EGUSPHERE-2023-2271

Reviewer Comments Response

RC1

The research has two great positive and innovative aspects. The first one is the spatial scale of the study, not limited to few stations, but widely distributed over the basin. The second one is the evaluation of the lag effect between SSC and the deforestation. To my knowledge, this was not assessed in the Amazon at basin scale.

1. It is mentioned in parts of abstract and introduction that the relationship between deforestation and fluvial sediment dynamics at large scales has not been extensively studied, in the Amazon or elsewhere. However, the authors also mention some counter examples, like in the Colombia and in Araguaia River. An example at global scale may be seen in: DETHIER, E. N.; RENSHAW, C. E.; MAGILLIGAN, F. J. Rapid changes to global river suspended sediment flux by humans. Science, v. 376, n. 6600, p. 1447–1452, 2022.
    a. Apologies if the phrasing was unclear. What was meant to be expressed is that there appears to be a gap in the literature on studies examining deforestation's impact *across and within* the entire Amazon. While there are many finer scale examinations of individual sub-basins, there did not seem to be a comparative analysis examining the spatial patterns across all Amazonian basins collectively. These statements have been rephrased in the abstract on lines 10-11 and in the introduction on lines 64-75.
    b. Thank you for sharing this paper. A mention of it has been included on line 62.
        i. "While the majority of these studies have been limited in scale, focusing on smaller basins or study areas (Bringhurst and Jordan, 2015; Latrubesse et al., 2009; Ochiai et al., 2015; Maina et al., 2013; Maeda et al., 2008), recent advancements in satellite and remote sensing technologies have allowed for larger, global scale analyses to take place (Dethier et al. 2022)."
2. As a suggestion, the authors could contextualize more what "large scale" means (quantifying the basin size, for example). Using more references about the topic is also welcome.
    a. Clarifications to the term "large-scale" have been made in a few places (lines 11 and 95)

3. About the use of rivers > 50 m: The Landsat images used have 30 meters of spatial resolution. Besides the nominal, I strongly suggest looking at the effective spatial resolution, which is not the same as nominal. Briefly, to use a section of about 50 m is

risky because could occurs spectral mixing with the riverbanks. In my experience using Landsat images, I used only sections > 100 m. In some cases, depending on the shape of the section, and the presence of sandbars, I used > 150 m. Is it possible to demonstrate that the width > 50 m does not affect the reflectance values of water?

a. We agree this is a concern in general, but we take a variety of measures to make sure we are only selecting and aggregating high quality river channel pixels including using highly accurate water masking algorithm (DSWE; DSWx) designed specifically for optically complex/turbid/vegetated waters like rivers and wetlands which. DSWE has been proven and validated in many publications (Jones 2015; Jones 2019; Taylor 2022; Huang 2018; Devries 2017) and has been adopted by NASA for water masking for many of their satellite products (https://www.jpl.nasa.gov/go/opera/products/dswx-product-suite). We select only pixels labeled as high confidence open water by DSWE and mask out all other "water" and "vegetated water" pixels. Also, note studies we cited in supplemental (e.g. Pahlevan et al., 2018) that suggest adjacency effects are probably not a concern for rivers actually, except perhaps where "dark waters" flowing through a "bright" desert landscape. Another measure we take is to calculate the median value for all high-quality river pixels within a river reach, which minimizes the impacts of potential outlier pixels matchups and SSC estimates which are derived from reach median surface reflectance values. Our methods are described in more detail in both Gardner et al., 2021 and 2023. Since our image analysis uses the exact same methods as our previously published work, decades of past research, and dozens of other publications, it is not appropriate to add this to the main text. Therefore, we added the text below to the supplemental as well as supplemental figure S1 which shows examples of high performing water masks in rivers at the limits of Landsat detection (30-50 meters wide).
b. "The dynamic surface water extent (DSWE) algorithm (Jones, 2015; 2019) was used to identify high quality open water pixels in each image (See S1). Only high confidence water pixels were selected for analysis (DSWE = 1) while all other water pixels identified as water by DSWE were removed including low confidence open water (DSWE = 2), high confidence vegetated water (DSWE= 3), and low confidence vegetated water (DSWE = 4). The Landsat quality assessment band generated by FMask was used to mask clouds, cloud shadow, snow, and ice (Foga et al., 2017; Zhu et al., 2015). The global Multi-Error-Removed Improved-Terrain (MERIT) DEM (Yamazaki et al., 2017) combined with Landsat image metadata on time, location, and solar zenith was used to calculate topographic shadow using the hillshadow GEE function to mask out shaded pixels. A cumulative cost algorithm then finds connected high confidence open water pixels connected to river centerlines to identify river channel water pixels. The remaining open water river channel pixels that are not removed by the

