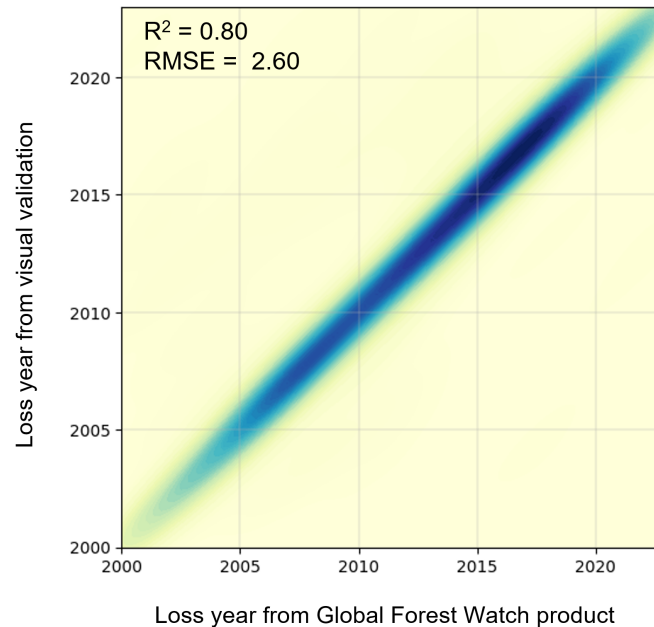
Supplementary Fig. S7. Location of sampled deforested pixels (n = 2000) and non-deforested pixels (n = 200) for Global Forest Watch forest-loss map.

[Figure Redacted]

Supplementary Fig. S8. Examples of validated GFW deforestation using Landsat time series from 2001 to 2023. **a**, The validation point overlaid on Google Earth Satellite Imagery (2024, Maxar Technologies, Map data 2024). **b**, The loss year of deforestation from Global Forest Watch within the mining areas from 2001 to 2023. **c**, Annual medoid value for NBR, NDVI, EVI time series for the validation pixel. **d**, Landsat true color visualization (RGB) from 2000-2023.

Supplementary Table S8. Number of samples and accuracy for Global Forest Watch deforestation in mining areas across continents.

		Reference					
GFW forest loss map	Continent	Loss	No Loss	Overall accuracy (OA)	Producer's accuracy (PA)	User's accuracy (UA)	
	Africa	Loss	140	22	85.96%	86.42%	97.90%
		No loss	3	13			
	Asia	Loss	797	83	98.41%	90.57%	98.64%
		No loss	11	69			
	Australia	Loss	82	19	82.88%	81.19%	100.00%
		No loss	0	10			
	Europe	Loss	102	23	83.33%	81.60%	100.00%
		No loss	0	13			
	North America	Loss	303	30	90.98%	90.99%	99.02%
		No loss	3	30			
	South America	Loss	413	66	87.29%	86.22%	99.76%
		No loss	1	47			
	Global	Loss	1837	243	91.77%	88.32%	99.03%
		No loss	18	182			



Supplementary Fig. S9. The density scatter plots for the loss year from Global Forest Watch dataset and visual validation for deforested pixels.

➤ **Comparison of GFW deforestation using Tropical Moist Forest (TMF) dataset**

GFW is the only product available for global-scale mapping worldwide, however, the Tropical Moist Forest (TMF) specifically covers tropical regions. To further address the uncertainty of GFW-based deforestation in mining areas, we have included additional analysis of deforestation assessment in tropical areas using the TMF dataset. As shown in **Supplementary Fig. S10**, the trends in deforestation from 2001 to 2023 are consistently aligned between the two datasets. Additionally, the uncertainty in deforestation in mining areas across tropical is estimated to be 12.43%, indicating a plausible accuracy. We have incorporated this analysis into our revised text (Page 7, Lines 241-248; Pages 10-11, Lines 377-388), which is duplicated as follows:

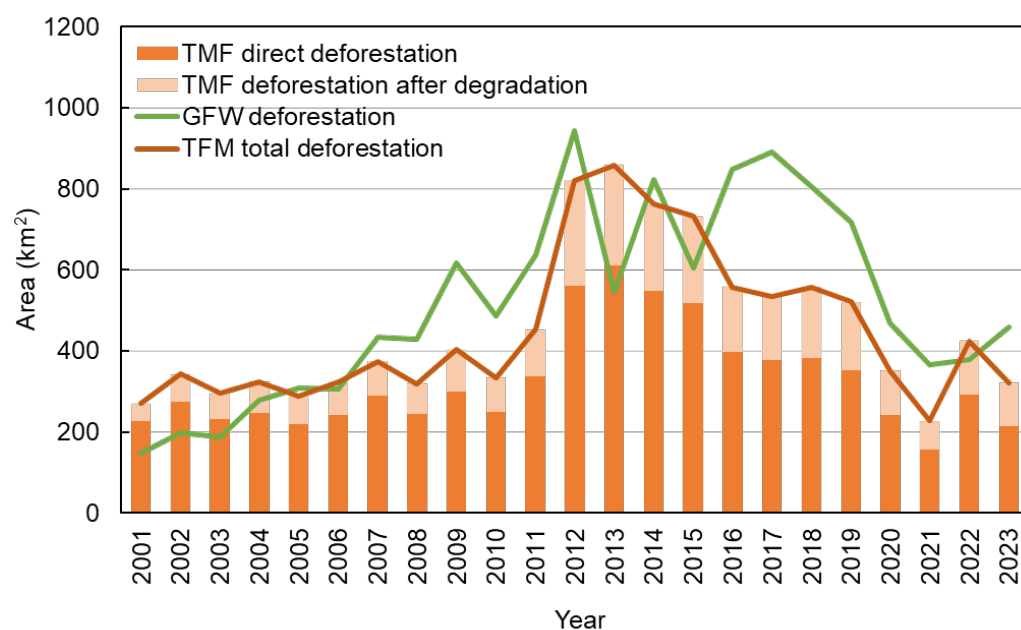
Page 7, Lines 241-248 (Discussion):

“On the other hand, we benchmarked the GFW data against the Tropical Moist Forest (TMF) dataset⁴¹ using a time-series cross-comparison. The deforestation trends from 2001 to 2023 are consistently aligned between the two datasets, with an uncertainty estimated at 12.43% in tropical regions, suggesting plausible accuracy. Given this satisfactory accuracy and consistent temporal trends, future work with high-resolution surveillance could still be invested such as incorporating satellites such as 10-meter Sentinel-1/2 and 3-meter PlanetScope to better capture small-scale mining activities across

different regions.”

Pages 10-11, Lines 377-388 (Methods):

“Additionally, we compared deforestation in mining areas across tropical regions using the Tropical Moist Forest (TMF, ‘TMFv2023’ version) dataset from the European Commission Joint Research Centre (JRC)³. The TMF dataset maps disturbances in tropical humid forest from 1990 to 2023 derived from the Landsat archive collection. The TMF dataset employed an algorithm that detects forest disruptions for each 30-meter pixel using available Landsat image, classifying the status of forest based on the duration these disruptions are visible⁴. The dataset identifies various types of forest disruptions, including both direct deforestation and deforestation after degradation. We compared the total deforestation reported by TMF dataset with that from GFW dataset in mining areas for the same period, from 2001 to 2023 (**Supplementary Fig. S10**). The result demonstrates that the deforestation in mining areas, as mapped by TMF (10,403.42 km²), closed aligns with GFW mapped deforestation (11,880.64 km²), showing an uncertainty of 12.43% in tropical regions (**Supplementary Fig. S10**).”



Supplementary Fig. S10. Deforestation in mining areas within tropical regions using Tropical Moist Forest (TMF) dataset and Global Forest Watch (GFW) dataset from 2001 to 2023.

➤ Deforestation assessment in small-scale mining areas

Regarding the point about small-scale mining areas, it should be clarified that the small-sized mining polygons, specifically those less than 0.1 km²,

accounted for only 3.05% of total mining areas and contribute 2.10% of total deforestation (415.66 km²). Therefore, these small-scale mining areas would not considerably impact on the overall assessment of global mining activities.

Supplementary Table S10. Total areas of mining polygons and mining-related deforestation from 2001 to 2023 categorized by different mining sizes.

Mining size (Area: km²)	Number of mining polygons	Mining area (km²)	Percentage of mining area	Mining-related deforestation 2001-2023 (km²)	Percentage of deforestation
<0.01	47828	232.05	0.16%	13.79	0.07%
(0.01, 0.05]	72990	1884.51	1.29%	167	0.84%
(0.05, 0.1]	32607	2334.54	1.60%	234.87	1.19%
(0.1, 0.5]	53732	12174.14	8.34%	1259.39	6.37%
(0.5, 1]	12015	8425.08	5.77%	927.92	4.69%
(1, 5]	12662	26641.16	18.26%	3270.99	16.55%
(5, 10]	2035	14235.77	9.76%	1879.96	9.51%
(10, 50]	1848	37750.5	25.87%	5113.96	25.87%
>50	311	42222.78	28.94%	6897.47	34.89%
Sum	236028	145900.5	-	19765.35	-

Additionally, there are uncertainties about how the authors account for mining areas located in regions with original natural vegetation less than 5 meters in height. Though not classified as forests in the conventional sense, these areas should still be accounted for, and the authors should consider methods to estimate the impact on natural vegetation in such regions.

> Thank you for your insightful comment regarding the differentiation of vegetation types in mining areas, particularly natural vegetation less than 5 meters in height. On the one hand, to clarify, our study primarily focuses on deforestation driven by mining activities. The definition of “forest” used in this study aligns with the GFW dataset. Other natural vegetation types, such as grassland, shrubs and lower vegetations, are not included in this study’s scope due to two main considerations: (1) These vegetation types generally regenerate more quickly, introducing uncertainties when estimating their impact on a global scale. (2) Some deforested areas transition into lower vegetation, such as shrubs or grasslands, complicating the assessment of mining impacts with higher uncertainties in time-series analyses. We have clarified this point in

the discussion (Page 7-8, Lines 251-262).

On the other hand, we appreciate the suggestion that mining activities may impact original natural vegetation under 5 meters in height. Therefore, we further incorporated the 30 m time-series land-cover dynamics monitoring product (GLC_FCS30D) (Zhang et.al., 2024) to quantify the impact on those lower-height vegetation from 2000 to 2022, including shrubland and grassland (Supplementary Table S7). The results indicate a loss of 1,253 km² of shrubland and 1,589 km² of grassland, which were converted to bare areas as a direct result of mining activities during this period from 2000-2022 (Supplementary Table S7).

Supplementary Table S9. Land transition of shrubland and grassland in global mining areas using GLC_FCS30D land cover map (Zhang et.al., 2024) from 2000 to 2022.

		Land cover 2000	
		<i>Shrubland</i>	<i>Grassland</i>
Land cover 2022	<i>Cropland</i>	945.52	1348.35
	<i>Forest</i>	812.64	616.41
	<i>Shrubland</i>	8716.74	1107.12
	<i>Grassland</i>	1166.64	7868.57
	<i>Tundra</i>	38.65	15.37
	<i>Wetland</i>	149.14	178.31
	<i>Impervious surface</i>	683.86	1123.99
	Bare areas	1252.64	1589.12
	<i>Water body</i>	241.04	292.05

Taken together, we have consolidated our discussion to address this point (Pages 7-8, Lines 251-262).

Pages 7-8, Lines 251-262 (Discussion):

“Additionally, this study assessed deforestation in forests exceeding 5 meters in height due to two considerations: (1) vegetation below 5 meters often regenerates more rapidly, introducing uncertainties when estimating global impacts; and (2) some deforested areas transition into lower vegetation types, such as shrubs or grasslands, complicating time-series assessments with higher uncertainties. However, mining activities may impact natural vegetation below 5 meters. To address this, we utilized the 30-meter time-series land-cover dynamics monitoring product (GLC_FCS30D)⁴² to quantify losses in shrubland and grassland from 2000 to 2022. Our analysis revealed a loss of 1,253 km² of shrubland and 1,589 km² of grassland, which were converted into

bare land as a direct result of mining activities during this period (Supplementary Table S9). Taken together, some next-step initiatives need to consider the indirect and cumulative forest and other vegetation impacts, beyond the immediate deforestation observed at surface levels³."

References

[1] M. C. Hansen et al., High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342,850-853(2013). DOI:10.1126/science.1244693.

[2] Zhang, X., Zhao, T., Xu, H., Liu, W., Wang, J., Chen, X., and Liu, L.: GLC_FCS30D: the first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method, *Earth Syst. Sci. Data*, 16, 1353–1381, <https://doi.org/10.5194/essd-16-1353-2024>, 2024.

A few minor points also require attention. On line 89, I believe the authors intended to write "1,215.18 thousand hectares" rather than "million hectares." To maintain consistency with the other measurements in the article, I suggest using km² instead of hectares. Additionally, please correct the unit of measurement in Table S2, where the title specifies km² but the data appears to use hectares, which is somewhat confusing.

> Thank you for pointing this out. We have revised our manuscript throughout to consistently use the units km² for all results. Additionally, we have updated Table S2 to ensure uniformity throughout. Please refer to our revised text.

In conclusion, while the manuscript is well-written and tackles important questions, the lack of robust error estimates in the deforestation data linked to mining activities weakens the validity of the results. Without addressing these concerns, the study is not yet ready for publication.

> We sincerely appreciate your insight comments, which have helped enhance the robustness of our assessment. Please refer to our refined manuscript.

Reviewer #2 (Remarks to the Author):

The work of Zhang et al. (2024) provides a new global dataset for mining activities, identifying documented and undocumented sites and quantifying the deforestation and CO₂ emissions associated with these mining polygons. Overall, the manuscript is well written and has good research questions. I have a few suggestions below to improve its understanding and outreach:

> Thank you for your detailed assessment and very positive comments. We greatly appreciate your insightful suggestions that have enhanced both understanding and outreach for our study. Please find our point-to-point response below.

The dataset considers only mining where there is land cover change, that is, artisanal small-scale and mining on shafts is unaccounted. This is partially discussed in the “Discussion” section but must be mentioned in the main text.

> Good point. We have added a dataset of the artisanal mining area, which includes 660 mining polygons across tropics, to our mining areas dataset. Additionally, the mining areas used in this study included artisanal mining activities. We agreed that mining activities in shafts areas are not accounted for. We have added these details to the main text for clarification (Page 3, Lines 87-95; Pages 8-9, Lines 299-302), as duplicated below.

