REVIEWER RESPONSE STATEMENT

For ESurf manuscript number egusphere-2023-156 *"*Sediment source and sink identification using Sentinel-2 and (kayak-based) lagrangian river turbidity profiles on the Vjosa River*" by Jessica Droujko, Srividya Hariharan Sudha, Gabriel Singer, and Peter Molnar.*

Dear editors and reviewers,

Thank you for the positive assessment of our manuscript. We have found the comments helpful in our review and believe that their implementation has positively contributed to our final manuscript. We have implemented the following major changes:

- *Figure 1 now includes the catchment elevation, catchment delineation, and beginning/end of the kayak measurements*
- *A further refined clarification of the methodology (cloud masking, deep-water pixel selection, regression analysis, etc.)*
- *A supplementary table including the correlation coefficients of all Sentinel-2 bands vs. log(Turbidity) for each station*
- *A supplementary figure showing the regression equation (1) in 3D space*
- *Removal of Table 3 (low information table, the information was instead added to the main text, lines 360-361)*
- *Better explanation of adjacency effects*
- *We have changed the title to better reflect the main content of the paper*

We hope that the final manuscript now reaches the standards of Earth Surface Dynamics and can be accepted for publication.

Sincerely,

Jessica, Srividya, Gabriel, and Peter

Reviewer 1:
The manuscript (MS) entitled "Sediment source and sink identification using Sentinel-2 and (kayak-based) lagrangian river turbidity profiles on the Vjosa River". Overall, the manuscript is well written. However, I have a few specific questions which have to be addressed for a better understanding of the concept. The following comments may help the authors to improve the manuscript.

1. **Page 3, Paragraph 2**: Although authors have mentioned literature for remote sensing-based SSC estimation methods, please categorized the existing methods, viz., semi-analytical approach and empirical approach. Subsequently, describe the pros and cons of approaches to motivate the present study.

   These approaches were categorized as empirical (line 84 – all line references are to the new article text). In addition, we have briefly described the pros/cons of this approach in lines 85-88 (e.g. this method has been successfully implemented/validated by other studies however, the work is not directly physically-based). We are not sure if the reviewer has seen that in a later section (3.1), we have *further categorized the various empirical equations that use reflectance bands to predict turbidity or sediment concentration.*

2. **Page 3, Paragraph 2**: For narrow-width rivers, there are several existing research studies are available using Landsat, Sentinel-2, Synthetic Landsat (derived from MODIS), and other commercial satellites (e.g., CubeSat, Planet lab satellites, etc.). Please highlight those studies in this section.

    For narrow rivers, we could not find any additional studies (to the ones already listed in section 3.1) that measure SSC and/or turbidity. After the reviewer's suggestion, we have found some narrow river studies investigating discharge and the presence of streamflow (using CubeSat). However, we feel that including these studies is not relevant to the overall message we are trying to deliver in the introduction because we are trying to come up with a methodology to extract SSC and/or turbidity from narrow rivers.

3. **Figure 1**: Please provide the elevation of the study site and demark the tributaries, main river, and catchment area distinctly.

    We have incorporated the suggested changes.

4. **Page 6, Paragraph 2, Section 2.2**: While highlighting the methodology, the authors mentioned that they have first established the regression model in the section where in situ data is available and then, the models are used in the entire mainstream. I agree with this hypothesis is true for lakes and estuaries where the water is nearly stagnant or less affected by lateral flows; however, adopting the same for dynamic river systems is still questionable.

    Although it is true that lateral flows certainly affect the turbidity of the river in the lateral direction, the degree to which lateral flows affect the turbidity differs from reach to reach and riffle to riffle. Here we wanted to see the limits of applying a single regression to a dynamic system across a long river reach representing an entire river network, and quantifying upstream/downstream changes in NTU from this regression. We know that larger errors are introduced from the application of one general regression to the entire stem (as opposed to individual regressions for the reaches around the individual in-situ stations - further explained in section 3.1) and we quantified these errors and compared the R^2 of the regressions (overall vs. single stations) in section 3.1.

5. **Page 6, Paragraph 2, Section 2.2**: What is the significance of the citation "Talluto (2023)" in this statement?

    We apologize. This was an incorrectly placed citation. The citation was removed (line 144).