aforementioned masking procedures are used to calculate the median surface reflectance for each band across all pixels within a SWORD river reach."

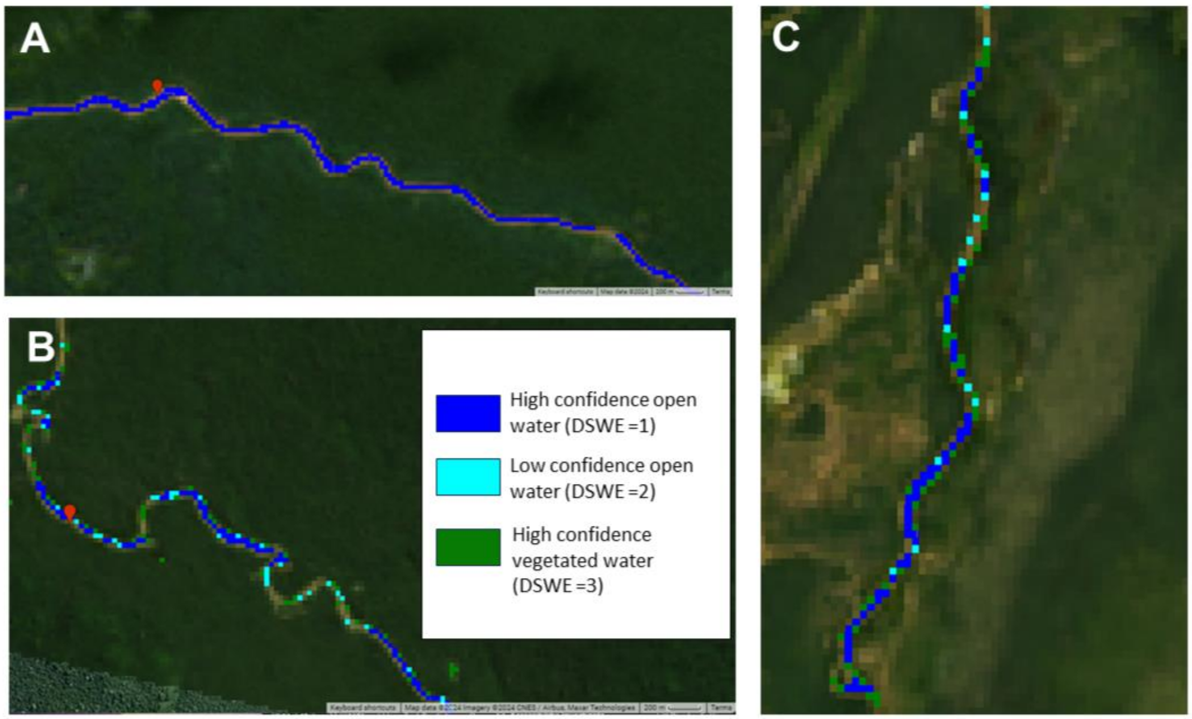


Figure S1. Examples of DSWE algorithm applied to Landsat images in three rivers in the Amazon basin that are at the edge of detection due to river widths of 30-50 meters. A) Small tributary to the Maranon River. Only DSWE =1 is shown to illustrate the selection of high-quality open water pixels only along the center of the channel (Landsat ID: LC08_L1TP_008063_20150629_20170407_01_T1). B) Rio Bacaja (Landsat ID: LC08_L1TP_225063_20150629_20170407_01_T1). C) Culuene River (Landsat ID: LC08_L1TP_225070_20150309_20170412_01_T1). Images B and C show low confidence open water and high confidence vegetated water which we exclude from analysis and are often located along the banks and sand bars. Note, any pixels not connected to the main river channel are also excluded from analysis in image processing.

DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Huang, W., Creed, I. F., & Carroll, M. L. (2017). Automated quantification of surface water inundation in wetlands using optical satellite imagery. *Remote Sensing*, *9*(8), 807.

Huang, W., DeVries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., & Carroll, M. L. (2018). Automated extraction of surface water extent from Sentinel-1 data. *Remote Sensing*, *10*(5), 797.

Jones, J. W. (2015). Efficient wetland surface water detection and monitoring via landsat: Comparison with in situ data from the everglades depth estimation network. *Remote Sensing*, *7*(9), 12503-12538.

Jones, J. W. (2019). Improved automated detection of subpixel-scale inundation—Revised dynamic surface water extent (DSWE) partial surface water tests. *Remote Sensing*, *11*(4), 374.

4. In the figures 1 and 3 (maps) the authors could include a scale bar and to increase a little bit the size of geographical coordinates.
    a. Thank you for this suggestion. The figures have been updated.

5. On the deforestation database: I suggest a view on the Mapbiomas dataset (https://brasil.mapbiomas.org/produtos/) which is a well-documented program that provides land-use and deforestation data for Brazil and South America.
    a. That is indeed an interesting database, thanks for sharing. A note about this database has been included in the discussion section (line 509-511)
        i. "Similarly, the use of land-cover datasets designed for Amazonian type landscapes such as MapBiomas (Souza et al., 2020) which covers Brazil, may unveil more regionally specific relationships than when using global land classification algorithms."

6. About the time of the year of images: acquiring images in the dry season is not an advantage due to coincidence with deforestation period because the sediment production occurs in the wet season.
    a. Thanks for bringing this up. A line noting this has been added (see lines 162-164).
        i. "While SSC data collected during the wet season may be preferable for studying deforestation driven changes in sediment dynamics, using wet season imagery was not possible due to high levels of cloud coverage."

7. Lines 160-162: About the estimation of high SSC values: Besides the factors mentioned by the authors, there is an important challenge related with the physical relationship between SSC and reflectance. As SSC increases to higher values (like > 1000, > 3000 mg/l), the reflectance sensibility to SSC decreases. So, little reflectance changes not related to SSC (low radiometric resolution, Sun glint effect, optical components mixtures, among any other) can cause large SSC estimations errors at high concentrations.
    a. Yes, there are many potential limitations of remote sensing of SSC especially at high values and we are well aware. Note, our manuscript does not focus on high SSC values, only relative change in concentrations over time. The purpose of this discussion is solely to inform potential users of the limitations associated with the published SSC data. We stated in the main text methods that our algorithm can estimate up to 2500 mg/L, which is below the values the reviewer suggested Landsat images may become saturated or increasingly non-linear. Machine learning is also inherently designed to handle non-linear relationships between surface reflectance across different bands and SSC. To our knowledge, there is not a known upper limit of SSC detection due to the sensor, but there are various

guesses. If the reviewers are aware of studies that have quantitatively found an upper limit of SSC detection from Landsat using machine learning, we would love to know about it. We already mentioned these known issues brought up by the reviewers (like sunglint) in supplemental material and added the text below to main text:

    i. "It should be noted that the SSC database focuses on surface concentration and may not accurately capture high SSC values due to factors such as cloud cover, sensor band saturation at high SSC, and a lack of high SSC field measurements for model training."