Page 3, Lines 87-95 (Results):

*“We combined four global available mining area datasets to integrate a comprehensive map of mining activities, incorporating global mining footprints^{20,24}, crowdsourcing mining polygons²⁸, manual delineated artisanal mining areas⁸, and S&P commercial mining sites²¹ (**Supplementary Table S1, S2**). Our dataset includes 236,028 mining areas (each over 1000 m²) worldwide, extending over 145,900.54 km² (**Fig. 1, Tabel 1**). The overall accuracy of these global mining areas through our independent validation is 87.37% (see Methods, **Supplementary Table S3, S4**). Noted, our mining data primarily relies on published mining land use datasets obtained through crowdsourcing and manual visual inspection methods; therefore, mining activities in shafts are not accounted in here (**Supplementary Table S1**).”*

Pages 8-9, Lines 300-303 (Methods):

*“These included all above-ground mining features recognizable from satellite images or community knowledge, such as open cuts, tailing dams, waste rock dumps, water ponds, and processing infrastructure, while mining in shafts may not be accounted in here (**Supplementary Table S1, Fig.1**).”*

We have the impression during the text that the authors created the dataset by classifying images and then combined this classification with existing datasets, but this is not the case. Please clarify this throughout the text.

> Thank you for your good suggestion. We now clarify the text throughout our manuscript:

Page 2, Lines 20-22 (Abstract): *“Here, we compiled a comprehensive inventory of global mining activities, including 236,028 mining areas with an overall accuracy of 87.37% to analyze deforestation within mining areas and the associated forest carbon emissions from 2001 to 2023.”*

Page 3, Lines 87-90 (Results): *“We combined four global available mining area datasets to integrate a comprehensive map of mining activities, incorporating global mining footprints^{20,24}, crowdsourcing mining polygons²⁸, manual delineated artisanal mining areas⁸, and S&P commercial mining sites²¹ (**Supplementary Table S1, S2**).”*

Page 7, Lines 220-225 (Discussion): *“First, although the mining datasets compiled in this study represent the most comprehensive global inventory of mining activities to date—integrating multiple sources such as satellite-based mining footprints and crowdsourcing mining polygons from OSM maps—they predominately focus on land surface mining activities. These datasets are unable to explicitly identify the extent of underground impacts caused by shaft openings.”*

Page 8, Lines 266-271 (Methods): *“To estimate the impact of global mining activities on natural forests, we first integrated a comprehensive inventory of global mining areas by combining four previously published mining areas databases. We categorized these mining areas into documented and undocumented mines based on integration with S&P metals and mining database. We then analyzed the mining-driven deforestation and forest carbon emission in these mining areas from 2001 to 2023 using the Global Forest Watch dataset and forest carbon fluxes. The flowchart is shown in **Supplementary Fig. S4**.”*

Page 8, Lines 273-275 (Methods): *“We combined four mining databases from multiple sources, including global mining polygons in 2019 and Global mining footprint in 2023=²⁰, artisanal mining areas in 2023⁸ and OpenStreetMap-based mining land use in 2024 (**Supplementary Table S1**).”*

The so-called validation of the dataset consists of the visual interpretation of 2,000 polygons (as shown in Fig. S2). My concern is that this represents less than 1% of the total polygons, so it is insufficient to corroborate the more than 87% accuracy of the dataset. Moreover, how many of these polygons were

documented and undocumented? Authors must provide a more robust validation of the dataset.

> Thank you for your good comment. We now have conducted an independent assessment of 11,878 mining polygons samples, approximately 5% of the total mining polygons across various mining sizes (Supplementary Fig. S5, Supplementary Table S3). We visually checked if there were or did mining activities occurred within these polygons using Google Earth satellite imagery. The overall accuracy of mining polygons is 87.37%. We revised our texts in the revised manuscript (Page 3, Lines 91-92, Page 8, Lines 304-316).

[Figure Redacted]

Supplementary Fig. S5. Mining polygons using for validation through a stratified random sample method ($n = 11878$). A total of 382 documented mining polygons (in black) in the SPX global mining dataset were confirmed.

Supplementary Table S3. Number of mining polygons for validation using stratified random sampling by mining size and the overall accuracy.

<i>Mining size (km²)</i>	<i>Number of validated mining polygons</i>	<i>Accuracy of samples (%)</i>
<i><0.01</i>	2387	70.05
<i>(0.01, 0.05]</i>	3653	86.15
<i>(0.05, 0.1]</i>	1629	94.17
<i>(0.1, 0.5]</i>	2690	95.02
<i>(0.5, 1]</i>	603	94.53
<i>(1, 5]</i>	657	97.41
<i>(5, 10]</i>	108	100
<i>(10, 50]</i>	125	100
<i>>50</i>	26	100
Sum	11878	87.37

Supplementary Table S4. Number of mining polygons for validation by continent and the overall accuracy.

Continent	Number of validated mining polygons	Accuracy (%)
Africa	1341	84.71
Asia	3877	92.55
Australia	325	91.08
Europe	3365	82.70
North America	2033	84.80
South America	937	90.82
Global	11878	87.37

Page 3, Lines 91-92 (Results)

*“The overall accuracy of these global mining areas through our independent validation is 87.37% (see Methods, **Supplementary Table S3, S4**).”*

Page 8, Lines 304-316 (Methods):

*“Additionally, we performed an assessment to estimate the accuracy for the mining polygons using stratified random sampling. We randomly sample ~5% of the total mining polygons across different mining size groups, resulting in randomly selected 11,878 mining polygons to assess the overall accuracy of the mining datasets, which also spanned different geographical regions (**Supplementary Fig. S5**). Specifically, we classify mining polygons into different categories based on the area of mining polygon and then randomly selected 5% mining polygons across different mining sizes (**Supplementary Table S4**). We then visually inspected manually sampled polygon using Google Earth satellite imagery to confirm the presence of mining activities in each mining polygon. The overall accuracy of the global mining areas was 87.37% from these validated mining polygons (**Supplementary Table S4, S5**). In general, larger mining areas show higher overall accuracy; specifically, those mines greater than $> 5 \text{ km}^2$ showed an accuracy $> 94\%$ (**Supplementary Table S5**). Among the sampled mining polygons, 383 mining polygons were documented by SPX mining dataset (**Supplementary Fig. S5**).”*

Most of the figures are confusing, sometimes combining 8 panels, and need to be improved. Please see specific comments below.

> Thank you for your valuable comments. We have reorganized all figures for better clarity. Please refer to our reply below in detail.

Authors confuse key terms related to deforestation, as they consider deforestation, forest loss, and forest cover the same throughout the text. For

example, forest loss encompasses other causes than deforestation such as hurricanes. You must define and revise these terms.

> Thank you for your good suggestion. We have revised the manuscript to focus more on deforestation. The definition of forest we used, and the cause of deforestation-derived forest loss have been clearly defined and clarified in the main text.

Page 2, Lines 22-25 (Abstract): *“Our results reveal that deforestation directly caused by mining activities is two to three times higher than previously estimated from widely used mining datasets, accounting for 19,765 km² of deforestation and 0.75 Pg CO₂ of carbon emission in the 21st century.”*

Page 4, Lines 119-122 (Results): *“As for the continental-scale comparison, Asia experienced the highest deforestation in mining areas, accounting for 42.88% (7,907.18 km²) of deforestation (**Table 1**), followed by South America (4,735.40 km²), North America (3,287.30 km²), Africa (1,600.09 km²), Europe (1,238.00 km²), and Australia (997.39 km²).”*

Page 4, Lines 125-126 (Results): *“Throughout the 21st century (2001-2023), 175 countries worldwide experienced deforestation in mining areas (**Fig.2a**).”*

Page 4, Lines 131-133 (Results): *“Indonesia experienced the highest deforestation in mining areas, representing 21.72% (4,292.33 km²) of total deforestation and 0.22 Pg CO₂ in carbon emissions (**Fig.2c**).”*

Page 7, Lines 220-223 (Discussion) *“First, although the mining datasets compiled in this study represent the most comprehensive global inventory of mining activities to date—integrating multiple sources such as satellite-based mining footprints and crowdsourcing mining polygons from OSM maps—they predominately focus on land surface mining activities.”*

I do not agree with the Global North x Global South comparisons made by the authors. This reproduces colonialist practices that omit the fact that the majority of the mining in the Global South is consumed in the Global North. So, I suggest removing this comparison.

> Thank you for this good suggestion. We have removed the analysis and related text in the revised manuscript.

Specific Comments:

Line 17: Please remove “believed” and adjust the sentence.

> Thank you. We have revised the sentence (Page 2, Line 18), which are duplicated as follows.

Page 2, Line 18 (Abstract): *“Mining is a major driver of deforestation.”*

Lines 37-38: This is an outdated reference. Could you please check if it still stands?

> We have checked and updated these references. Please see our new references in the revised text (Page 2, Lines 35-37), which are duplicated as follows.

Page 2, Lines 35-37 (Main):

“This surge in mineral demand has led to significant environmental impacts, including disturbance of forest ecosystems and protected areas^{3,4}, biodiversity loss⁵⁻⁷, soil pollution^{5,8}, and water scarcity^{8,9}.”

References

- [1]. Sonter, L. J. *et al.* Mining drives extensive deforestation in the Brazilian Amazon. *Nat. Commun.* **8**, 1013 (2017).
- [2]. Ladewig, M., Angelsen, A., Masolele, R. N. & Chervier, C. Deforestation triggered by artisanal mining in eastern Democratic Republic of the Congo. *Nat. Sustain.* 1–9 (2024) doi:10.1038/s41893-024-01421-8.
- [3]. Gerson, J. R. *et al.* Amazon forests capture high levels of atmospheric mercury pollution from artisanal gold mining. *Nat. Commun.* **13**, 559 (2022).
- [4]. Yin, X., Ang, L. P. & Chang, Z. Y. Rare earth element mining threatens Malaysia’s biodiversity. *Science* **384**, 1182–1182 (2024).
- [5]. Bradley, S. Mining’s impacts on forests aligning policy and finance for climate and biodiversity goals. *Energy Environ. Resour. Programme* 1–29 (2020).
- [6]. Luckeneder, S., Giljum, S., Schaffartzik, A., Maus, V. & Tost, M. Surge in global metal mining threatens vulnerable ecosystems. *Glob. Environ. Change* **69**, 102303 (2021).

Lines 38-40: Now you have returned to the global scale. This is confusing.

> We have revised these sentences for better clarify and logical flow (Page 2, Lines 38-49), which are duplicated as follows.

Page 2, Lines 38-49 (Main):

“In the pursuit of economic incentives through mining operations, a considerable amount of mining activities have been reported in several African and developing countries, along with the associated environmental issues¹⁰. Mining is now regarded as one of major human-driven contributors to deforestation, second only to agriculture, infrastructure, and urban expansion¹¹. The deforestation induced by mining activities not only disrupts ecosystem

functioning but also reduces carbon sequestration capacities, posing considerable risks to global climate stability^{12,13}. A pantropical assessment reported 3,264 km² of forest was directly lost due to industrial mining from 2000 to 2019². A plenty of case studies have documented the substantial impacts of mining activities in tropical rainforest, as seen in countries such as Brazil³, Colombia¹⁴, Congo¹⁵, and Indonesia¹⁶. In light of the negative impacts of mining activities on forest ecosystem (e.g. deforestation and degradation), local, regional, and global initiatives have invested considerable effort to address these issues⁷."

Lines 42-43: Which decades?

> We revised the sentence for clarity (Page 2, Lines 43-45), which are duplicated as follows.

Page 2, Lines 43-45 (Main):

"A pantropical assessment reported 3,264 km² of forest was directly lost due to industrial mining from 2000 to 2019²."

Lines 37-45: Text flow is not good in here. It keeps shifting from global to regional scale.

> Thank you for spotting this. Same response to the previous comment. We have revised these sentences for better clarify and logical flow (Page 2, Lines 38-49), which are duplicated as follows.

Page 2, Lines 38-49 (Main):

"In the pursuit of economic incentives through mining operations, a considerable amount of mining activities have been reported in several African and developing countries, along with the associated environmental issues¹⁰. Mining is now regarded as one of major human-driven contributors to deforestation, second only to agriculture, infrastructure, and urban expansion¹¹. The deforestation induced by mining activities not only disrupts ecosystem functioning but also reduces carbon sequestration capacities, posing considerable risks to global climate stability^{12,13}. A pantropical assessment reported 3,264 km² of forest was directly lost due to industrial mining from 2000 to 2019². A plenty of case studies have documented the substantial impacts of mining activities in tropical rainforest, as seen in countries such as Brazil³, Colombia¹⁴, Congo¹⁵, and Indonesia¹⁶. In light of the negative impacts of mining activities on forest ecosystem (e.g. deforestation and degradation), local, regional, and global initiatives have invested considerable effort to address these issues⁷."

Lines 45-46: Which negative impacts?

> We specifically highlight the negative impacts of mining activities on deforestation. Accordingly, we have revised the sentence for better clarification (Page 2, Lines 45-49), which are duplicated as follows.

Page 2, Lines 45-49 (Main):

"A plenty of case studies have documented the substantial impacts of mining activities in tropical rainforest, as seen in countries such as Brazil³, Colombia¹⁴, Congo¹⁵, and Indonesia¹⁶. In light of the negative impacts of mining activities on forest ecosystem (e.g. deforestation and degradation), local, regional, and global initiatives have invested considerable effort to address these issues⁷."

Fig 1. It is impossible to evaluate the heatmap results due to the overlap with the points related to the primary commodities. Numbers are hard to read in a. What is the meaning of the numbers in d?

> Thank you for your good comments. We have refined Fig. 1 and added the details to its caption. The numbers in d represent *"The count of protected areas conflicted with mining activities"*.

[Figure Redacted]

Fig. 1. Geographic distribution of global mining activities. **a**, The heatmap of global mining areas (visualization at <https://ee-xzrscph.projects.earthengine.app/view/global-mining-watch>). **b1-b6**, zoomed-in examples of mining areas, gravel and sand mine in USA (b1), cropper mine in Peru (b2), diamond mine in South Africa (b3), brick kiln in Nepal (b4), gold mine in Myanmar (b5) and coal mine in Indonesia (b6). **c**, Documented and undocumented mining areas (see Methods) for the top largest 20 mining countries. **d**, The count of protected areas conflicted with mining activities.

Table 1. What did you mean by count? Number of mining polygons?

> Yes, we have revised the term to “Number of mining polygons” to clarify. Please refer to our updated Table 1.

Table 1. The mining areas and mining-related deforestation and forest carbon emission by documented mines and undocumented mines across different regions for 2001-2023.

