

Dear Editor and Reviewers,

Thank you for the helpful review of our manuscript. We have completed the revision, per your suggestions. Our revisions are summarized below, with reviewer text in standard font and our responses in bold font. We appreciate your time and look forward to your further evaluation.

Sincerely,
James Stegen (on behalf of all co-authors)

###

Reviewer 1:

General comments:

Laan et al. submitted a well-written manuscript with a clear structure. The key strengths of this work are that Laan et al. could reliably measure ecosystem respiration in the water column (ERWC) at a broad spatial scale in environmentally diverse rivers, which are located at different positions in the Yakima River network and that they measured a comprehensive set of explanatory variables for ERWC ranging from basic physicochemical measures to high-resolution DOM data.

Thank you for the encouraging remarks.

Laan et al. hypothesized that ERWC values increase from upstream to downstream (approximated by drainage area) and conclude from their results that rather local factors are driving ERWC in river networks. Although I, in general, agree with the conclusion, I do not think that the data analysis is robust enough to really rule out the effect of drainage area (and regional processes in general) on ERWC because of two potential problems:

1. Laan et al. use LASSO regressions for variable selection and regularization. LASSO analysis does not provide p-values per se, however, one could argue that only important variables are selected in the model. Total drainage area was selected in the final model with a low positive mean beta coefficient suggesting a slight increase of ERWC values (i.e. rates slow down) from upstream to downstream. However, based on Fig. 3b, I see a decreasing trend of ERWC (i.e. rates fasten) from upstream to downstream. I suspect that this discrepancy is due to strong effects of other variables on ERWC, which also strongly correlate with drainage area (Fig. S4). I suggest to try partial regression plots to illustrate the net effects of predictors on ERWC. For Fig. 3b this could mean that you show how ERWC changed as a function of the drainage area after statistically holding the effect of other important predictors (e.g. TDN) constant. This strong collinearity among predictors (Fig. S4), however, may have even more influence on the analysis: If there are highly correlated predictors, LASSO's variable selection might only select one of them. I suggest to provide an explanation why you do not think that this is an issue for your analysis or a

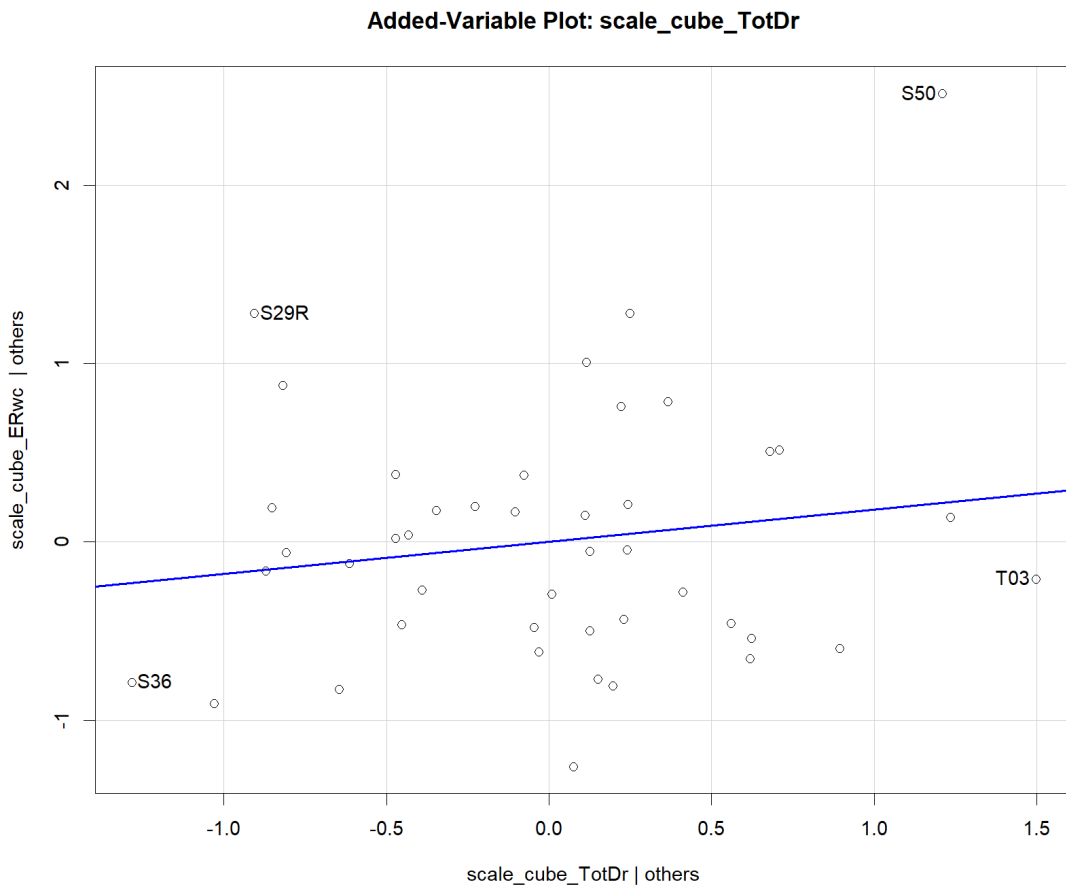
comparison of your current method with methods using elastic net penalties (combination of penalties of the LASSO and ridge method) which may deal better with multicollinearity.