8. About the regression analysis: Is not necessary a problem the small R^2 value of 0.13 but should be included the p-value (is significant the 0.13 value?). Besides that, as the regression can be easily affected by outliers, it is interesting to show the plot for this case (could be in the supplementary materials).
    a. Thank you for this suggestion. After double checking the R^2 values, it appears that there was a mistake where r was calculated instead of R^2 and the shifts were not applied prior to performing the analysis. The values in the chart have been updated to reflect r values and their associated p-values. *Please note however, this does not change the conclusions of the analysis.*
    b. Further, plots have been included in the supplementary materials document (S5, S6, S7)

9. In the discussions I suggest including a topic about the limitation related to evaluating only concentration values rather than sediment transport values. The amount of sediment transported by a river is a simple function of SSC times discharge, but the lack of discharge data at the same scale of SSC data would make your research inexecutable. But the discharge is a great source of uncertainty, because at the same SSC, different discharge values yield different sediment transport rates. Perhaps, in the more deforested areas, the larger discharge values (which decreases SSC values) could be masking an even larger effect of deforestation, for example.
    a. Thank you for this suggestion. An additional paragraph was added to address this potential limitation (lines 515-526)
        i. Another limitation of this study relates to the use of concentration values over sediment transport values (flux; kg/s) to assess deforestation-sediment relationships. In regions characterized by extensive deforestation, it is possible that the presence of larger discharge values (which inversely decrease SSC), act as a masking factor, potentially obscuring a more substantial impact of deforestation on sediment dynamics than what is observed in this study. Deforestation has been

observed to increase surface runoff in many parts of the world (Guzha et al., 2018; Potić et al., 2022; Zhao et al., 2022) leading to the dilution of SSC, and making it challenging to discern the true extent of the influence of deforestation on sediment concentration. While these observations are often more present in water limited watershed (Zhang et al., 2017) and observations of this phenomenon have been somewhat limited in the Amazon (Lucas-Borja et al., 2020; Voldoire and Royer, 2004), there may be overlooked decreases in SSC associated with increasing discharge values. Utilizing sediment transport values may unveil an even more profound effect of deforestation on sediment dynamics in these heavily deforested areas, not elucidated in this study. To better capture the true extent of deforestation's impacts, future examinations on deforestation-sediment dynamics in the Amazon should consider using both SSC and discharge data in their analyses.

RC2

This study utilizes a novel remotely sensed river suspended sediment concentration (SSC) dataset, and investigates the relationship between deforestation and SSC dynamics, which is lacking in this field, especially in the Amazon. An innovative XGBoost method is applied for SSC inversion, effectively enhancing accuracy. Deforestation and SSC dynamic analysis uses time-lagged cross-correlation (TLCC) analysis, capturing significant patterns of hydrogeomorphic response to reforestation on SSC time delays. The conclusions indicate that deforestation has a significant impact on sediment dynamics. This article provides valuable insights and references for environmental management and policy-making in the strategically important Amazon region.

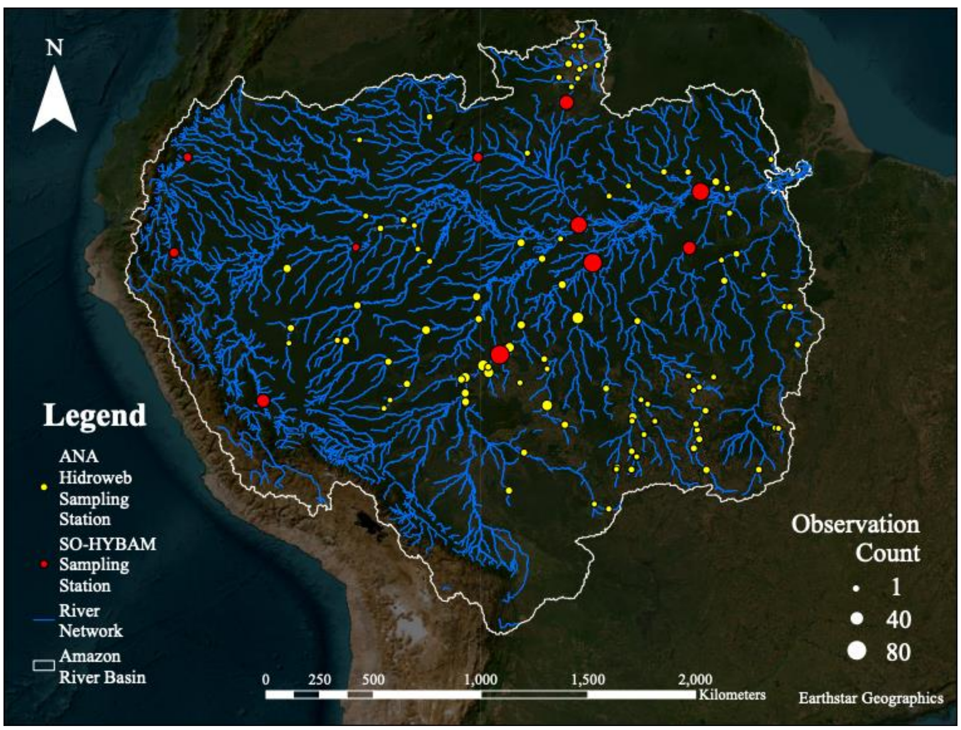
General comments are:

1. The introduction and discussion could be further improvement. The introduction could better highlight the advantages of remote sensing, such as large spatial scale, long-time coverage, and low cost, and when combined with machine learning it can build high-quality river SSC datasets. For the discussion, more references and comparisons with deforestation vs SSC in other regions should be comprehensively compared and this will help place the study in a larger context and bigger framework.
    1. Thank you for this suggestion. Lines on the benefits of remote sensing have been included in the introduction (lines 102-109). An additional comparison to the finding of other studies/broader impacts has been included in the discussion section (lines 463-481).
        1. Lines 102-109: Recent advancements in remote sensing technologies have revolutionized our capacity to monitor and analyze environmental changes. The application of remote sensing in environmental sciences

offers unparalleled advantages, including the ability to cover large spatial extents, provide long-term data coverage, and ensure data acquisition at lower costs compared to traditional methods. When coupled with machine learning algorithms, remote sensing data can be transformed into high-quality, comprehensive datasets (e.g., Global Forest Change Dataset (Hansen et al., 2013), World Settlement Footprint (WSF; Marconcini et al., 2020), and LandCoverNet (Alemohammad and Booth, 2020). This synergy is particularly effective in building high-quality river SSC datasets, enabling a more nuanced understanding of sediment dynamics over vast geographical areas and extended timeframes.

2. Lines 463-481:
    1. Other factors, such as basin size can also affect the discernable influence of deforestation on sediment. In previous studies, strong relationships between deforestation and sediment are observed within relatively small river basins . For example, in New Zealand's Waipoa River System, which encompasses an area of 1,987 km2, Kettner et al. (2007) observed a sixfold increase in suspended sediment discharge at the river outlet due to deforestation. In Wisconsin's North Fish Creek (drainage area of 122 km2), deforestation and human settlement was observed to increase sediment to 4-6 times pre-settlement rates (Fitzpatrick and Knox, 2000). In larger catchments, however, the influence of deforestation on sediment appears to be much lower as more variability is introduced into the relationship. For example, within Spain's Ebro River basin (85,530 km2), long term anthropogenic land use was revealed to increase sediment by 35%, from 30.5 Mt yr.−1 to 47.2 Mt yr.−1 over a 4000-year period (Xing et al., 2014). In the Magdalena River Basin (273,459 km2), Restrepo et al. (2015) observed a 9% increase in sediment load attributable to deforestation. Our observations within Amazonian major tributary basins, align with these overall trends. The majority of basins within our study were each greater than 100,000 km2, therefore it is to be expected that the discernable influence of deforestation on sediment dynamics may exhibit greater variability or attenuation compared to smaller basins. Some studies have found that the influence of land use and land cover change on runoff (Blöschl et al., 2007) and discharge (Zeilhofer et al., 2018; Rodriguez et al., 2010) decreases with watershed size. Likely, in the case of sediment transport, transport processes, such as deposition and dilution of the deforestation-sourced sediment, are magnified at these larger basin scales. The increased occurrence of these processes allows for large basins to have a greater buffering capacity, and therefore produce a small sediment delivery ratio (Walling, 1983, Walling 1999). These relationships likely result in the observed variable relationship strengths (Figure 6).