<i>Continent</i>	<i>Number of mining polygons</i>		<i>Mining area (km²)</i>		<i>Mining-related deforestation (km²)</i>		<i>Mining-related forest carbon emission (Pg CO₂)</i>	
	<i>Undocu mented</i>	<i>Documen ted</i>	<i>Undocum ented</i>	<i>Documen ted</i>	<i>Undocume nted</i>	<i>Documented</i>	<i>Undocume nted</i>	<i>Documented</i>
<i>Africa</i>	22695	726	36,053	40,749	5,430	5,522	0.0202	0.0236
<i>Asia</i>	75043	1930	24,744	16,914	3,678	2,007	0.2012	0.1034
<i>Australia</i>	7124	661	19,972	13,595	1,932	2,287	0.0063	0.0038
<i>Europe</i>	70207	574	4,188	9,920	767	245	0.0188	0.0040
<i>North America</i>	38112	1507	15,609	7,604	757	2,071	0.0420	0.0320
<i>South America</i>	16887	562	6,757	2,949	279	566	0.2307	0.0600
<i>Global</i>	230068	5960	107,323	91,731	12,842	12,698	0.5192	0.2268

Fig. S2 appears before Fig. S1 in the text.

> Thank you. We have checked the text and reordered the number. Please refer to our revision.

Fig 2. There is a non-described graph inside the map. What does it mean? The figure is related to deforestation, so why did you add CO2 and budget information on it? Why divide in Global South and Global North in e? Meaning of the black line in c? In b, divide the emissions in documented and undocumented too?

> Thank you for your comments. We have refined Fig. 2 for greater clarity, as shown below. Additionally, we have removed the analysis regarding the global budget, as well as the subplot about to the Global South and North.

[Figure Redacted]

Fig. 2. Deforestation from mining areas during 2001-2023. a, Country-level deforestation in mining areas from 2001 to 2023. b, The deforestation in mining areas and forest carbon emission in different climate zones. c, The top 10 largest countries of deforestation from mining areas. d, Annual deforestation in mining areas and the fitted trend for global annual deforestation using segmented linear regression (shown in black line).

Line 109: You have CO₂ information in Fig. 2 too. Why? Is it Fig. 2 or Fig. 3?

> We apologize for the incorrect citation regarding our figure. It should refer to Fig. 2. We have revised our text accordingly (Page 4, Lines 122-124), which are duplicated as follows.

Page 4, Lines 126-128 (Results):

“Tropical forests were seriously affected, with 10,824 km² of mining-driven deforestation accounting for 0.56 Pg CO₂ in forest carbon emissions, making them as hotspots. (Fig.2b)”

Line 121: Deforestation from mining.

> Thank you. We have revised the text (Page 4, Lines 124-126), which are duplicated as follows.

Page 4, Lines 124-126 (Results):

“Cold and temperate regions also faced considerable deforestation, with 5,162 km² and 3,470 km² of deforestation from mining, respectively.”

Lines 125-128: I do not see the point of this North x South analysis. This might lead to conclusions that reflect the colonialism process.

> Thank you for your concerns. We have removed the related analysis and corresponding text.

Lines 129-130: Did you perform a statistical analysis to corroborate this so-called trend?

> Yes, we applied segmented linear regression to assess breakpoint and the trend in global annual deforestation, identifying 2012 as a breakpoint. We have incorporated the fitted linear trend in Fig 2d and clarified this analysis in the revised text (Page 5, Lines 138-141), which are duplicated as follows.

Page 5, Lines 138-141 (Results):

*“The two-decade trajectory reveals an increasing trend in global mining-driven deforestation from 2001 to 2012 (+90.26 km²/year, $p < 0.001$), followed by a decline since 2013 (-50.73 km²/year, $p < 0.001$) (**Fig. 2d**). This temporal pattern corresponds closely with the trend in global exploration budgets observed throughout the 21st century (**Supplementary Fig. S3**).”*

Fig. 3. In a, missing names and values. C appears after d in the caption. Please, clarify in the graphs that you are considering only gold mines. Not sure what f means.

> Thank you for your good comment. We have reformulated Fig. 3 and added additional details in the captions to better clarify. Please refer to our revised Fig. 3.

[Figure Redacted]

Fig. 3. Deforestation estimated from different mining types during 2001-2023. a, Deforestation with different mining commodities. **b,** The top 10 largest countries with gold mining-related deforestation. **c,** Deforestation differentiated by documented and undocumented gold and coal mining. **d,** The top 10 largest countries with coal mining-related deforestation by coal mines.

Lines 151-153: Also needs a statistical analysis to affirm this.

> Thank you for the suggestion. We have added a statistical analysis for the trends. Please refer to the refined Fig 2d and related text (Page 5, Lines 138-141):

Page 5, Lines 138-141 (Results):

“The two-decade trajectory reveals an increasing trend in global mining-driven deforestation from 2001 to 2012 (+90.26 km²/year, $p < 0.001$), followed by a decline since 2013 (-50.73 km²/year, $p < 0.001$) (Fig. 2d). This temporal pattern corresponds closely with the trend in global exploration budgets observed throughout the 21st century (Supplementary Fig. S3).”

Lines 163-137: How did you aggregate the minerals in these classes?

> We aggregated these minerals into renewable/non-renewable energy production based on their critical renewable energy production and infrastructures. The classification follows the framework established by Sonter et al.²³. Additional details have been included in the Results section (Page 5, Lines 164-166), and Methods (Page 9, Lines 334-343) to clarify this classification further.

Page 5, Lines 164-166 (Results):

*“We grouped this mining-related deforestation by the mining commodity types to categories the mining demand by renewable/non-renewable energy production²³ with the energy transition (**Supplementary Table. S4**, see Methods).”*

Pages 9-10, Lines 335-344 (Methods):

*“Additionally, we aggregated this mining by commodities types to determine the mining demand by renewable/non-renewable energy production, following the classification by Sonter et al.³⁰ (**Supplementary Table. S6**). The classification is based on the projected material demand under 2050 low-carbon energy scenarios⁴⁶. This projection primarily focused on two renewable-energy technologies: wind turbine manufacturing and solar photovoltaic installations, along with one energy-using technology: storage batteries for electric vehicles. Other technologies, such as carbon capture and storage installation, nuclear electricity generation installation, electric vehicle manufacturing, and lithium-Ion batteries, are also considered part of the renewable energy sector³⁰. The mining properties that are not critical for renewable energy are regarded as material for non-renewable energy production.”*

Supplementary Table S6. Mining properties targeting materials required for renewable energy technology and infrastructure and other non-renewable commodities, following the classification by Sonter et al.²³.

Critical for renewable energy	Not critical for renewable energy
Gold, Copper, Silver, Zinc, Lead, Iron Ore, Nickel, Molybdenum, Cobalt, Platinum, Palladium, Lanthanides, Lithium, Tin, Tungsten, Manganese, Graphite, Vanadium, Bauxites, Tantalum, Chromite, Antimony, Titanium, Niobium, Zircon, Yttrium, Scandium, Chromium, Alumina, Aluminum, Platinum Group Metals	Coal, U3O8, Diamonds, Phosphate, Potash, Ilmenite, Rutile, Borates

Line 178: Forest cover x forest loss x deforestation.

> Thank you for pointing out this. We have corrected all related terms as “deforestation”. Please refer to our previous reply.

Line 204: Missing statistical analysis to support this.

>Thank you for pointing out this. We have added a statistical analysis in Fig 2d to support this statement. Please refer to our previous reply.

Line 213: How did you derive the mining land use from OSM?

> Thank you for your question. We have added more details in Methods to clarify (Page 8, Lines 289-295), which are duplicated as follows.

Page 8, Lines 289-295 (Methods):

*“Additionally, we extracted all mining areas categorized as “quarry” in land use type from OpenStreetMap (OSM, collected by January 2024), resulting in 170,620 mining-related polygon features (**Supplementary Table S1, Supplementary Table S2**). OpenStreetMap, one of the most popular open data communities, has already contributed to filling existing data gaps, especially in socioeconomic domains^{43,44}. In OSM, a quarry is defined as an area of land used for surface extraction (open-pit mining) of mineral or organic resources like gravel, sand, coal, or other land²⁸.”*

Lines 245-248: Which images were used? Planet? How did you classify them?

> Thank you for the question. We have clarified that our mining polygons dataset was created by integrating four mining databases, which combined published mining data, rather than being classified directly through satellite imagery. Please refer to our revised text for further details (Page 8, Lines 273-297), which are duplicated as follows.

Page 8, Lines 273-298 (Methods):

*“We combined four mining databases from multiple sources, including global mining polygons in 2019 and Global mining footprint in 2022²⁰, artisanal mining areas in 2023⁸ and OpenStreetMap-based mining land use in 2024 (**Supplementary Table S1**). The global mining polygons from ref²⁴, which delineated all underground and above-ground mining infrastructure visible on satellite images. This dataset, derived visual interpretation of Sentinel-2 images for 2019, contained 44,929 polygon features covering 101,583 km² (**Supplementary Table S1, Supplementary Table S2**). However, the dataset only inspected areas within a 10 km buffer around the coordinates from SNL mining sites²⁴. Global Mining Footprint in 2023 provided by Tang et al.,²⁰ also was collected here, as a source for mining land use. The Global Mining Footprint was mapped from remote sensing analysis of high-resolution, comprising 74548 mining polygons, covering 65,585 km² of features like waste rock dumps, pits, water ponds, tailings dams, heap leach pads and processing/milling infrastructure (**Supplementary Table S1, Supplementary Table S2**)²⁰. The artisanal mining areas were visually identified and outlined through a combination of literature and popular media review, mining company and national reporting on mining activity, social media investigations, and extensive manual satellite and aerial imager review⁸. This dataset includes 660*

mining polygons that are significant for artisanal mining activities in tropical regions, which primarily focus on river mineral mining, including selected global and demand mining⁸ (**Supplementary Table S1, Supplementary Table S2**). Additionally, we extracted all mining areas categorized as “quarry” in land use type from OpenStreetMap (OSM, collected by January 2024), resulting in 170,620 mining-related polygon features (**Supplementary Table S1, Supplementary Table S2**). OpenStreetMap, one of the most popular open data communities, has already contributed to filling existing data gaps, especially in socioeconomic domains^{43,44}. In OSM, a quarry is defined as an area of land used for surface extraction (open-pit mining) of mineral or organic resources like gravel, sand, coal, or other land²⁸. We combined these four mining datasets spatially to create a comprehensive global database for mining activities (**Supplementary Fig. S4, Supplementary Table S2**).”

Line 255: 1 or 0.1 hectare (Line 88)?

> The threshold is set at 0.1 hectare. The previous study indicate that mining areas are generally larger than 1,000 km² (Tang and Werner (2023)); therefore, we defined the threshold for mining areas at 1,000 m² to remove very small-scale mining areas that may have low accuracy (Page 8, Lines 298-299). Additionally, in response to Reviewer #1's suggestion, we have now converted all results to m² and km² throughout the manuscript.

Page 8, Lines 298-299 (Methods):

“Additionally, mining polygons smaller than 1000 m² were filtered out due to the potential low accuracy associated with very small mining areas²⁰.”

Reference:

[1] Tang, L., Werner, T.T. Global mining footprint mapped from high-resolution satellite imagery. *Commun Earth Environ* **4**, 134 (2023).

Line 266-277: Please discuss the uncertainties on the SNL database.

> Thank you for the good suggestion. We now added details about the uncertainties of SNL database in Discussion (Page 7, Line 225-231) and Methods (Page 9, Lines 323-331), which are duplicated as follows.

Page 7, Line 225-231 (Discussion):

“Furthermore, we defined “documented mining” based on spatial overlays with data from the S&P Global Market Intelligence database, which reports mining sites as point locations. However, this database still suffers from its inherent

biases and limitations due to reliance on public disclosures, potentially leading to uncertainties regarding undocumented informal mining areas^{31,37}. Incorporating more diverse datasets and integrated approaches will be critical for improving the identification of mining activities in different forms and quantifying their associated impacts.”

Page 9, Lines 323-331 (Methods):

“The SNL mining sites include both active and inactive mining operations, reflecting various stages from grassroots exploration to closure, and provide insight into current investments and industrial mining activities^{18,24,30}. However, mining activities located outside project boundaries are generally not covered by the database^{24,45}. Similarly, scan reporting mining and informal mining activities such as artisanal and small-scale mining areas typically missed in S&P database. Despite these limitations, the S&P database is one of the most comprehensive and up-to-date sources of mining data, and it is commonly referenced in previous studies documents mining sites, acting as a proxy for reported mining sites⁴⁵.”

Line 342: Some references are outdated (e.g., 7, 13, and 31).

> Thank you. We have removed the outdated references and replaced them with new references. Please see our revised manuscript and references.

Response letter

> We appreciate the editor and reviewers for their time spent in assessing this manuscript and for their thoughtful and valuable comments and suggestions, which have significantly enhanced the overall quality of our study.

> Please find our point-by-point response (in blue) to the comments below. To make it easy to track our revisions, all modified text is marked in blue in the revised manuscript.

Reviewer #3 (Remarks to the Author):

Dear authors,

Thank you for your paper titled “Overlooked deforestation from global mining activities in the 21st century”. After the first review round, the paper is written well and makes an excellent contribution to the mining and responsible sourcing of minerals debate, i.e. with regards to the consideration of small-scale mining.

Authors: We greatly appreciate your positive feedback on our revision and thank you for the valuable and insightful comments that have contributed to improving our manuscript.

My background is in mining and sustainability, hence a number of my comments below are based on this background, i.e. concerning your use of mining-related terms, which I think can be improved. I also used the opportunity of Springer to involve an early career researcher, one of my PhD students, in this review. The comments are also below.

Authors: Thank you for the great comments and suggestions from you and your team. We have carefully reviewed them and revised relevant sections of the manuscript to improve clarity and accuracy accordingly. Please see our point-to-point response below addressing your concerns.

In general, you incorporated the comments of the first-round reviewers well. I think, however that your paper needs another revision. See especially the comments concerning lines 94 and 96 below.

Authors: Thank you for your acknowledgment of our revision effort. We have carefully checked and revised to improve the overall quality of the manuscript according to your feedback.

Regarding your concern about the use of mining-related terms in lines 94-96, we clarify that our mining areas were derived from four datasets, which include mining polygons or land-use parcels with visibly identifiable mining features. The definition of a mining area in our study refers to the footprints of mining operations, including underground mining activities such as waste rock dumps, as well as various quarries (e.g., sand,

gravel), as reported in previous studies (Tang and Werner, 2023a; Werner et al., 2020). We have revised the relevant sections to clarify the definition of mining areas in the Results (**Page 1, Lines 94-97**), Methods (**Page 9, Lines 330-337**) and to discuss the uncertainty in the Discussion (**Page 7, Lines 238-242**). Please see our detailed response below.

Overall: there are still a number of grammatical and spelling mistakes in the text that need correction, e.g. Line 45: “A plenty of case studies have documented”. Delete “A”. Line 47: “ecosystems”, etc.