6. **Page 6, Section 2.2.2**: Why the ACOLITE correction was adopted in this study? Although there are several other atmospheric correction algorithms are available for water colour

inversion studies in the literature, why its only ACOLITE? It may be SNAP (specially developed for Sentinel) or SeaDAS etc.

Although there are many atmospheric correction algorithms for water, ACOLITE was specifically developed for atmospheric correction over inland waterbodies, and from our understanding it is a widely used correction algorithm in RS also with Sentinel data. We did not intend to compare the effect of different atmospheric correction algorithms in this paper, but this may indeed be a relevant and interesting question for future studies.

7. **Page 6, Paragraph 2, Section 2.2**: What is the average depth of water for deep water? Is there any threshold defined for deep water?

The average depth of the water for deep water pixels is unknown. An explanation of how the threshold (in sr^{-1}) was selected is explained in section 2.2.2, but it was mostly done through trial and error by visual inspection of the Sentinel images in addition to personal observations during the kayak sampling. We have added an additional clarification in lines 174-176 in the revised manuscript.

8. **Page 6-7, Paragraph 1, Section 2.2**: When the authors considered the rivers are narrow, especially the tributary, the water pixels are more pronounced to the land adjacency effect. How to address those effects while considering this framework?

To investigate the effect of land adjacency, in the sense of terrestrial effects on reflectance properties in vicinity of water bodies, as did Paulino et al. (2022) would go beyond the scope of this work. Past studies (Pahlevan et al., 2018) suggest that narrow rivers may be less affected by land adjacency than large water bodies. However, adjacency effects are important and should be investigated on narrow (and large) rivers. We have clarified this in lines 183-188. We have also added a small section in the outlook addressing these issues (lines 416-418).

Paulino, Rejane S., et al. "Assessment of adjacency correction over inland waters using sentinel-2 MSI images." Remote Sensing 14.8 (2022): 1829.
Pahlevan, N., Balasubramanian, S. V., Sarkar, S., and Franz, B. A.: Toward long-term aquatic science products from heritage Landsat missions, Remote Sensing, 10, 1337, 2018

9. **Page 7, Paragraph 1, Section 2.2**: Please mentioned the type of interpolation followed during the processing of the images.

We simply split each pixel into four pixels (to fit into the 10x10m grid). We adjusted the wording from "interpolation" to "resampled" to make this clearer (line 171).

10. **Page 8, Section 3.1**: From the results, it was found that the best correlation of the *in-situ* turbidity was found in the blue band of electromagnetic radiation also represented by Eqn. 1. However, it is completely different from the findings of other studies as the majority of the literature tested and proved the sensitive bands are red and NIR bands

(as elaborated in *Page 9*, *Section 3.1*) which is also obvious due to the peak occurred in that region only. So, on what basis, the blue band will be considered?

This is a valid question, and one that we discussed and considered in much detail in our work. It is true that most previous studies do use red and NIR bands, but as the reviewer probably knows, some of the outlined studies in section 3.1 also used the green band. Our take on this was, that we do not want to prescribe a-priori bands for the regression model, but rather let the data dictate what is the best fit. This is because there are many factors that affect the optical properties of inland water bodies, and since these empirical equations are not physically-based, we decided to investigate all of the available bands and fit models to several combinations of bands. Using the Bayesian Information Criterion (BIC), we then selected an empirical model using only those bands that act as the best predictors for turbidity in our catchment. We acknowledge that this may be a site-dependent outcome and is something worth looking into in future research.

11. **Page 8-9, Section 3.1**: What are the correlation values in the red and NIR region when the reflectances of ROI (region of influence) are compared against the in-situ SSC or Turbidity? Please provide scattered plots for individual stations (like region-specific).

    We see the benefit of the referee's suggestion to investigate the effect that turbidity has on the red and/or NIR bands (since these bands were discussed already in section 3.1 and used in several past studies). We have provided a correlation table in the Supplementary Materials of log(Turbidity) and all of the Sentinel-2 bands for all four stations separately and together. In this table which will accompany the revised paper, it can be seen that the red band in some of the sites compares to the blue band or even surpasses it locally, however, the NIR band is not a good predictor. We refer to the Supplementary Table in line 210.

12. **Page 8-9, Section 3.1**: Band ratio techniques are a very popular technique for both SSC and turbidity estimation along the river claimed in many literatures. Include observations while comparing the developed model with the existing models.