2. Data and methods section does not provide important information on the source of the field-measured data, and the total number of match-ups used to train machine learning algorithm;.

1. We added new text (lines 167-176) and a new figure (Figure 3) which provides more in-depth information regarding match-up points.
    1. An Extreme Gradient Boosting (XGBoost) model was trained using 1112 matchups, or satellite and SSC field observations that occurred within the same day or +/- 1 day following Ross et al., (2019). Field observation data used in the match ups were obtained from both gauging stations and grab samples spread throughout the basin (Figure 3). Data from 121 gauging stations were sourced from ANA Hidroweb (Water Resources National Agency, 2020). ANA operates a network of automatic monitoring stations equipped with sensors to continuously measure various hydrological parameters, including sediment concentration, water level, flow velocity, and other water quality parameters. Additionally, grab sample data were collected from fourteen SO-HYBAM gauging stations (Institut national des sciences de l'Univers, 2021). While the number of sampling locations is limited, these stations are strategically positioned throughout the Amazon Basin to represent diverse hydrological features such as main rivers, tributaries, lakes, and other water bodies. These grab samples are collected manually by field personnel at regular intervals from the water surface (Institut national des sciences de l'Univers, 2021).
    2. 

3. Many statements should be supplemented with statistical values to enhance readability and persuasiveness.

Good point, additional statements have been included in the results section on lines 335- 337, and 382-383and additional reference to the statistical values have been made in the first two paragraphs of the discussion section.

- Lines 335- 337:
  - These results indicate that deforestation has a strong, direct impact on sediment dynamics, with more intensive deforestation activities leading to quicker hydrogeomorphic responses in the affected basins.
- Lines 382-383:
  - This further implies that deforestation intensity directly influences the timing of sediment response, with more immediate hydrological alterations occurring in basins experiencing higher deforestation rates.

Specific comments:

Main text:

1. Abstract: line 11, "...or elsewhere...", As far as I know, there have been studies looking into how vegetation degradation affects the dynamics of SSC.
    1. The language in this section has been updated.
2. Introduction: The third and fourth paragraphs could be merged, mainly referring to the progress and limitations of the study; the fifth and sixth paragraphs could be merged, mainly referring to the difficulties in obtaining data and the limitations of the model.
    1. Thank you for this suggestion. We have opted to keep the third and fourth paragraphs separate as they detail two different studies. However, the fifth and sixth paragraphs have been merged.
3. Line 142, "Suspended Solid Concentration (SSC) concentration (mg/L)...." Please use suspended sediment concentration (SSC, mg/L) consistently throughout the text.
    1. The language has been updated to “Suspended Sediment Concentration”.
4. Section 2.3, No details have been shared on the training parameters of the XGBoost and how were they chosen. The correct spelling of xgboost is ‘XGBoost’, plus the full name. Model evaluation metrics I think should include RMSE, which is widely used and accepted; data matching should give more details, e.g. what is the source of the measured data, what is the specific matching strategy, and does it add additional manual control criteria?
    1. The RMSE was added to the main text and Figure 4 as requested. Note, the remote sensing of water quality community has coalesced around error metrics such as MAE and relative error (sometimes referred to as relative bias), see Seegers et al., 2018, hence our emphasis on those metrics. We edited the name XGBoost and added the full name as requested. We added more details on the XGBoost model, see below sentences added to main text (lines 184-186) :