Authors: Thank you for pointing out this issue. I have double checked the language throughout the entire manuscript and revised relevant terminologies and corrected typos for grammar and spelling. The example you spot now reads as below:

Page 2, Lines 46-49 (Main):

“Plenty of case studies have reported the substantial impacts of mining activities in tropical rainforest, as seen in countries such as Brazil³, Colombia¹⁵, Congo¹⁶, and Indonesia¹⁷. In light of the negative impacts of mining activities on forest ecosystems (e.g. deforestation and degradation), local, regional, and global initiatives have invested considerable effort to address these issues⁷.”

Line 31: I suggest you replace “mining commodities” with the term “mineral products”.

Authors: Thank you for your suggestion. Revised.

Page 2, Line 31 (Results):

“Numerous industries globally rely on mining products¹.”

Line 32: “..., supports billions of people in mining-related activities, ...This is the first time, that I hear of billions of people IN mining. Should this be millions (see line 204)? Or do you consider indirect impacts from mining activities– if so, I suggest you replace “in” with “through”.

Authors: Thank you for pointing this out. We have clarified the sentence in **Page 2, Lines 31-33**, which now reads:

Page 2, Lines 31-33 (Main):

“The mining sector contributes 2-4% to global Gross Domestic Product (GDP), supports millions of people directly through mining-related activities, and accounts for approximately 4-7% of global greenhouse gas emissions¹.”

Line 94/95: “...therefore, mining activities in shafts are not included here”. There is no such thing as mining activities in shafts! Shafts (as well as portals and ramps) are only

ways on how to enter an underground mine. So this raises a question: Do you mean that underground mines are not included? Or just their access points? This is important, because whilst underground mines have generally smaller land footprints (also depending on the mining method, i.e. use of backfill (little to no impact on the surface) vs. caving (large subsidence areas with high impact, see e.g. Kiruna in Sweden)) than surface mines, they still have a footprint, i.e. they also have processing and waste / tailings storage facilities, etc. This is also relevant for line 224 onwards and 303, and is not clear to me from the description there as well. Please clarify.

Authors: Thank you for your valuable insights. Our study focuses on the land surface footprint by mining activities, including both open pits and land footprints of underground mines. Below are a few examples illustrating the footprints of underground mines (such as processing infrastructure, tailings storage facilities) (**Supplementary Figure S1**). We have revised the definition of mining areas in our study for better clarifications in the Results (**Pages 3-4, Lines 94-98**) and Methods sections (**Page 9, Lines 327-337**).

Pages 3-4, Lines 94-98 (Results):

*“Noted, our mining data primarily relies on published footprints of mining activities obtained through crowdsourcing and manual visual inspection methods^{8,21,29,30}, including surface mining, placer mining, and land footprints directly related to underground mining operations (**Supplementary Figure S1**). ”*

Page 9, Lines 329-337 (Methods):

*“We combined these four mining datasets spatially to create a comprehensive global footprint for mining activities (**Supplementary Fig. S7, Supplementary Table S2**). By combining these mining datasets, our study captures a broad spectrum of mining activities, from large-scale industrial to smallholder mining operations, with varied delineation scales ranging from detailed individual infrastructures to broader land use categories. These included all surface mining footprints and underground mining features recognizable from satellite images or community knowledge, such as open pits, quarries, water storage ponds, milling infrastructures, waste rock dumps, and tailings storage facilities which are directly associated with mining operations (**Supplementary Table S1, Supplementary Fig. S1**).”*

[Figure Redacted]

Supplementary Figure S1, Examples of the mining activities footprints overlaid on Google earth satellites imagery (Google, 2023 Maxar Technologies).

Line 96: “We identified a total of 230,068 undocumented mines...”. I suggest you delete “undocumented” in this sentence- as you then explain that these mines are just not in the S&P database. After this explanation, using the term is fine for me. Otherwise, “undocumented” has this unwritten link to “informal” and/ or “illegal” right away, which is certainly not the case for Europe, where the high percentage of “undocumented mines” are simply small mines and quarries that are too small for S&P but certainly they are permitted and hence not illegal or informal. In line 184 you make exactly this link yourself and talk of “undocumented informal mining operations”. “Undocumented” in S&P and “informal” are not the same thing! And then, in line 191, you only use the word “informal” anymore. Did you generally analyze the “undocumented” mines in such a way that you are able to say that they are “informal”? In lines 199-204 making this link is ok for me, as there you do so specifically and supported by references.

Authors: Great point; we highly appreciate your feedback. Our study highlights the unreported mining activities in S&P (now SNL) mining database. Due to the lack of consistent standards and reference datasets across different countries and regions currently for identifying informal/illegal activities, this study is not about classifying the formal and informal/illegal mining activities for our mining areas.

Yes, a lot of unreported mining areas in Europe mainly were small quarries, sand and gravel. We have clarified this context of undocumented mines in main text (**Page 4, Lines 102-105**). We agree with your insights and have removed the term “informal” in the main text (**Page 4, Lines 98-100; Page 6, Lines 192-195**) and further clarifying this point in our revised Methods (**Page 10, Lines 369-372**).

Page 4, Lines 102-105 (Results):

“Europe show high rates of undocumented mining activities at 81.57% (Table 1),

*primarily contributed by small mines and quarries that are often unreported in the SNL mining database (**Supplementary Figure S2, Supplementary Tables S5**)."*

Page 4, Lines 98-100 (Results):

*"We defined mines not recorded in the SNL mining database as undocumented mines¹⁹, resulting in a total of 230,068 polygons, accounting for 93,596 km² (**Table 1**)."*

Page 6, Lines 192-195 (Discussion):

"This study provides a more comprehensive inventory for mining areas than those used in previous studies^{2,32}, encompassing 236,028 mining locations worldwide, including a substantial number of previously undocumented mining operations sites not captured in SNL mines database."

Page 10, Lines 369-372 (Methods):

"The undocumented mines highlight the potential underestimation of the impact of these mining activities by the suage of SNL mining database in previous studies. These mines may be subjected to permitted local regulations but have not been captured in the SNL database, including many small mines and quarries."

Line 143 onwards: Do all four of the datasets used for your analysis provide you with the same quality of commodity data or how did you align them? E.g. S&P gives you main products and by-products. Gold can be a by-product of copper mining – considered for copper only? Did you then use allocation and if so which method? How is that in the other databases? Lines 335 onwards in the methods description does currently not explain these aspects – this should be added.

Authors: Good question. In our analysis, we incorporated mining areas from all multiples mining footprints and land use datasets. These were merged spatially to create a comprehensive mining database, which we then independently validated to assess the overall accuracy and ensure consistency. However, it is worth noting that the S&P mining database (now SNL) only provides information on the primary products for each mine under the attribute "primary commodity". For consistency across datasets, we focused exclusively on the main product for the mining sites and did not consider by-products (e.g., gold as a by-product of copper mining). We acknowledge that this approach may overlook certain secondary commodities, but it ensures comparability across datasets. We have clarified this point and added further details in the Methods section (**Page 9, Line 372-376**).

Page 9, Line 372-376 (Methods):

*"We further identified the primary products of the mining polygons by analyzing the spatial nearest-neighbor mining activities, as reported in the SNL mining database, based on the Euclidean distance (**Supplementary Fig. S10**). Noted, the by-products of mining activities were not considered in this study due to the limitations in the attribute information provided by SNL mining database."*

Comments from the early-stage researcher:

Lines 96-102 & 113-124 & 300-301: It is recommended that the results section be reviewed, since there are inconsistencies between the values reported in the text and the data presented in the corresponding tables.

Authors: Thank you for spotting this. The updates for Table 1 were inadvertently missed during the revision process. We have now thoroughly checked and revised Table 1 in the results section. Please refer to the updated **Table 1** below.

Table 1. *The mining areas and mining-related deforestation and forest carbon emission by documented mines and undocumented mines across different regions for 2001-2023.*

Continent	Number of mining polygons		Mining area (km ²)		Mining-related deforestation (km ²)		Mining-related forest carbon emission (Pg CO ₂)	
	Undocumented	Documented	Undocumented	Documented	Undocumented	Documented	Undocumented	Documented
Africa	22,695	726	7639	6370	770	830	0.0202	0.0236
Asia	75,043	1,930	38110	18933	5174	2733	0.2012	0.1034
Australia	7,124	661	4608	5252	562	436	0.0063	0.0038
Europe	70,207	574	12264	2770	1001	237	0.0188	0.0040
North America	38,112	1,507	13684	9985	1933	1354	0.0420	0.0320
South America	16,887	562	17292	8995	3709	1027	0.2307	0.0600
Global	230,068	5,960	93596	52304	13149	6617	0.5192	0.2268

Lines 306-307 & 362-364: It would be useful to include a justification on the sample size used for the accuracy assessment, both in the selection of the 11,878 mining polygons and in the 2,200 deforestation pixels analyzed. On what criteria is the determination of these values based and why was a larger sample size not chosen?

Authors: Thanks for your suggestion and question. Both the validation of mining polygons and deforestation pixels require a manual visual check on a one-by-one basis, which is inherently time-consuming. The visual assessment for the mining polygons took over more than one week (approximately 50 hours, averaging 0.5 minutes per sample), while the evaluation of the 2,200 deforestation pixels required about three weeks for one person (approximately 110 hours, averaging 3 minutes per sample). This time was necessary due to the need to access and pre-process annual Landsat images for 2000-2023 via Google Earth Engine for each validation. Therefore, a larger

sample size was not chosen due to practical constraints, including time limitations, and the resources required for visual validation.

We are confident that the selected sample sizes provide a solid foundation for our accurate assessment. We have selected 11,878 mining polygons using a stratified random sampling method grouped by mining sizes, totally representing about 5% of the total mining polygons reported in our results. A sample size of 5% is generally considered sufficient for assessing overall accuracy for land cover change (Olofsson et al., 2014). For the 2,200 deforestation pixels, we chose a sample size significantly larger than the original assessment from the Hansen Global Forest Change, which includes 1,500 pixels for the global map (Hansen et al., 2013). It is also important to note that mining areas cover only about 5% of global forest cover. Therefore, we believe this sample size adequately reflects the overall accuracy of our mining-related deforestation. We have further incorporated this explanation into the revised Methods for clarity (**Page 10, Lines 341-345; Page 11, Lines 400-406**).

Page 10, Lines 341-345 (Methods):

*“Additionally, we performed an assessment to estimate the accuracy for the mining polygons using stratified random sampling⁵⁰. We randomly sample ~5% of the total mining polygons across different mining size groups, resulting in randomly selected 11,878 mining polygons to assess the overall accuracy of the mining datasets, which also spanned different geographical regions (**Supplementary Fig. S8**).”*

Page 11, Lines 400-406 (Methods):

*“To assess the accuracy of GFC forest loss in mining areas, we perform an independent assessment of deforestation in mining areas using stratified random sample methods (**Supplementary Data**). We randomly sample 2000 deforested pixels and 200 non-deforested pixels in our mining areas across six continents (**Supplementary Fig. S11**), which represents a much larger sample population than the original assessment of global product (v.1.0, $n = 1500$)¹². As the mining areas are about 5% of the global forest cover, the 2200 samples adequately reflected the overall accuracy of mining-related deforestation.”*

References:

1. Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* 148, 42–57. <https://doi.org/10.1016/j.rse.2014.02.015>
2. Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342, 850–853. <https://doi.org/10.1126/science.1244693>

Line 320/365: It is recommended to correct the typographical error on line 320, where “23,6028” appears to be misspelled. Also, in line 365, it would be helpful to clarify the use of the term 'Medoid'. Does it refer to a median-based composite?

Authors: Thanks for spotting this typo, and we have revised them in our manuscript in Page 10, Lines 356-357.

Page 10, Lines 356-357 (Methods):

“We categorized the 236,028 mining areas as documented and undocumented mines based on integration with SNL metals and mining database.”

Additionally. For the concern about the medoid value, which is different with median-based composite. The medoid approach looks at all unmasked values for each pixel, calculates the median value mathematically, but then identifies the actual pixel observation (i.e., the specific image from the collection that year) that is closest to the median in *n*-dimensional spectral space. Thus, the method tends toward the representative spectral value but uses the observations from an actual image (rather a contrived value calculated as the median of each band). Standard use of medoid composites is intended to minimize the influence of vegetation phenology, effectively controlling for seasonal variability prior to modeling temporal trajectories (Pasquarella et al., 2022). We have further incorporated this explanation and reference into the revised manuscript for clarity in support the usage of medoid value (**Page 11, Lines 408-412**):

Page 11, Lines 408-412 (Methods):

*“We then aggregated the annual medoid values (the actual data in a dataset that is most centrally located)⁵⁵ for Landsat time series (level-2 product with atmospheric correction, masked for cloud cover) and plotted the time series for the Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) from 2000-2023 using Google Earth Engines (**Supplementary Fig. S12**).”*

Reference:

1. Pasquarella, V.J., Arévalo, P., Bratley, K.H., Bullock, E.L., Gorelick, N., Yang, Z., Kennedy, R.E., 2022. Demystifying LandTrendr and CCDC temporal segmentation. *Int. J. Appl. Earth Obs. Geoinformation* 110, 102806. <https://doi.org/10.1016/j.jag.2022.102806>

Lines 364-378: It would be useful to clarify the methodology used to define deforestation: was a specific NDVI/NBR/EVI threshold applied, and was its impact on the classification evaluated to avoid errors? In addition, it is recommended to indicate if the Landsat images were atmospherically corrected and if topographic effects in mountainous regions were considered.

Authors: Thank you for your comments and suggestion. We visually examined whether deforestation occurred for each sample pixels, rather than relying on threshold method. Deforestation refers to the conversion of forest cover to other land use (FAO, 2020). We plotted the annual time-series trends of NBR, NDVI, and EVI for each sample pixel to support our validation of these samples (see Supplementary Fig. S8). These trends provided contextual evidence to assist analysts in distinguishing instances of deforestation. We believe that visual validation offers an independent reference of accurate assessment. We have revised the relevant text for a better clarity in **Page 11, Lines 412-416**.

Page 11, Lines 412-416 (Methods):

*“We visually interpreted these pixels to validate whether deforestation occurred and when deforestation took place using Landsat imagery with true color visualization (RGB). The annual time series of NBR, NDVI and EVI for these samples assisted was plotted to support the visual validation of the sample pixels (**Supplementary Fig. S12**), which are widely used indicators for deforestation monitoring⁵⁶.”*

Additionally, for the concerns about Landsat images, we utilized Landsat imagery from the Landsat Level 2 product available on Google Earth Engine (<https://developers.google.com/earth-engine/guides/landsat>), which has been atmospherically corrected by the USGS. As clarified in our previous response, we visually validated the occurrence of deforestation for the sample pixels rather than relying on specific thresholds. In this process, the Landsat true color images, along with the NDVI, EVI, and NBR indices, served as references for visual validation. The impact of terrain effects on our assessment was minimized by also considering the surrounding pixels (Supplementary Fig. S8). Topographic effects did not significantly influence our visual validation of deforestation, as we did not use threshold-based methods to determine deforestation for the validated samples. We have further clarified this in the Methods section to enhance transparency (**Page 11, Lines 408-416**).