In the revised manuscript, we addressed the high-level point of collinearity influences on parameters chosen in the LASSO regression in the second paragraph of section 3.3. We emphasized relationships between LASSO-identified variables and correlated ion concentrations that seem to indicate agricultural influences. We retained our approach of emphasizing N, rather than TDN or NO_3 , in paragraph 3 of section 3.3 to avoid assumptions about which of those variables is most important. We also refrained from placing an order of importance to the variables to maintain a focus on each as explaining ER_{wc} across the Yakima River basin.

To address the more specific point about total drainage area, we made the following five observations and plan to take actions summarized in each bullet.

- 1) To help minimize spurious outcomes that could arise due to collinearity across parameters, we used LASSO regressions across 100 random seeds, averaging the model coefficients. This reduced the likelihood of arbitrarily selecting one variable over another, which could happen with high collinearity. This revealed relatively small standard deviations of β coefficients compared to mean β coefficient values, indicating that the four most important variables are consistent across seeds, even when one variable is chosen over another. In the revised manuscript, we added more text explaining this comparison between seeds in section 2.7.
- 2) In preliminary analyses, not shown in the paper, we tested the sensitivity of our LASSO results to collinearity by removing highly correlated explanatory variables. In this case, we had to select between co-correlated variables and include one of them in the LASSO. We selected the variable to keep based on it having the higher (univariate) Pearson correlation coefficient with ER_{wc} . This did not change identities of the most important explanatory variables, as identified via LASSO. This consistency suggests that collinearity did not meaningfully alter the key findings from the LASSO modeling. These analyses are not included in the revised manuscript.
- 3) The variables identified as the most important via LASSO also had the largest Pearson correlation coefficients with ER_{wc} . This consistency between multivariate and univariate methods provides additional confidence in the robustness of the LASSO outcomes. This is already discussed in the first paragraph of section 3.3 of the paper and we retained that text.
- 4) A partial regression based on total drainage area, controlling for all the variables that had non-zero β coefficients in the LASSO model, is shown below. The relationship between total drainage area and ER_{wc} remains weak in this partial

regression plot. Additionally, the direction of the relationship is opposite from the univariate correlation; ER_{wc} was faster moving downstream in the univariate analysis but slower moving downstream in the LASSO model. We do not plan to add the partial regression plot to the manuscript, but added more discussion of how the relationship between total drainage area and other explanatory variables may relate to ER_{wc} as described below.



- 5) The highest Pearson correlation of total drainage area was $r = 0.63$ with temperature, followed by an $r = 0.58$ correlation with TDN. Temperature and TDN are indicated by the LASSO model as being important for explaining variation in ER_{wc} . Clear mechanistic connections between ER_{wc} and both temperature and TDN suggest that any relationship between drainage area and ER_{wc} is not causal. Instead, we infer that the relationship between total drainage area and ER_{wc} is simply because of correlations among the explanatory variables, and not a

mechanistic connection. This means that drainage area itself is not the key factor, but instead may act as a proxy for influences of temperature and TDN. Further, TDN is strongly correlated with other explanatory variables, including NO_3 , Cl, and SO_4 . The covariation among these variables likely reflects increasing influences of agricultural inputs that, in turn, lead to faster ER_{wc} by supporting microbial metabolism. We collectively infer that increasing temperature and nutrients, potentially from agricultural inputs, are the most likely drivers of ER_{wc} rates in the Yakima River basin. In the revised manuscript, we added more discussion of these points, including addressing the sign change between analyses, in the second paragraph of section 3.3.

2a. The model does not account for spatial autocorrelation, which could further lead to an overseen stronger shift of ERWC from upstream to downstream. Either you directly use spatially explicit methods for your analysis, such as spatial stream network models (e.g. Peterson, E. E. & Ver Hoef, J. M. A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology* 91, 644–651 (2010)) or you discuss this issue more thoroughly. For example, when looking at TDN along the drainage area (Fig. S4), I assume that TDN concentrations increase from upstream to downstream along two major branches of the network, which finally merge into the most downstream sites, suggesting a strong spatial autocorrelation of TDN concentrations. Hence, regional processes (transport of water chemical from upstream to downstream) might have an indirect effect on ERWC by shaping the spatial TDN pattern across the entire river network.