    This is correct, and we have included in our fitting procedure band combinations as well, perhaps this was not clear to the reviewer. We are unsure what the reviewer means by "include observations." Fitting of band reflectance is conducted with observations of turbidity. During our BIC analysis, we had investigated empirical relationships using typical band ratios, however eqn. (1) was the best predictor of turbidity in our catchment overall. We have added a clarification in lines 211-212 that we have also investigated models with band-ratios as predictors.

13. **Page 8-9, Section 3.1**: As this study emphasized narrow-width rivers, please include the studies in this section that were exclusively conducted along rivers using the finer resolution satellite images (Sentinel, Landsat, or any derived satellite products).

    Here, we wanted to give an impression of the performance range (R^2) in similar studies and there were very few studies conducted on narrow rivers. For this reason, *we included studies from wide rivers and other inland water bodies as well. If the reviewer is thinking of a particular narrow width study that we have missed please let us know.*

14. **Page 8-10, Section 3.1**: If the particular methodology will shift to another river system worldwide, how to implement this? Please add a comprehensive discussion regarding this.

    Thank you for this comment, it is in fact important for applications. We have added a short section in the outlook (lines 418-422).

15. **Page 10, Section 3.2**: How the land adjacency effect was addressed while mapping the turbidity along the narrow-width river section?

    Addressing the effects of adjacency goes beyond the scope of this work (see reply to comment #8).

16. **Page 14, Section 3.3**: With reference to #Comment 15, while averaging the 106 number of image time series river pixels, the uncertainty of Turbidity estimation is still existing and due to the land adjacency effect, the estimation may be quite high near the bank. In this aspect, it will be very much difficult to judge the concept provided in this section. Please provide substantial evidence regarding the claim.

    We assume that land adjacency affects all of the river pixels. It would be interesting to see a future study that investigates at which point (e.g. at which river width) does land adjacency not affect the pixels in the middle of the river. With this information, we could determine where and when to remove the effects of land adjacency. We have added a small section in the outlook addressing these issues (lines 416-418).

Reviewer 2:

This manuscript uses satellite images and field data to generate SSC profiles along a relatively un-impacted river to identify potential sediment sinks and sources and calculate sediment flux. I think this is a nice proof of concept showing the use of SSC profiles. The methods, analysis, and interpretation are sound. I think the most impactful piece is the ability to identify sources/sinks from a SSC profile. No need to work it out further in this proof of concept paper, but I do recommend further emphasizing/exploring sources/sinks if possible.

There are a few points that need clarifying in writing (and potentially analysis). One, I was expecting the kayak-based turbidity data to be more important to the manuscript since it was in the title. The kayak data is super cool, but my understanding is this data wasn't used for anything and only appears in 1 paragraph. I may recommend simplifying the title (i.e. removing "kayak-based lagrangian") and/or finding a way to integrate the lagrangian data more into the results/interpretation. See these papers for more info on how to interpret snapshot vs lagrangian data.

We thank the reviewer for their assessment. And we have read the suggested papers which derive the rate of SSC change (and separate the changes in the SSC rate over time vs. the changes in SSC rate over space) along a simulated river reach (Ensign et al., 2017) and coupled the longitudinal profile with a kinetic model (Hensley et al., 2014). We believe that undertaking such an analysis would go beyond the scope of this work because the aim of this paper was to see if Sentinel-2 data can be used in a narrow river to identify systematic changes that may be related to sediment source activation with minimum calibration and ground-truthing. For this reason, we have revised the title of the article to "Sediment source and sink identification using Sentinel-2 and a small network of turbidimeters on the Vjosa River." We have used the kayak-based survey as a proof of concept of variability in SSC along the river system only, because we cannot compare this data easily with the snapshot of the river by satellite at a given instance in time.

Ensign, Scott H., Martin W. Doyle, and John R. Gardner. "New strategies for measuring rates of environmental processes in rivers, lakes, and estuaries." *Freshwater Science* 36.3 (2017): 453-465.

Hensley, Robert T., Matthew J. Cohen, and Larry V. Korhnak. "Inferring nitrogen removal in large rivers from high-resolution longitudinal profiling." *Limnology and Oceanography* 59.4 (2014): 1152-1170.