Page 11, Lines 408-416 (Methods):

*“We then aggregated the annual medoid values (the actual data in a dataset that is most centrally located)⁵⁵ for Landsat time series (level-2 product with atmospheric correction, masked for cloud cover) and plotted the time series for the Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI) from 2000-2023 using Google Earth Engines (**Supplementary Fig. S12**). We visually interpreted these pixels to validate whether deforestation occurred and when deforestation took place using Landsat imagery with true color visualization (RGB). The annual time series of NBR, NDVI and EVI for these samples assisted was plotted to support the visual validation of the sample pixels (**Supplementary Fig. S12**), which are widely used indicators for deforestation monitoring⁵⁶.”*

Best regards,

Michael Tost

Professor for Sustainable Mining Technology

Chair of Mining Engineering and Mineral Economics

Montanuniversität Leoben, Austria

Dear Micheal,

We sincerely appreciate you and your student for the time and effort you dedicated to improving our work. Your valuable feedback has significantly improved our manuscript and enhanced its clarification and transparency.

Thank you once again for your support for our study.

Sincerely,

All authors.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Authors: Thank you for your effort in a detailed assessment and your constructive feedback.

Reviewer #5 (Remarks to the Author):

PREAMBLE

Firstly, I would like to apologise for the additional time it has taken me to render a review. I had been unexpectedly indisposed for health reasons for a few weeks, and I hope you will forgive the delay on the spaces.

All accounts indicate that I have reviewed a manuscript that has already been revised. All the same, and perhaps annoyingly for the authors, my review encompasses all elements of the manuscript, not just the revised portions. As the manuscript is new to me, I address all its facets without discrimination, as they are all new to me.

Authors: We greatly appreciate your effort in providing a detailed assessment, which has significantly enhanced both the clarity and rigor of our study. Please refer to our detailed responses and the revised manuscript.

MAIN COMMENT

The manuscript is, in effect, a simple or, better terms, straightforward overlay analysis. Two datasets are gathered up, and overlaid to see how they intersect. I refer here specifically to a 'new' composite data on mining, produced by combining existing datasets, and on the other hand's a well-worn global dataset on gross forest loss over the last 2 decades. There's nothing wrong with this approach, nor with the data brought to bear here. My main preoccupation when reviewing the study was to ensure that its messages were clear and findings were sufficiently rigorous and detailed.

Authors: Thank you for your thoughtful and constructive feedback. We acknowledge that our study presents a straightforward and effective way to assessing the overall impact of global mining activities on forests. Your review has helped us refine our presentation to ensure that the study's objectives, limitations, and contributions are clearly communicated. We have carefully addressed your concerns in our revised manuscript, providing additional methodological details and clarifying key findings.

I observed no fatal flaws in the data or analysis per se - nothing that would unquestionably prevent publication, at least. That said, there are many points that are unclear, sometimes seemingly in part due to English being perhaps a second language for some authors, I presume, and in part for other unknown reasons, speculatively and unfamiliarity with certain datasets or reporting standards. These points of lack of clarity, sometimes bordering on under-reporting. This will have to be addressed before anything is actually publishable.

So I am recommending 'major revision' to clarify and tighten up reporting and a few key analytical details of the work already done. I don't recommend any new analysis per se, notwithstanding one elaboration of an analysis already realised, in relation to protected areas.

In the following texts I raise points as I observe them in the manuscript, in order of

reading them, also organized by their headings and subheadings. My comments do not appear in priority order. I do however indicate in my text which issues are relatively important, by saying just so in my individual comments, below.

Authors: Thank you for all the good words and insightful comments. We have made a thorough revision to address your concerns and suggestions, particularly providing additional methodological details and clarifying key findings. Please find our detailed point-by-point responses below.

MAIN RESULTS

> A GLOBAL INVENTORY...

1. In the first few paragraphs of the results, I note that it is reported that's over 80% of mining areas in the authors composite mining data sets are supposedly undocumented. This means that the mines were not recorded in the S&P mining dataset the authors adopted to differentiate between documented and undocumented mines. It is impossible for me to believe that 81% of mining areas in Europe are unknown or unregulated by local governments. This, of course, is not what 'undocumented' means, in the sense that the authors use the term, though it still may be fairly the reader's initial impression. So, on this basis, I become sceptical of the ultimate worth of the S&P mining data set, even if it is, as the authors allude, the best available global data set on mining activity.

Authors: Thank you for your insightful observation regarding the classification of undocumented mining areas within our composite mining dataset. In our study, the term "undocumented" specifically refers to mining areas that are not included in the SNL mining dataset; this does not necessarily imply that these areas are unregulated or unknown to local governments. It is possible that these mines are subject to local regulations but have not been captured in the SNL dataset due to various reasons, such as inconsistent reporting practices or challenges in data acquisition.

For the high percentage of undocumented mining areas in Europe, as mentioned by Reviewer#3, most of mining areas was composed of small mines and quarries, which are generally unreported in the SNL mining database. We further analyzed the documented and undocumented areas across different mining sizes, as shown in **Supplementary Tables S5**. The results indicate that undocumented areas are primarily contributed by small-scale mining polygons. These small-scale mining activities mainly consist of sand and gravel operations (**Supplementary Figure S2**). These small-scale mining areas often were omitted by public attention, and even scientific communities as there are often slight impacts. Additionally, most of these small-scale mining areas in our mining polygons are sourced from OpenStreetMap (OSM). In Europe, OSM tends to provide higher accuracy and more detailed reporting, as contributed by volunteers, compared to developing regions such as Africa and Southeast Asia (Herfort et al. 2023). Therefore, the undocumented areas in Europe are associated with the large number and widespread distribution of quarries. We have revised this point in the Results provide further clarification (**Page 4, Lines 102-105**)

Page 4, Lines 102-105 (Results):

“Europe show high rates of undocumented mining activities at 81.57% (Table 1), primarily contributed by small mines and quarries that are often unreported in the SNL mining database (Supplementary Figure S2, Supplementary Tables S5).”

[Figure Redacted]

Supplementary Figure S2. *a*, Mining areas and SNL mining sites in Europe. *b*, Small-scale mining areas in east Europe. *c*, Examples for some small mining in Spain (c1), Germany(c2), Austria (c3), Italy (c4).

Supplementary Tables S5. Undocumented areas and documented mining areas in Europe.

<i>Mining size (Area: km²)</i>	<i>Undocumented mining areas</i>	<i>Number of Undocumented mining polygons</i>	<i>Documented mining areas</i>	<i>Number of documented mining polygons</i>	<i>The ratio of documented mining areas</i>
<i><0.01</i>	<i>83.75</i>	<i>18660</i>	<i>0.04</i>	<i>6</i>	<i>99.96%</i>
<i>(0.01, 0.05]</i>	<i>586.22</i>	<i>22665</i>	<i>0.65</i>	<i>22</i>	<i>99.89%</i>
<i>(0.05, 0.1]</i>	<i>722.25</i>	<i>10095</i>	<i>1.67</i>	<i>23</i>	<i>99.77%</i>
<i>(0.1, 0.5]</i>	<i>3281.04</i>	<i>14905</i>	<i>46.54</i>	<i>161</i>	<i>98.60%</i>
<i>(0.5, 1]</i>	<i>1535.40</i>	<i>2242</i>	<i>67.45</i>	<i>92</i>	<i>95.79%</i>

(1, 5]	2674.23	1411	367.86	171	87.91%
(5, 10]	924.95	133	255.86	35	78.33%
(10, 50]	1713.93	91	1354.79	56	55.85%
>50	741.75	5	675.52	8	52.34%
Sum	12263.53	70207	2770.38	574	81.57%

Additionally, we also have removed the term “informal” throughout our manuscript and focused on the unreported mining activities in SNL mining dataset (**Page 4, Lines 98-99; Page 5, Page 5, Lines 159-161; Page 6, Lines 192-198**). Additionally, we have further clarified the definition of “undocumented mines” in accordance with (Maus and Werner, 2024) and clarify the undocumented may be subjected to permitted local regulations but have not been captured in the SNL database, such as small mines and quarries (**Page 6, Lines 201-203; Page 10, Lines 369-372**). Please see our detailed revised text below:

Page 4, Lines 98-99 (Results): “We defined mines not recorded in the SNL mining database as undocumented mines¹⁹, resulting in a total of 230,068 polygons, accounting for 93,596 km² (**Table 1**).”

Page 5, Lines 159-161 (Results): “Our results also reveal that 73.21% of gold mining-related deforestation was due to undocumented mining activities during 2001-2023, which facilitated the persistence of unrecorded mining activities on SNL databases (**Fig. 3c**)”

Page 6, Lines 192-198 (Discussion): “This study provides a more comprehensive inventory for mining areas than those used in previous studies^{2,32}, encompassing 236,028 mining locations worldwide, including a substantial number of previously undocumented mining operations sites not captured in SNL mines database. Building upon this enhanced inventory, we provide an up-to-date and spatially explicit assessment of these mining activities’ direct impact on forest changes, revealing that mining has led to a total loss of 19,765 km² of deforestation, contributing to 0.75 Pg CO₂ emissions from 2001 to 2023.”

Page 6, Lines 201-203 (Discussion): “As a result, they have significantly underestimated the overall impacts of global mining activities, including undocumented operations in previous studies using SNL mining database.”

Page 10, Lines 369-372 (Methods): “The undocumented mines highlight the potential underestimation of the impact of these mining activities by the suage of SNL mining database in previous studies. These mines may be subjected to permitted local regulations but have not been captured in the SNL database, including many small mines and quarries.”

References:

1. Maus, V., & Werner, T. T. (2024). Impacts for half of the world's mining areas are undocumented. *Nature*, 625(7993), 26-29.
2. Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2023). A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. *Nature Communications*, 14(1), 3985.

2. The authors reports that 139 protected areas, plus an additional simply sent from the wilderness areas, are intersected or otherwise overlaid by mining areas as they observe them in their composite dataset on mining areas. The authors report that these protected areas are of “the highest category and protected area”, but do not give any criteria by which category of protected area is defined, or even what it means. Protected areas are commonly categorised by six or so categories the IUCN (IUCN protected area categories - Wikipedia), but these categories are often not listed for many protected areas in the world database unprotected areas the author is used. All the a fraction of all protected areas in Africa, for example, have attribute data on the IUCN management category, so it is doubtful to me that this categorical attribute was used to describe protected areas as reported, and I am left with no other guests as to how protected areas were described in terms of their management category. The authors seem equally unclear, and the reporting on their selection and categorization of protected areas in the methods section effectively said nothing on the topic or otherwise. This must be corrected.

Authors: Thank you for your constructive comment. The categorization of protected areas in this study follows the IUCN protected areas categories. 139 strict nature areas and 77 wilderness areas protected areas are both according to the IUCN categories.

In the revised manuscript, we have clarified in the main text (**Page 4, Lines 106-118; Pages 12-13, Lines 451-464**) that our Methods section now provides a detailed description of the categorization of protected areas (**Supplementary Table. S11**). We acknowledge that many protected areas were not listed/assigned/reported in WDPA databases due to their inherent limitations.

Page 4, Lines 106-118 (Results):

“Our results show that a total of 11,425 global protected areas have been or are currently affected by mining activities, based on the intersection with World Database on Protected Areas (WDPA)³¹ (Fig. 1d, Supplementary Fig. S3). Specifically, 139 Ia-level protected areas (strict nature reserves) and 77 Ib-level protected areas (wilderness areas) have been observed with mining activities, which was strictly prohibited mining activities and infrastructural access, according to the International Union for Conservation of Nature (IUCN) protected areas categories (Methods). Additionally, 528 II-level protected areas (national parks) and 371 III-level protected areas (natural monuments are features) overlap with mining areas; these are generally also forbidden with mining extractive activities. Furthermore, 2,163 protected habitat or species management areas overlap with mining areas, and 2,184 protected nature

landscape and seascape areas have observed mining activities. Noted, 5550 other protected areas have also observed mining activities; however, these areas are not reported or assigned management categories on WDPa, demonstrating the potential extensive impacts of mining operations on protected areas worldwide.”

Pages 12-13, Lines 451-464 (Methods):

“Protected areas are some of the most critical areas for the conservation of natural landscapes, wildlife, and biodiversity and hence mining activities and deforestation in these areas pose a threat to the global protected areas. To explore the potential impacts of mining activities on protected areas, we calculated the number of protected areas overlay in mining areas based on World Database Protected Areas (WDPa)³¹. WDPa protected areas are carried out by the Union Nation in collaboration with governments, non-governmental organizations, academia and industry. This provides the most comprehensive database for global protected areas. According to the International Union for Conservation of Nature (IUCN), protected areas are classified into six distinct management categories: Ia: strict nature reserve; Ib: wilderness area; II: national park; III: natural monument or feature; IV: habitat/species management area; V: protected landscape/seascape; VI: protected area with sustainable use of natural resources (details in **Supplementary Table. S11**). These categories are not listed for some protected areas due to omissions in the WDPa dataset, but they does not reveal any information whether management objectives are in place or if the site is of importance^{31,57}.”

Supplementary Table. S11. Protected areas category and description from World Database Protected Areas (WDPa)(IUCN & UNEP-WCMC, 2017).

Protected areas category	Description	The limitation of human activities
Ia. strict nature reserve	Areas set aside to protect biodiversity and possibly geological or geomorphological features, which are usually managed for scientific research.	Limited to scientific research and education; no extractive or recreational uses allowed.
Ib. wilderness area	Areas are a protected domain in which biodiversity and ecosystem processes (including evolution) are allowed to flourish or experience restoration if previously disturbed by human activity.	Allows non-motorized recreational activities (e.g., hiking), but no infrastructure or motorized access.
II. national park	Large natural areas set aside to protect natural biodiversity and provide for environmentally compatible recreation.	Allows for recreational activities, education, and tourism but restricts damaging activities (e.g., mining and logging).

III. natural monument or Feature	Areas that are set aside to protect a specific natural monument, which can include landforms, sea mounts, or a combination of features that are significant for their environmental or cultural value.	May permit limited recreational activities, but extraction activities are generally prohibited.
IV. habitat/species management area	Areas managed to protect particular species or habitats, often involving active management intervention.	May allow for sustainable level extraction or land use but focuses on conservation of habitats and species.
V. protected landscape/seascape	Areas that are protected to maintain the interaction between people and nature, where the landscape is shaped by human activity but still retains ecological significance.	Encourages sustainable land use and conservation of cultural landscapes, allowing for agricultural and recreational activities.
VI: protected area with sustainable use of natural resources	Areas that aim to protect ecosystems while allowing for sustainable use of natural resources, balancing conservation and resource extraction.	Allows for sustainable extraction and use of resources (e.g., forestry, fishing) that do not compromise conservation goals.
Other	- No application: The protected area does not fit the IUCN's definition or management categories - Not Assigned: The IUCN category has not yet been assigned to the protected area. - Not reported: The IUCN category exists but has not been reported to the WDPA.	Noted: The absence of a management category does not in any way reduce the importance of a protected area, nor does it imply that the site is not being adequately managed or should be excluded from analyses.