    1. "XGBoost has four hyperparameters that were tuned using a grid search across all possible hyperparameter combinations with each parameter range centered around the default parameter values."
2. The field measurement sources were added to methods in main text (lines 172-174):
    1. "Data from 121 gauging stations were sourced from ANA Hidroweb (Water Resources National Agency, 2020). ANA operates a network of automatic monitoring stations equipped with sensors to continuously measure various hydrological parameters, including sediment concentration, water level, flow velocity, and other water quality parameters. Additionally, grab sample data were collected from fourteen SO-HYBAM gauging stations (Institut national des sciences de l'Univers, 2021). While the number of sampling locations is limited, these stations are strategically positioned throughout the Amazon Basin to represent diverse hydrological features such as main rivers, tributaries, lakes, and other water bodies. These grab samples are collected manually by field personnel at regular intervals from the water surface (Institut national des sciences de l'Univers, 2021)."

3. The matchup strategy is standard and conservative. We find satellite observations +/- 1 day from the field measurement and extract surface reflectance around the sampling coordinates. This remote sensing side of matchups was explained in supplemental. And yes, we perform manual and automatic quality control on matchups. We added details about generating and quality controlling the matchups to supplemental:
    1. Main text: "An Extreme Gradient Boosting (XGBoost) model was trained using 1112 matchups, or satellite and SSC field observations that occurred within the same day or +/- 1 day."
    2. Supplemental text: "The dynamic surface water extent (DSWE) algorithm (Jones, 2015; 2019) was used to identify high quality open water pixels in each image (See S1). Only high confidence water pixels were selected for analysis (DSWE = 1) while all other water pixels identified as water by DSWE were removed including low confidence open water (DSWE = 2), high confidence vegetated water (DSWE= 3), and low confidence vegetated water (DSWE = 4). The Landsat quality assessment band generated by FMask was used to mask clouds, cloud shadow, snow, and ice (Foga et al., 2017; Zhu et al., 2015). The global Multi-Error-Removed Improved-Terrain (MERIT) DEM (Yamazaki et al., 2017) combined with Landsat image metadata on time, location, and solar zenith was used to calculate topographic shadow using the hillshadow GEE function to mask out shaded pixels. A cumulative cost algorithm then finds connected high confidence open water pixels connected to river centerlines to identify river channel water pixels. The remaining open water river channel pixels that are not removed by the aforementioned masking procedures are used to calculate the median surface reflectance for each band across all pixels within a SWORD river reach. We applied the same procedure to extract Rs

over field measurement sites to generate matchups. Matchups must occur within a +/- 1 day difference between field and satellite measurements and the median surface reflectance values were calculated for high confidence water pixels within a 500-meter buffer from the field sampling coordinates. Quality control steps include removing any match-up or reach level pixel aggregates that had less than 5 remote high confidence water pixels to remove observations impacted by neighboring non-water pixels. Matchups were also manually inspected to remove field sampling sites with coordinates that to not correspond to Landsat visible rivers and field measurements > 7500 mg/L."

5. lines 155-160, The models are very accurate, which is to be congratulated, but the training data is different from other studies, and the study area is different, so I don't think this is an appropriate comparison. I think it would be appropriate to run the models through your data.
    1. Thanks, this is a good point. We recognize these citations use different training data, image processing, and workflows, and we are only suggesting that our error metrics are reasonable and within the same range of SSC algorithms also developed over large extents (e.g., countries, continents, global) using data-driven approaches and Landsat. This paper is not a model comparison, and comparing different published models is beyond its scope. Such an analysis would be inappropriate as data-driven algorithms for estimating water quality are often not transferable to different waters or regions with different optical properties not included in the training data. We revised the text as follows (lines 190-194):
        1. "In comparison, Gardner et al. (2023) reported a relative error of 0.59 for rivers in the USA, while Dethier et al. (2020) reported a relative error of 0.73 for rivers on a global scale. However, we focus on MAE and relative error as suggested by Seegers et al., (2018). While these studies are based on different regions and training datasets, they provide valuable benchmarks for evaluating the performance of the model in predicting SSC across diverse geographic and environmental settings.
6. line 149, "......time series analysis (Roy et al., 2016; Gardner et al., 2021)." Please specify whether Roy's or Gardner's sensor fusion method is used.
    1. The citation has been updated to reflect Gardner et al. (2021).
7. Section 3.3, Table 4. and Table 5., I am not sure why the Coefficient of Determination for the L0 lag group is 0.13 while the P-value is 0.0022. The authors also mention that the regression results for the other groups are poor, suggesting that sediment dynamics may be affected by a combination of other factors such as dam building, mining, agricultural practices, and urbanization. I would suggest that the authors may consider attributing these drivers in future work or using machine learning methods? A recent paper by Zhou and Li et al. "Distinguishing the multiple controls on the decreased sediment flux in the Jialing River basin of the Yangtze River, Southwestern China." Catena 193 (2020): 104593" quantitatively discussed these issues and it could be considered as a potential reference for comparison.