See specific comments for a few other points that could be clarified (mostly the water masking/deep water masking).

Specific comments

Line 34: Be careful about framing SSC as pollutant, which it can be in small/clear/pristine streams, but in this river, I expect SSC is not a pollutant, correct? Also, the conclusion highlights how SSC in mountain areas is an important sediment source for downstream sites and this may confuse readers if SSC was presented as a pollutant in the intro. Consider reducing language about SSC as pollutant.

We have revised the language to make it clear that one aspect of high turbidity is its function as a pollutant, e.g. for drinking water, and it is a standard water chemistry attribute (line 32). But we agree with the reviewer that the main focus on SSC here is to identify catchment sediment sources, fluxes and yields, and in terms of quantifying catchment erosion and sediment budgets.

Line 44. Consider breaking this into 2 sentences.

Line 72: Technically, colored dissolved organic matter (CDOM) since not all DOM is visible. Consider adding "colored" or "chromophoric" in front of organic matter.

We have incorporated these last two suggestions (lines 48 and 77).

Line 165: Please clarify the water masking procedure that was used. Also, please clarify how the deep water threshold was used, was this used to filter "shallow water" so only deep water pixels were used in analyses?

> *The water masking procedure was used to eliminate all water except the deep water. In the Sentinel-2 image SWIR band, we selected a cutoff value (pixel values are from 0-1, cutoff value was selected based on visual inspection of known deep sections of the river). We have added a clarification in the text (lines 172-176). We select deep water because in shallow water the satellites measure the reflectance of the river bed under low turbidity conditions.*

Line 170: what cloud cover % threshold was used? I think you said cloud free, but does this mean only imaged with 0% cover? Please show threshold used.

We used the default parameters in FMask (MATLAB-based masking algorithm "Function of Mask" (Qiu et al., 2019))

> *for Sentinel-2 images, therefore a cloud probability threshold of 20% was used in the processing. The mask does not work by selecting images with a certain percentage of cloud cover in the image (e.g. an image with 30% cloud cover may still be used in our analysis). The mask works by detecting the clouds (and cloud shadows) present in the image and removing these parts of the image (or "masking" them). We have added a clarification in lines 180-182.*

Line 178: What was the maximum time offset you allowed between satellite image and in-situ turbidity? Please state this.

> *The maximum time offset between the satellite image in situ turbidity acquisition times is 15 minutes. We have added this info at line 198.*

Line 222: If reporting all other concentrations in mg/L, consider changing the g/L to mg/L.

> *Thank you, we have changed the values and units in lines 244 and 245.*

Line 345: I do not understand why SSC measurements from satellites from 2019-2021 were combined with flow data from 1958-1989, please clarify. This assumes seasonal SSC is the same from 1958-2021. Was there a discharge data limitation, why not use flow data 2019-2021? If there is a data limitation and this calculation is just for proof of concept, just state that.

> *Yes, the discharge data was only available for the 1958-1989 period, and is unfortunately not available for the current period. We have clarified this in the text (lines 382-384). It has to be noted that the past streamflow data was only used to give rough estimates of potential suspended sediment yields in the river in the past, and to compare with other studies. Yes this would be true under the assumption of the same interannual and long-term variability in streamflow in the past as today, as the reviewer correctly points out. As the reviewers may know, the Vjosa River was being considered as a site for hydropower system construction in Albania, and the issue of the high sediment loads which would be disrupted was a central part of the*

successful argument to stop this construction. This was the motivation to add in the paper a sediment yield estimate from the RS analysis.

Reviewer 3:
General Comments

This paper characterizes the concentration and change of turbidity profiles and fluxes using in situ field measurements and built relationships to Sentinel-2 images supplemented by two lagrangian field studies on the Vjosa River. In general, this is an interesting combination of methods with promise when including point and longitudinal measurements as well as remote sensing to better understand river water quality dynamics for an entire main stem river. I believe that the article is well written and overall clear, but I have some specific comments and questions for further clarification going forward.

Specific Comments

The use of the Lagrangian profile is a nice addition, but I would appreciate more in both the Introduction and Discussion concerning this kind of sediment assessment. For example, is your use of a Lagrangian profile for sediment validation novel? How do others use or incorporate this kind of data?