3. A related point of interest to me and my field, regarding protected areas impacted by mining, is that there does already exist a global database on this phenomenon - PADDTracker.org. This online database is compiled on the basis of any and all reports and datasets on phenomenon like mining or logging that effect protected areas, leading to their degradation, downgrading in status, or de-gazettment. Reports and listings in the online data set are vetted by an expert panel, not just random notes by citizen scientists. The author is really should, even must, compare their list of protected areas impact by mining against that compiled by this online database of the same.

Authors: Thank you for bringing the PADDTracker.org database to our attention. We recognize the value of this resource. In our current study, we relied on publicly available protected area boundary datasets and registered geographic information, but we did not incorporate data from PADDTracker.org. We agree that comparing our findings with this existing database could enhance our analysis and provide an additional validation layer.

We have access to a total of 3,500 protected areas from the PADDTracker Data Release Version 2.1 (Figure R1a), shared by the World Wildlife Fund via Zenodo, as shown below (Figure R1, Figure R2). The PADD covers some countries while omitted protected areas for most countries compared to WDPA database (Figure R1b). The PADD dataset only reported 130 protected areas have been impacted by mining activities (Figure R1c). We conducted a detailed review for these 130 PADD-protected areas affected by mining activities (Figure R2b) and suggested the limitation of PADD at large-scale assessment of mining impacts.

In some PADD-protected areas, no evidence of mining activity was found—for example, in certain regions of Cambodia (Figure R2b)—which we further verified using Google Earth imagery. This likely leads to an underestimation of the global mining impact on protected areas, since PADD primarily focus some protected areas across the United States and Amazonia. Moreover, although the PADD records are validated through peer review, many protected areas that have been caused or affected by mining activity may still be missing from the database.

[Figure Redacted]

Figure R1. **a**, Protected areas from PADDTracker Data Release Version 2.1. **b**, Protected areas from WDPA database. **c**, PADD protected areas distributed by mining. **d**, WDPA protected areas overlaid by mining areas.

[Figure Redacted]

Figure R2. a, Protected areas and mining areas in our study in USA. **b,** Protected areas and mining areas in Cambodia.

Given the limited coverage of PADD dataset, we believe the World Database on Protected Areas (WDPA) provides more comprehensive global maps and information on protected areas. Therefore, our study is primarily based on WDPA data. Additionally, we acknowledge and discuss the limitations of the WDPA dataset. Accordingly, we have revised our manuscript to better articulate our results, clarify the usage of WDPA (**Pages 12-13, Lines 451–464**), and highlight the limitations of the WDPA, particularly regarding the management category data (**Page 4, Lines 106-118**), as shown in previous response.

References:

1. Conservation International, & World Wildlife Fund. (2021). PADDTracker Data Release Version 2.1 (2.1) [Data set]. Zenodo.<https://doi.org/10.5281/zenodo.4974336>.
2. Golden Kroner, R. E., Qin, S., Cook, C. N., Krithivasan, R., Pack, S. M., Bonilla, O. D., ... & Mascia, M. B. (2019). The uncertain future of protected lands and waters. *Science*,364(6443), 881-886.
3. IUCN & UNEP-WCMC. The world database on protected areas (WDPA). (2017).

>DEFOR FROM MINING...

4. Perhaps the biggest source of uncertainty for me is that what qualifies as a so-called “mining area” is not so transparent. Seemingly, no single standard or criterion to describe a spatial area as dedicated to mining, rather than simply being in proximity of mines, has been applied. This of course reflects the fact that the authors’ composite data on mining areas is derived from multiple independent datasets, so at the very least consistency in this respect cannot be expected. But even the authors’ description

of these data in the relevant section of the methods text does not detail well enough for my liking the criteria by which a given area surrounding mining activity but itself not dedicated to mining activity may be included as a “mining area”. In effect, the criteria and/or distance threshold that defines lands around mining activity as part of the “mining area” of observation here remains an open question. I am, of course, reasonably presuming that the entirety of a nominal “mining area” is not, in fact, dedicated to mining (hole digging) itself.

Authors: Thank you for highlighting this important issue. We recognize that the lack of a single standard or criterion for classifying areas as dedicated to mining versus those in proximity can lead to ambiguity. Our composite dataset is derived from multiple independent sources, which may result in variations in how mining areas are defined.

To clarify, we did not apply specific distance thresholds to delineate mining areas. Instead, the mining footprints shown represent land footprints directly associated with mining activities, including open pits, water storage ponds, milling infrastructure, waste rock dumps, and tailings storage facilities, as shown in Supplementary Figure S1. These footprints encompass both surface mining and footprints related to underground mining areas, although they do not capture the extent of hole digging.

To address this, we have provided a detailed explanation of the criteria used to define mining areas (**Pages 3-4, Lines 93-97; Page 9, Lines 333-337**) and added a figure illustrating examples of mining activity footprints (**Supplementary Figure S1**). This should improve clarity and understanding of our definitions.

Pages 3-4, Lines 93-97 (Results):

*“Noted, our mining data primarily relies on published footprints of mining activities obtained through crowdsourcing and manual visual inspection methods^{8,21,29,30}, including surface mining, placer mining, and land footprints directly related to underground mining operations (**Supplementary Figure S1**).”*

Page 9, Lines 333-337 (Methods):

*“These included all surface mining footprints and underground mining features recognizable from satellite images or community knowledge, such as open pits, quarries, water storage ponds, milling infrastructures, waste rock dumps, and tailings storage facilities which are directly associated with mining operations (**Supplementary Table S1, Supplementary Fig. S1**).”*

a. The reason the spatial definition of a mining area (i.e., mining polygon) matters for this analysis is that the authors quantified deforestation within the spatial bounds of a so-called mining area (polygons). Hence, if the bounds of these mining areas are liberally defined, Relative to the actual extent of explicit mining (hole digging), then the estimates of deforestation attributed to mining can likewise be liberal. In many areas of mining I know in central Africa, mines attract new settlers, and so settlement expansion, then also informal / artisanal mining around the corporate mines, new agriculture and related forest conversion, and so on - basically the whole gamut of

human-driven forest loss. Not all of this forest loss is related to the mining itself, except perhaps indirectly. So, the more liberal the normal mining areas of this study were defined spatially, the more of such secondary deforestation may be captured and attributed to the mining itself, again somewhat liberally. About 1/3 and 1/9 of the mining areas defined by the authors' composite dataset are greater than 1 km² and 5 km², respectively, as shown in Table S2, while a select few can be as large as 50 km², suggesting that liberal definitions of mining areas may apply here, even if to a minority of the mines observed. To assuage all preoccupation on the part of your readers, I would ask you to comment and clarify, for each individual input dataset, the criteria, means, scales etc. by which the nominal mining areas of your study were delineated spatially around the actual mining activity of interest (hole digging).

Authors: Thank you for your detailed and insightful comment, which can help strengthen the transparency and interpretability of our findings. In our study, the mining footprints are based on composite datasets that primarily represent footprints directly associated with mining infrastructure and operations, as shown in previous response. We acknowledge that these footprints may be broad and may encompass nearby lands, especially since some mining areas delineated through the manual visual interpretation, rather than pixel-based classification using satellite image. However, currently, no high-quality, wall-to-wall products capable of global mining activities with high precisions.

To address this issue, we conducted a detailed inventory of some large-scale mining polygons using Google Earth imagery (Figure R3). Additionally, we incorporated the ESA World Cover land cover map from 2021 due to its superior spatial resolution compared to other global land cover datasets (Xu et al., 2024). We calculated the total area of built-up land within large mining areas (as defined by those greater than 50 km²). Our results show that only approximately 346 km², accounting for 0.94% of the total mining areas (36,813 km²), is classified as built-up. Some of these areas include mining infrastructures (Figure R3c1). We acknowledge that their contribution of new settlements is marginal relative to the overall impact considered in our analysis.

[Figure Redacted]

Figure R3. Examples of large-scale mining areas in Ghana. **a.** mining polygons overlaid with Google earth satellite imagery. **b,** zoom-in view of mining areas with mining infrastructures (b1) and small settlements (b2), **c,** ESA Worldcover 10m land cover map for the zoom-in view.

Therefore, based on our best knowledge, we believe that the surrounding areas in mining (the footprint quantified in this study) had a slight impact on our estimation of direct deforestation from 2001 to 2021. Nonetheless, we advocate that our academic community and related stakeholders will promote the development of global, pixel-based mapping of mining activities in the future. We have added new text in discussion section to highlight this limitation and promote further accurately mapping the extent of various mining operations (**Page 7, Lines 238-247**), and we clarify the criteria, methods, and scales used for each individual dataset in the Methods section to ensure transparency (**Page 8-9, Lines 298-337**).

Page 7, Lines 238-247 (Discussion):

“Our study is subject to several levels of uncertainty. First, although the mining datasets compiled in this study represent the most comprehensive global inventory of mining activities to date—integrating multiple sources such as manual delineated mining footprints and crowdsourcing mining polygons from OpenStreetMap (OSM) map—they predominately focus on polygons or land-use parcels broadly defined surrounding visibly identifiable mining areas. As a result, these datasets are limited in their ability to precisely capture the actual extent of mining operations. To improve individual-level mining recognition and delineation, future efforts should prioritize multi-scale, wall-to-wall detection and segmentation of mining areas using high-resolution satellite imagery. This would enable more accurate delineation of the mining extents and their associated

land cover changes.”

Page 8-9, Lines 298-337 (Methods):

*“We combined four mining databases from multiple sources, including global mining polygons in 2019 and Global mining footprint in 2022²¹, artisanal mining areas in 2023⁸ and OpenStreetMap-based mining land use in 2024 (details in **Supplementary Tables S1,2**). The specific criteria and definition for each dataset are as following:*

(1) The global mining polygons²⁵ delineated all underground and above-ground mining infrastructure visible in satellite images. This dataset, derived from the visual interpretation of Sentinel-2 images for 2019, contained 44,929 polygon features covering 101,583 km². The polygons represent mining-related footprints that were delineated spatially around the actual mining activities, focusing on mining infrastructures and signatures indicative of active mining operations.

(2) Global Mining Footprint in 2022 provided by Tang et al.,²¹ also was collected here, as another source for mining areas. The dataset was mapped from remote sensing analysis of high-resolution imagery, comprising 74,548 mining polygons, covering 65,585 km². The boundaries of the mining footprints were manually delineated based on visual interpretation of the imagery, utilizing spectral, textural, and contextual cues to identify active extraction zones. The delineation focused primarily on features representing the operational core—namely, the areas of actual excavation and processing—rather than peripheral or ancillary land cover modifications.

(3) The artisanal mining areas⁸ were visually identified and outlined through a combination of literature and popular media review, mining company and national reporting on mining activity, social media investigations, and extensive manual satellite and aerial imager review. This dataset includes 660 mining polygons that are significant for artisanal mining activities in tropical regions, which primarily focus on river mineral mining, including selected global and demand mining⁸. The delineation of artisanal mining areas was based on visual interpretation and corroborating secondary sources, emphasizing the spatial boundary of active artisanal mining sites across tropical rivers.

(4) We also extracted all mining areas categorized as “quarry” in land use type from OpenStreetMap^{48,49} (collected by January 2024), resulting in 170,620 mining-related polygon features. The quarry is defined as an area of land used for surface extraction (open-pit mining) of mineral or organic resources like gravel, sand, coal, or other land²⁹. The OSM data was contributed by volunteers and reflects widely quarrying operations, defined by open pits or surface mining activities.

*We combined these four mining datasets spatially to create a comprehensive global footprint for mining activities (**Supplementary Fig. S7, Supplementary Table S2**). By combining these mining datasets, our study captures a broad spectrum of mining activities, from large-scale industrial to smallholder mining operations, with varied delineation scales ranging from detailed individual infrastructures to broader land use categories. These included all surface mining footprints and underground mining features recognizable from satellite images or community knowledge, such as open*

pits, quarries, water storage ponds, milling infrastructures, waste rock dumps, and tailings storage facilities which are directly associated with mining operations (Supplementary Table S1, Supplementary Fig. S1).”

Reference:

1. Xu, P., Tsendbazar, N. E., Herold, M., de Bruin, S., Koopmans, M., Birch, T., ... & Zanaga, D. (2024). Comparative validation of recent 10 m-resolution global land cover maps. *Remote Sensing of Environment*, 311, 114316.

>GOLD AND COAL

5. As a suggestion for future research, I would urge the authors to explore whether the increasing then decreasing trends they observe for some mining related deforestation track the global prices for the mined commodities in question.

Authors: Thank you for your insightful comments. We have indeed conducted some preliminary analysis on the correlation between mineral prices—specifically for gold and coal—and deforestation, as illustrated in Figure R4.

We noted an increase trend where deforestation correlated with mined commodity prices before 2010; however, this relationship has not held consistently since then. For instance, while gold prices have declined since 2012, deforestation began to decrease only in 2018 (Figure R4). This lag suggests that the impact of commodity prices on deforestation may not be immediate. The reduction in deforestation attributable to gold mining activities can be primarily linked to increased efforts to curb illegal mining, particularly in Peru (Caballero Espejo et al., 2018).