1. Apologies- an error has been corrected where the r values in Table 4 were reported as R^2 .
2. To clarify, the correlation coefficient (r) found in Table 4 is unrelated to the p-values found in Table 5. The r values found in Table 4 is calculated by plotting the annual SSC values against the percent of the basin deforested that year (for basins in each lag group). The variation in r values between lag groups suggest that the geomorphic response to deforestation is highly specific to each sub-basin (i.e., no strong association) except for regions with a relatively high intensity of deforestation (L0).
3. The p-value found in Table 5 describes the significance of the Mann Whitney U test. The small p-value (<0.05) for the L0 group indicates that there are significant differences in the amount of deforestation occurring within the L0 group. Groups with a negative standardized SSC value (indicating a decreased level of SSC over the 20-year period) tended to have less deforestation compared to groups with a positive standardized SSC value.
4. Thank you for sharing that paper. An expansion on the matter was added beginning on line 436 (in the discussion section) referencing Zhou et al. (2020) (lines 506-509).
    1. "Future work could explore attributing changes in SSC to specific anthropogenic activities and regenerative processes in the Amazon. For example, attribution methods previously established to examine controls of sediment flux in other basins such as the Jialing River Basin (Zhou et al., 2020) and the Yellow River Basin in China (Wang et al., 2016) could be adapted to the Amazon."

Figures:

Fig.1: Missing scale; longitude and latitude labelling text is not readable; compass and the actual bearing do not match; captions in the description of the line width represent the width of the river, please give a reference value to be placed in the legend.

1. Please see the revised Figure 1. To our knowledge, the compass and actual bearing is correct.

Fig.3: Figure a should reflect the number of matches N, section 2.3 mentions that the matchups are 1200, but the figure doesn't seem to have such a large number. Figure b should be modified as Fig.1, and the font size should be consistent.

1. The figure represents the validation (hold-out test) data. We've updated the caption to reflect this.
2. Corrected, please see the revised Figure 4 (was listed as Figure 3 before).

Fig.7: Explanation of missing box-and-line diagram elements.

1. The figure and figure caption has been updated. The caption now includes additional information on the removed outlier points.

Supplementary:

The appendix seems the same as Gardner et al., 2021, in which case I would suggest that the application of the Gardner et al. method should be stated directly in the text, without the need for a supplementary, and that the appendix leaves a lot of questions unanswered.

Thanks, we added statements to methods section 2.3 saying methods are based on Gardner et al., 2021 or 2023. This was also mentioned in supplemental. If the reviewers would prefer to remove supplemental entirely, or add supplemental text to the main text, that is fine. But it seems that all reviewers requested more methods details which can largely be found in our previously published work. Therefore, we chose to put details in supplemental material so previously published work does not take up space in the main text.

"Suspended Sediment Concentration (SSC, mg/L) data was acquired using Landsat Collection 1 and machine learning using the methods described in Gardner et al. (2023) and (2021) ..."

"The model was built using methods described in Gardner et al., (2023) and ..."

More details have been added to the supplemental material. The reviewers did not advise which details they wish to see, but we hope this is satisfactory. If not, please see our previously published work using these exact methods and the previously established methods we applied and cited from the literature.