> *Per the suggestion of reviewer 2, we have changed the title of the article since we do not give much attention to the Lagrangian river profile, except in section 3.2. Therefore, we will not dedicate space in the Introduction or Discussion regarding the profile, however we have added some sentences in section 3.2 about how others have used this kind of data in the past (lines 283-284). As explained above, the kayak-based profile was only used as a proof of concept of variability in SSC along the river system, because we cannot compare this data easily with the snapshot of the river by satellite at a given instance in time.*

The introduction should be more explicit on the connection between riverine sediment inputs and coastal/ocean nutrient processing and productivity. For this kind of study, which focuses on a mostly free-flowing river with direct inputs to the Adriatic sea, this should be a major reason why understanding sediment processes are important.

> *Thank you for this suggestion. We have added a few sentences in the introduction pointing to the importance of nutrient delivery to the coast/oceans (lines 35-39).*

Line 181: Have you considered or tested certain band ratios in your OLS regression? Many studies test different combinations as well and it is unclear if you do the same.

> *Yes, we have done the same. We have added a line for clarification (line 211-212). Please also see the response to Reviewer 1.*

Line 342: I am confused on the number of water grab samples, their locations, and field/processing methodology for these samples over the field campaigns. It seems like you used them for the Dorez section, but the samples are from the entire main stem? Do you have any sediment-turbidity relationships to validate other sections and your source/sink conclusions on other sections of the river?

> *We created a sediment concentration-turbidity relationship from samples taken across the entire main stem. We did this because we want to create one relationship (sentinel_bands-turbidity) that can be applied to the entire main stem (as opposed to different relationships applied to different reaches, as explained in Section 3.1). When estimating the yield at Dorez, we similarly applied this catchment-wide equation of turbidity-SSC. This relationship can be applied to any section of the river but of course, a relationship with turbidity and SSC taken at only Dorez would be more accurate (however, we have neither turbidity nor SSC in this reach). We have clarified the processing methodology and from where the samples were taken in lines 372-379 in the revised manuscript.*

Line 354: While there are comparisons between total sediment yield on the Vjosa to other major rivers and past work, it would be helpful to compare your results to other more similarly sized mountainous catchment rivers (if mention this if it doesn't exist).

> *Thank you for the comment. We have included a comparison with some similarly sized and nearby rivers (lines 394-396).*

Line 357: I appreciate the contextual information about a potential Vjosa National Park and its implications (no hydropower; uninterrupted sediment flow) for this study.

> *Thank you.*

Technical Corrections

Figure 1: Add a secondary line indicating the kayak profiled reaches of the river and the long-term discharge measurement location

> *The long-term discharge measurement location, Dorez, is marked in the figure. We have added arrows indicating the start and end of the kayak profiled reaches. Thank you for the suggestion.*

Table 1: Consider adding a figure plotting NTU vs Rrs reflectance with the regression equation to help contextualize the equation and fit (how does it perform at low and high turbidity values). Also add the number of data points used to fit each equation.

> *We have added the number of data points used to fit each equation (line 220 and Table 2). And we have added a figure to the Supplementary Material showing the 3D regression plot (we refer to the figure in line 224).*

Figure 3: The map text and legend are too small, consider increasing the font size for easier interpretation.

Thank you. We have significantly increased the font text size.

Line 165: What are the average number of pixels, or was there a threshold used to calculate means at each ROI, reach, and tributary? If these are narrower river reaches, how many pixels constitute deep water reaches? This can help determine whether to use the mean or median band values. Since this is a smaller scale study and you can visually examine each image for cloud/shadow contamination the mean should work. Other studies use the median.

The average number of pixels within each ROI for every image was 280 pixels or 28,000m^2 (after removing the clouds, cloud-shadows, and non-deepwater pixels). After taking the average of the pixels within the ROI, we could obtain an average Rrs for each band (at each location, on every acquisition day). The average area and number of pixels used was added in lines 196.

Figure 7: Do you have any theories for the large and opposite trends to the rest of your data in the Fall 2020 lagrangian profile in the upstream portion of the reach? It is also unclear if highest boxplots are reaches or tributaries.

Yes, we theorize that the river is too narrow in the upstream section to extract proper reflectance data. Also, the river is much clearer upstream so our measurements may fail (added explanation at lines 326-328). And have also added arrows in the figure to make it clear whether the last two boxplots are tributaries or reaches.