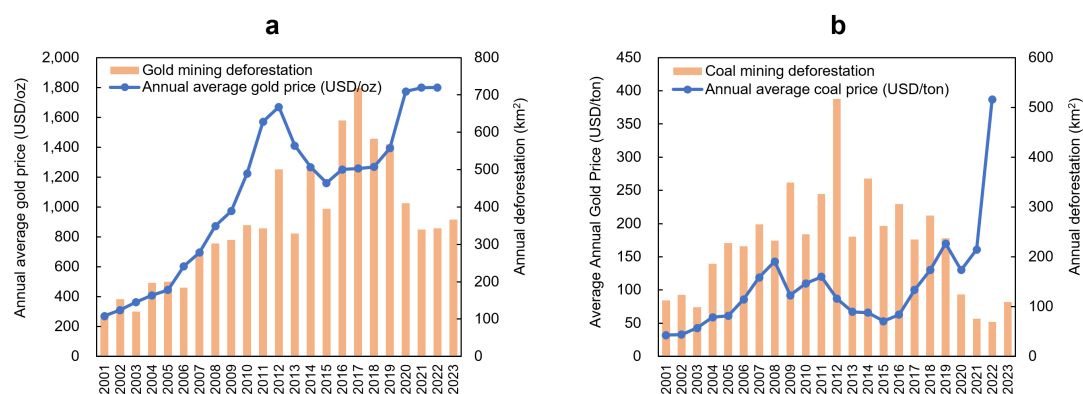


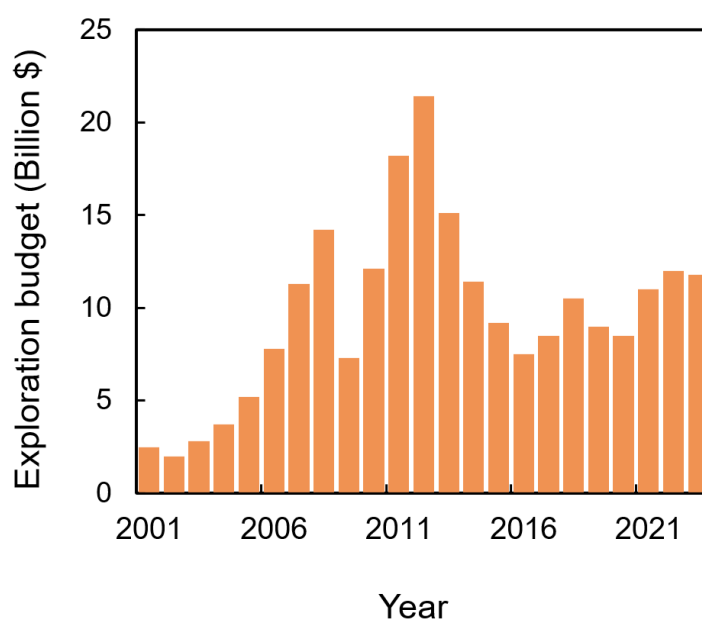
Figure R4. a. Annual average gold prices and deforestation by gold mining activities. **b.** Annual average coal prices and deforestation by coal mining activities.

Further exploration of the increasing and then decreasing trends between mined commodity prices and deforestation is warranted in future studies, especially considering the effects of local government policies and management practices aimed at reducing deforestation in the mining industry.

Additionally, we have further examined the changes in deforestation in relation to the global development of the mining industry, indicated by the annual budget for nonferrous exploration (**Supplementary Fig. S4.**) Our findings suggest that this temporal pattern aligns with the trends in global mining exploration budgets from 2001-2021. We have included a discussion of this intriguing insight to promote further exploration about the drive by mining-related deforestation in the discussion section (**Page 7, Lines 231-237**):

Page 7, Lines 231-237(Methods):

*“Moreover, the temporal changes in mining-related deforestation appear to be influenced by global exploration budgets of nonferrous minerals, potentially reflecting the demand dynamics driven by the international mineral market (**Supplementary Fig S5**). Previous studies have shown that fluctuations in the prices of special minerals can trigger the expansion of mining activities^{44,45}. Therefore, understanding the drivers of mining-related deforestation necessitates an integrated consideration of both global mineral market trends and the effects of regional and local mining management practices.”*



Supplementary Fig. S4. Annual nonferrous exploration budgets from 2001 to 2023. Source: S&P Global Martin Intelligence.

Reference:

1. Caballero Espejo, J., Messinger, M., Román-Dañobeytia, F., Ascorra, C., Fernandez, L. E., & Silman, M. (2018). Deforestation and Forest Degradation Due to Gold Mining in the Peruvian Amazon: A 34-Year Perspective. *Remote Sensing*, 10(12), 1903. <https://doi.org/10.3390/rs10121903>.

6. You write "Our results indicate that between 2001 and 2012, 66.20% of deforestation in mining areas was attributed to mines extracting minerals for renewable energy production (see Methods, Fig. 4b)." The language here is too presumptive that the minerals extracted will be used for renewable energy production. The language does not give due recognition to the fact that these minerals might, or might not, be used for renewable energy production. In reality, one would imagine that the bulk of most minerals are not used for this purpose, even when appropriate for that purpose. Please revise accordingly.

Authors: Thank you for your feedback. We agree our statement should be enhanced. We have rephrased this statement more accurately, as follows:

Page 4-5, Lines 178-180:

*"Our results indicate that between 2001 and 2012, 66.20% of deforestation in mining areas occurred in region where minerals were extracted, which could potentially be used for renewable energy production (see Methods, **Fig. 4b**)."*

METHODS

7. There is no such thing as the "Global Forest Watch dataset". The data you mean to refer to, notably uncited at the beginning of your methods text, are typically referred to as global forest change data from the university of maryland, or similarly global forest loss data from the university of Maryland. Global forest watch is merely a website that disseminates these data.

Authors: We have corrected this terminology to Global Forest Cover (GFC) throughout our manuscript. Please see our revised manuscripts for the updates.

>A COMPREHENSIVE DATASET...

8. The large paragraph at the outset of this subsection is written somewhat awkwardly, and not very clearly. Here, but generally everywhere to a lesser degree, a thorough re-checking by a native English speaker would be of benefit. Some of the revised sentences is blue do not make any sense at all here. One feature lacking clarity for me here was the listing of each of the four input datasets. I strongly suggest that you list them here listwise, as by numbered bullet points, rather than within the current large block of text.

Authors: Thank you for your suggestions. We have reformulated this paragraph and listed points in a numbered format for improved clarity (**Pages 8-9, Lines 298-328**), which now read as follows:

Pages 8-9, Lines 298-328 (Methods):

*"We combined four mining databases from multiple sources, including global mining polygons in 2019 and Global mining footprint in 2022²¹, artisanal mining areas in 2023⁸ and OpenStreetMap-based mining land use in 2024 (details in **Supplementary Table S1, Supplementary Table S2**). The specific criteria and definition for each dataset are*

as following:

(1) The global mining polygons²⁵ delineated all underground and above-ground mining infrastructure visible in satellite images. This dataset, derived from the visual interpretation of Sentinel-2 images for 2019, contained 44,929 polygon features covering 101,583 km². The polygons represent mining-related footprints that were delineated spatially around the actual mining activities, focusing on mining infrastructures and signatures indicative of active mining operations.

(2) Global Mining Footprint in 2022 provided by Tang et al.,²¹ also was collected here, as another source for mining areas. The dataset was mapped from remote sensing analysis of high-resolution imagery, comprising 74,548 mining polygons, covering 65,585 km². The boundaries of the mining footprints were manually delineated based on visual interpretation of the imagery, utilizing spectral, textural, and contextual cues to identify active extraction zones. The delineation focused primarily on features representing the operational core—namely, the areas of actual excavation and processing—rather than peripheral or ancillary land cover modifications.

(3) The artisanal mining areas⁸ were visually identified and outlined through a combination of literature and popular media review, mining company and national reporting on mining activity, social media investigations, and extensive manual satellite and aerial imager review. This dataset includes 660 mining polygons that are significant for artisanal mining activities in tropical regions, which primarily focus on river mineral mining, including selected global and demand mining⁸. The delineation of artisanal mining areas was based on visual interpretation and corroborating secondary sources, emphasizing the spatial boundary of active artisanal mining sites across tropical rivers.

(4) We also extracted all mining areas categorized as “quarry” in land use type from OpenStreetMap^{48,49} (collected by January 2024), resulting in 170,620 mining-related polygon features. The quarry is defined as an area of land used for surface extraction (open-pit mining) of mineral or organic resources like gravel, sand, coal, or other land²⁹. The OSM data was contributed by volunteers and reflects widely quarrying operations, defined by open pits or surface mining activities.”

9. In regards to your accuracy assessments of your composite dataset on mining areas, you mention a SPX data set. But this SPX data was not previously defined or introduced. Did you mean to write ‘SNL’ and/or ‘S&P’?

Authors: Yes, the correct dataset is the "SNL" dataset. We have corrected the typographical errors in the related text thoroughly.

Page 10, Lines 352-353 (Methods):

“Among the sampled mining polygons, 383 mining polygons were documented by SNL mining dataset (**Supplementary Fig. S8**).”

>DEFINITION OF UNDOCUMENTED...

10. The authors seemingly refer to the 'SNL' and 'S&P' dataset interchangeably, causing some confusion. Please check that you used a consistent term when referring to the data set in question.

Authors: Thank you for bringing this to our attention. The SNL mining dataset is provided by the S&P Global Intelligence Platform. We have revised the manuscript to consistently use the term "SNL" in accordance with previous studies (Maus et al., 2022; Tang and Werner, 2023b). Please refer to our revised manuscript for further details.

References:

1. Maus, V., & Werner, T. T. (2024). Impacts for half of the world's mining areas are undocumented. *Nature*, 625(7993), 26-29.
2. Tang, L., & Werner, T. T. (2023). Global mining footprint mapped from high-resolution satellite imagery. *Communications Earth & Environment*, 4(1), 134.

11. This statement is unclear: "However, mining activities located outside project boundaries are generally not covered by the database^{24,45}" What are project boundaries in the data, and how do they matter to the present analysis?

Authors: Thank you for pointing out the unclear statement. We have revised the related text to enhance clarity. The revised statement on Page 10, Lines 362-364 now reads:

Page 10, Lines 362-364 (Methods)

"However, scan reporting mining and informal mining activities such as artisanal and small-scale mining areas typically unrecorded in SNL mining database^{25,49}. Similarly, the historic and abandoned mining sited areas typically missed⁴⁹."

12. This sentence is unclear: "Similarly, scan reporting mining and informal mining activities such as artisanal and small-scale mining areas typically missed in S&P database. ". So is this one: "The mining areas interacting with SNL mining sites"... Are you referring to spatial intersection by the term 'interacting with'.

Authors: Thank you. We have revised the statements as shown below in Methods section (**Page 10, Lines 362-364; Page 10, Lines 367-369**):

Page 10, Lines 362-364 (Methods)

"However, scan reporting mining and informal mining activities such as artisanal and small-scale mining areas typically unrecorded in SNL mining database^{25,49}. Similarly, the historic and abandoned mining sited areas typically missed⁴⁹."

Page 10, Lines 367-369 (Methods):

"Mining areas that spatially interact with SNL mining sites were classified as

documented mines, whereas mining polygons that did not overlaid with the SNL mining database were categorized as undocumented mines according to ref⁵¹. “

13. Finally, this sentence is also unclear, but the lack of clarity here seems much more weighty than above: "We categorized the mining polygons based on the nearest-neighbor mining activities as reported in the SNL mining database.". Nearest neighbour mining activities reported in the SNL database seems to be a key term, wholly undefined notably, and one that determines the "categorization" of the mining polygons in your main composite dataset, This categorization is also unclear to me. What are we categorising? That is what are the categories? Perhaps documented versus undocumented?. But returning to the term "nearest neighbour" , having a definite spatial meaning, the sense I get from your text is that your composite dataset was intersected spatially with the SNL database, but some spatial liberties were taken to allow for intersections between elements that were not, in fact, strictly spatially coincident per se. For example, a mine in your composite that datasets could thus intersect with a mine in the SNL database, and be "categorised", Even if the two minds were some distance apart from each other. Perhaps a few 100 metres, or perhaps a few kilometres, your text does not say, but it really should specify this, and clarify what the upper distance threshold was in this approach to spatial pairing of mines.

Authors: Thanks for your good question. For each mining polygon, we identified the nearest mining site reported in the SNL mining database and categorized the primary products of mining areas accordingly. Since the database reports mining sites as points, this approach was necessary to associate primary products with spatially related sites. It is worth noted that, we did not apply a specific distance threshold; instead, we directly linked the attributes of primary products of mining activities from attributes reported by the nearest mining site, assuming the closest spatial alignment. While SNL metals and mining database have some limitation, it remains the most comprehensive source of information on primary products for mining activities at global scale (Maus and Werner, 2024b). In this context, we used the mining type reported by the nearest mining sites as the basis for categorizing mining activities.

To enhance clarity, we have added a new figure illustrating examples of how we categorized mining types for mining polygons based on the nearest mining sites (**Supplementary Fig. S8**). Additionally, we revised the Methods section to better explain our methodology (**Page 10, Lines 372-376**).

Page 10, Lines 372-376 (Methods):

*"We further identified the primary products of the mining polygons by analyzing the spatial nearest-neighbor mining activities, as reported in the SNL mining database, based on the Euclidean distance (**Supplementary Fig. S10**). Noted, the by-products of mining activities were not considered in this study due to the limitations in the attribute information provided by SNL mining database."*

[Figure Redacted]

Supplementary Fig. S10. Examples of categorized mining types based on the primary products of nearest-neighbor mining sites using Euclidean distance, as reported in the SNL mining database. a. The SNL mining types and mining polygons overlaid on Google Earth imagery. b. Categorized mining types of mining polygons using the nearest-neighbor SNL mining sites.

Reference:

1. Maus, V., & Werner, T. T. (2024). Impacts for half of the world's mining areas are undocumented. *Nature*, 625(7993), 26-29.

> MINING DEFOR

14. You write: "The result demonstrates that the deforestation in mining areas, as mapped by TMF (10,403.42 km²), closely aligns with GFW mapped deforestation (11,880.64 km²), showing an uncertainty of 12.43% in tropical regions (Supplementary Fig. S10)." . It is unclear what the uncertainty pertains to. ...Upon re reading, I now understand that the 12% reported here is a differential. The area deforested according to the TMF data differs from the area deforested according to the GFW data by 12% of the latter area. Please clarify your reporting key.

Authors: Thank you. We have revised the sentence for clarity on **Page 8, Lines 266-269**. It now reads:

Page 8, Lines 266-269 (Discussion):

*“The deforestation trends from 2001 to 2023 are consistently aligned between the two datasets, with the TME-based deforestation differing from GFC-based deforestation by an uncertainty of 12.43% in tropical regions, suggesting plausible accuracy (Methods, **Supplementary Fig. S6**).”*

Response letter

> We appreciate the editor and three reviewers for their valuable comments to improve our manuscript.

> Please find our point-by-point response (in blue) to all comments below. To make it easy to track our revisions, all modified text is marked in blue in the revised manuscript.

Reviewer #3 (Remarks to the Author):

Dear authors,

thank you for considering and incorporating our comments. We think that the manuscript now reads really well, we offer no further comments and think that the manuscript can be published.

> We sincerely appreciate your insightful comments and valuable feedback, which have greatly helped to improve our work. Thank you for your recognition on this study.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

> Thank you for your great effort in reviewing our manuscript and for your constructive comments.

Reviewer #5 (Remarks to the Author):

----SUMMARY OPINION----

The authors' responses have largely satisfied my earlier queries. I recommend publication, following some final housekeeping, namely the revision of a few terms, the inclusion of a few caveats, and the reporting of some simple data already held by the authors.

> Thank you for your encouraging words and recognition of this study. We highly appreciate all the detailed comments and suggestions you provided along with the revision. Please find our point-by-point responses and actions to your comments below.

My comments below are in no particular order, save for 'major' and 'minor'. Here, 'major' comments are not those requiring 'major revision', just the most important of comments I offer here.

----MAJOR COMMENTS----

1. Documented vs Undocumented Mines

Comment no. 1 of my earlier review raises the issue of the use of the terms 'documented' and 'undocumented' in reference to the mining areas (polygons) compiled by the present study. A mining area is said to be 'undocumented' where it does not spatially match a mining point from the separate SNL global mining database, taken here as the global 'gold standard' spatial database on mining activity. The authors have revised to explain this definition, and while clearer, I still take issue with the term 'undocumented'. The term 'undocumented' implies that such a mining area is 'unknown' or 'unregulated' by local authorities, not simply omitted from the SNL database. Again, the authors have clarified this explicitly – an undocumented mining area is one merely unrecognised by the SNL database, and may still be subject to local regulation – yet the 'undocumented' term simply says, or rather implies, that opposite, inadvertently. I strongly recommend the use of more appropriate terminology, instead of 'documented' and 'undocumented'. For example, 'recognized' vs 'unrecognized', 'observed' vs 'unobserved', or 'recorded' vs 'unrecorded', would be better. Doubtless the authors can think of even better terms.

> Thanks for your in-depth thinking about the “documented vs undocumented” terminologies used. We acknowledge that the terms may carry unintended connotations related to regulation and recognition. Accordingly, we have revised into “recorded vs unrecorded” as suggested.

2. --In Page 4 Lines 106-118—

To write to greater clarity, please use the term "IUCN Management Class" before referring to a given protected-area designation, as thusly: 'Protected areas with IUCN Management Class of Ia (strict nature reserves) or Class Ib (wilderness areas) are something something...'.

> Thanks for the suggestion. We have revised according, which is duplicated below.

Page 4, Lines 107-111 (Results):

“Specifically, according to International Union for Conservation of Nature (IUCN) management categories, 135 protected areas with IUCN management class of Ia (strict nature reserves) and 68 protected areas with IUCN management class of Ib (wilderness areas) have been observed with mining activities (Methods), which was strictly prohibited mining activities and infrastructural access.”

In this same text, you wrote:

"Noted, 5550 other protected areas have also observed mining activities; however, these areas are not reported or assigned management categories on WDPA, demonstrating the potential extensive impacts of mining operations on protected areas worldwide."

These 5550 other protected areas (PAs) are half of all PAs intersected by your mining areas. It would be worthwhile to further report on the nature of these PAs. While these

PAs will not have IUCN Management classes recorded in the WDPA, you may still at least identify those that are designated as 'national parks' - typically, PAs with such a designation have a high level of protection, consistent with IUCN Class II. In the WDPA, one of its attribute fields, possibly [Desig] (i.e., designation), or [Desig_Eng] (i.e., designation in English), will specify whether a given PA is a national park or some other designation (e.g., forest reserve, game reserve, etc.).

In acknowledging that many PAs are not described by any IUCN Management Class in the WDPA, you may cite Sloan et al. (<https://doi.org/10.1111/aje.12377>), who quantify how the WDPA omits IUCN Management Class information for something like 80% of all African PAs. This fact raises the strong possibility that the omissions of IUCN Management Class data in the WDPA is greatest in 'developing countries', where environmental regulations for sustainable mining are probably weakest. This is a valid point for your discussion.

> Thanks for the insightful comment. We have cited the relevant reference and discussed this point to support our statement that these PAs are not assigned to any IUCN Management Class in the WDPA, as described in the Results section (**Page 4, Lines 117-121**).

Page 4, Lines 117-121 (Results):

"Noted, 2,491 other protected areas have also observed mining activities; however, these areas are not reported or assigned management categories on WDPA³², particularly in developing countries where environmental regulations for sustainable mining are probably weakest. This highlights the potential extensive impacts of mining operations on protected areas worldwide."

In this same text, you wrote:

"Furthermore, 2,163 protected habitat or species management areas overlap with mining areas, and 2,184 protected nature landscape and seascape areas have observed mining activities"

Please cite the corresponding IUCN Management Classes for these two classes of protected areas, as above.

> Thanks. Revised as suggested:

Page 4, Lines 114-117 (Results):

"Furthermore, 1,749 protected habitat or species management areas with IUCN management class of IV overlap with mining areas, and 1,800 protected nature landscape and seascape areas with IUCN management class of V have observed mining activities."

3. --Methods--

-On the use of near-neighbour algorithm to match mining-area polygons compiled by

the authors to mining commodities recorded for mines in the SNL database.

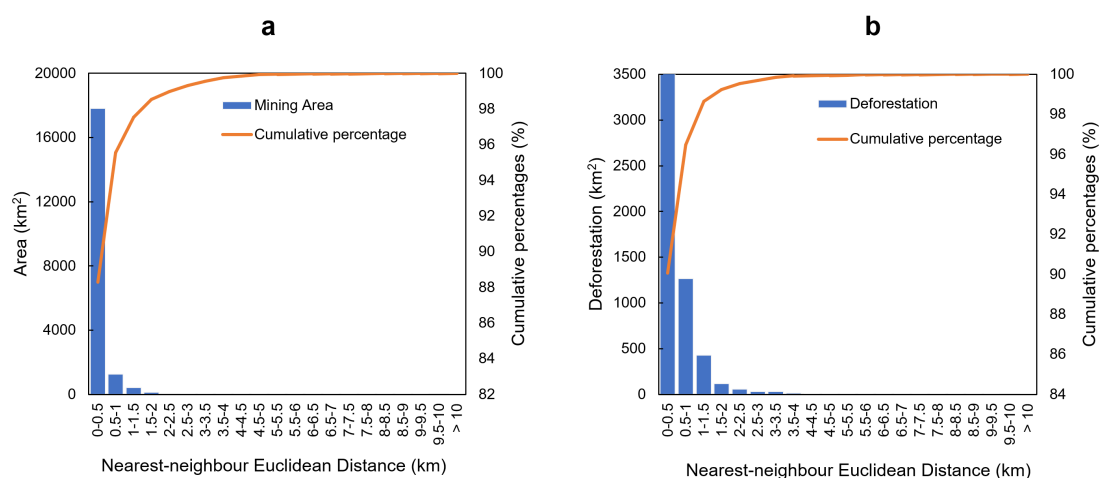
Comment 13 of my earlier review addressed your use of the nearest-neighbour (NN) algorithm. I understand your response. Your use of the NN algorithm is not unreasonable per se. I however would certainly have specified an upper limit on the allowable distance between a given mining area and a given SNL mining commodity, when establishing linkages between them via the NN algorithm. You did not. As a result, it is possible that some mining areas (polygons) in your study were assigned mineral commodities according to SNL spatial data that were, in fact, quite distant, even if also 'nearest'.

Your run of the NN routine should however at least have recorded the Euclidean distance between each mining area (polygon) and its nearest SNL data point. Given this data, I ask you to increase transparency by reporting the average, median, and standard deviation distance between each mining areas (polygons) and each nearest SNL data point, for those mining areas submitted to the NN algorithm.

> Thanks for your point about our use of the nearest-neighbor (NN) algorithm. To enhance transparency, we have now recorded the nearest distance between each mining area (polygon) and its nearest SNL data point, alongside estimates of mining-related deforestation for different nearest Euclidean distances. As shown below (Supplementary Fig. S11), approximately 95.55% of mining areas have their nearest SNL data point within 1 km, and these mining areas account for 96% of the total global mining-related deforestation. Therefore, we do believe that using the NN algorithm to link mining commodities introduces minimal bias and remains the most suitable method currently available for connecting mining polygons with the nearest SNL data points. To ensure transparency, we have detailed this information in the Methods section (Page 10, Lines 377-378):

Page 10, Lines 377-378 (Methods):

“Most mining polygons (96%) have associated with their nearest SNL mining sites within 1-km surrounding areas (Supplementary Fig. S11)”



Supplementary Fig. S11. The Euclidean distance between each mining polygons and its nearest SNL data point. a, Total mining areas and the cumulative percentage of

mining areas for various nearest distances groups. b, Mining-related deforestation and the cumulative percentage for various distance groups.

----MINOR COMMENTS----

4. --Methods--

The term "suage" appears. It is unfamiliar to me, as it will be to 99% of readers. It is seemingly misused in the text. Please replace it.

> Thank you for pointing out the typo. Revised as suggested below:

Page 10, Lines 367-369 (Results):

"The unrecorded mines highlight the potential underestimation of the impact of these mining activities by the usage of SNL mining database in previous studies."

5. --Methods--

-On the double-counting of PAs intersected by mines-

In your Methods text, it will be worth noting that counts of protected areas (PAs) intersected by mines will be slightly inflated, due to double counting, in some cases. The reason for this double counting is that, in the WDPA, each discrete designation of a PA is mapped by its own spatial polygon. If a given protected area has two separate designations, it will have two separate and highly congruent polygonal records in the WDPA. For example, where a National Park is also separately designated as a UNESCO World Heritage Site, the same area will effectively be mapped by two polygonal records in the WDPA - one for the National Park, and another for the World Heritage Site. Given this fact, your GIS overlay analysis would have identified two PAs intersected by the same mine, not just one, in this example above; hence, the double counting of PAs I mention.

In my experience, this double counting tends to be most common in the WDPA where national parks are also designated as some other, typically high-level protected area, as in the example above. However, the possibility of double counting is not restricted to PAs of high levels of conservation. Polygonal record in the WDPA can be messy and often partially overlap each other. Sloan et al., cited above, confronted this (inevitable) issue of double counting in Africa. At the global scale of your study, resolving this issue is probably not feasible. The only workable approach would be to spatially 'dissolve' all PAs in the WDPA into spatially discrete polygons. Doing so would however remove all attribute data from such polygons, thus providing a revised number of intersected PAs, now free of double counting, but absent any information on their IUCN Management Classes or similar. Again, Sloan et al. probably adopted this approach of spatial dissolving to ensure accurate counts of intersected PAs - it's been a while since I wrote that paper, but I have at least implemented that approach somewhere.

> Thank you for the point regarding overlapped protected areas in the WDPA dataset. We have reviewed the dataset and found that some polygons overlap with different

IUCN Management Classes. To address these issues, we carefully dissolved overlapping polygons based on specific criteria: we firstly dissolved polygons where the overlapping area exceeds 50% of each polygon's area, rather than dissolved all overlapped polygons. This approach accounts for cases where protected areas have different boundaries but are spatially joined. For these merged polygons, we assigned the management class based on the highest level of IUCN Management Classes present in dissolved polygons, following this hierarchy: Ia > Ib > II > III > IV > V > VI > Not Applicable/Reported/Assigned. We have revised the Results (**Page 11, Lines 107-120**) and Methods (**Page 13, Lines 468-475**) section to include the process details and updated the reported count of protected areas.

Page 11, Lines 105-121 (Results):

“Our results show that a total of 7,441 global protected areas have been or are currently affected by mining activities, based on the intersection with World Database on Protected Areas (WDPA)³¹ (Fig. 1d, Supplementary Fig. S3). Specifically, according to International Union for Conservation of Nature (IUCN) management categories, 135 protected areas with IUCN management class of Ia (strict nature reserves) and 68 protected areas with IUCN management class of Ib (wilderness areas) have been observed with mining activities (Methods), which was strictly prohibited mining activities and infrastructural access. Additionally, 497 protected areas with IUCN management class of II (national parks) and 348 protected areas with IUCN management class of III (natural monuments) are features overlap with mining areas; these are generally also forbidden with mining extractive activities. Furthermore, 1,749 protected habitat or species management areas with IUCN management class of IV overlap with mining areas, and 1,795 protected nature landscape and seascape areas with IUCN management class of V have observed mining activities. Noted, 2491 other protected areas have also observed mining activities; however, these areas are not reported or assigned management categories on WDPA³², particularly in developing countries where environmental regulations for sustainable mining are probably weakest. This highlights the potential extensive impacts of mining operations on protected areas worldwide.”

Page 13, Lines 468-475 (Methods):

“We filtered protected areas in the WDPA datasets where mining activities were observed within their boundaries. Considering the issue of overlap protected areas in the WDPA datasets³², we dissolved polygons where the overlapping area exceeds 40% of each polygon's area. For the merged polygon, we assigned the management class based on the highest IUCN management classes present in dissolved polygons, following this management level: Ia > Ib > II > III > IV > V > VI > Not Applicable/Reported/Assigned. We reported the counts of protected areas across various IUCN management categories that are potentially threatened by mining activities.”

6. --Methods--

-On the Global Forest Cover data used-

These data are cited by reference no. 12. This citation is the correct citation of the study that first created these data, but it is not an adequate citation of your source of the data themselves. Since the 2013 study by Hansen et al. (reference no. 12), the data have been updated and revised. Please therefore also cite the online source (url) of the updated data you used, including the observation years of the data, its version (e.g., version 2, version 3...), url, etc., for transparency.

> Thank you for your suggestion. We have included the information and observation years of the data. Additionally, the URL has been provided in the "Data Availability" section:

Page 11, Lines 387-390 (Methods):

"To quantify forest cover changes associated with the expansion of mining areas from 2001 to 2023, we used a high-resolution map of global forest cover change that maps tree cover change at a spatial resolution of 30m provided by Global Forest Change 2000-2023 (GFC, Version 1.11)¹² (Supplementary Fig. S4)."

Page 13, Lines 476-479 (Methods):

"Global Forest Change dataset v1.11, which annual forest loss information, can be accessed through Google Earth Engine at https://developers.google.com/earth-engine/datasets/catalog/UMD_hansen_global_forest_change_2023_v1_11."

7. --TYPO--

Line 266-267, reports "TME-based deforestation". 'TME' should be 'TMF'.

> Thank you and we have corrected it to "TMF," as shown below:

Page 8, Lines 263-267 (Discussion):

*"On the other hand, we benchmarked the GFC data against the Tropical Moist Forest (TMF) dataset⁴⁶ using a time-series cross-comparison. The deforestation trends from 2001 to 2023 are consistently aligned between the two datasets, with the TMF-based deforestation differing from GFC-based deforestation by an uncertainty of 12.43% in tropical regions, suggesting plausible accuracy (Methods, **Supplementary Fig. S6**)."*

Sean Sloan

Canada Research Chair (T2)

Human Dimensions of Sustainability

Dept. of Geography

Vancouver Island University

Office: Bld 180 Rm 220

Tel: 250-740-6128

sean.sloan@viu.ca

Reviewer #5 (Remarks on code availability):

NIL

Dear Sean,

We sincerely appreciate the time and effort you dedicated to commenting on and improving our work. Your valuable and detailed review has greatly enhanced our manuscript, making our methodology more robust and transparent.

Thank you once again for your support for our study.

Sincerely,

All authors.