We added more discussion on the possibility of regional watershed processes influencing key explanatory variables such as total dissolved nitrogen to the second paragraph of section 3.3. We find this very helpful and we certainly agree that the chemistry observed at a given site is the outcome of processes occurring throughout the upstream drainage area (watershed).

2b. Further, I suggest to change negative ERWC values to positive values (consumption of O_2), as I think it makes the interpretation of positive and negative beta coefficients in the LASSO analysis easier and it is more common in the literature. This would also solve ambiguity when writing “ERWC increases”, because in these cases I was often not sure if the authors mean whether ERWC rates increase (i.e. rates fasten) or ERWC values increase (i.e. rates slow down). Overall, considering that my major concerns and specific comments will be addressed in a major revision, this study warrants publication in Biogeosciences.

To clarify interpretation in the text, we removed the occurrence of “rates increase” in the abstract, and replaced the occurrences of “rates increase” with revised language in the last paragraph of section 1 to explicitly state that we are referring to rate magnitudes (i.e., faster or slower). All language was changed to ‘faster or slower’ to avoid ambiguity. Slightly positive ER_{wc} values likely stem from very slow oxygen consumption and are

associated with instrument noise. However, at the advice of the reviewer, we changed these slightly positive values to 0 and redid analyses. This did not change the conceptual interpretation of the story. Section 2.6 in the methods and analyses throughout the results and discussion have been updated to reflect this. We retained the negative sign of the values, where more negative values indicate faster oxygen consumption. The main reason to retain this approach is to enable easy comparison to our original results, which are now in the supplementary information. Those original results include the slightly positive ER_{wc} rates such that we can't simply change the sign of the rates.

Specific comments:

26: Reading about positive ER values in the Abstract is very confusing. After reading the method section I understand that you kept positive values below 0.5 because they are difficult to distinguish from zero. Still, I wonder if you could avoid stating positive ER values in the Abstract to avoid confusion by the readers early on. But also see my specific comment to line 292.

We improved clarity of results in the Abstract by changing this sentence to "... 0 to -7.38..."

26: Although you clearly state that you "did not test this directly", I think it is still too much to say "that the contribution of ERWC rates to reach-scale ER_{tot} rates across the Yakima River basin are likely highly variable", because a comparison of a set of ERWC values with another set of ER_{tot} values, which both derive from a wide range of streams differing in size, stream order, environment, etc. does not give you that information. You might come closer to such a result if you apply the same classification to the streams from the literature as you have done to your study streams and then compare ERWC and ER_{tot} within classes.

To clarify that this comparison was exploratory and intended to assess the potential overlap between our measured ER_{wc} values and ER_{tot} reported in the literature, we added a sentence describing these limitations in the third paragraph of section 3.1. In this section, we also addressed that the overlap in ER_{wc} measured in the Yakima River basin and literature ER_{tot} suggests that ER_{wc} could contribute a variable amount to ER_{tot}, which may be substantial in some cases, though this is not guaranteed.

28: I suggest to keep the statement that you "did not observe a clear increase in ERWC" more neutral. Because at that point of the manuscript, it is not clear why you expect an increase in absolute ERWC values from upstream to downstream. But see also my next specific comment.

We changed this sentence to: We observed that ER_{wc} is locally controlled by temperature, dissolved organic carbon, total dissolved nitrogen, and total suspended solids, which explained 40% of ER_{wc} variability across the basin using Least Absolute Shrinkage and Selection Operator (LASSO) regression.

71: I miss an explanation why you expect absolute ERWC values to increase. At the beginning

of this paragraph you state that water column processes are expected to become increasingly important from upstream to downstream due to a shift from benthic-dominated processes to water column dominated processes. This statement is about the relative contribution of water column processes. However, I do not think that this easily translates to your hypothesis that absolute ER_{WC} values will increase from upstream to downstream, because ER_{tot} is expected to decrease (rates slow down) from upstream to downstream (e.g. Segatto, P.L., Battin, T.J. & Bertuzzo, E. (2021). The Metabolic Regimes at the Scale of an Entire Stream Network Unveiled Through Sensor Data and Machine Learning. *Ecosystems*, 24, 1792–1809.).

We added the following additional material to the third paragraph of the Introduction to provide more rationale for the hypothesis that ER_{WC} increases moving downstream:

“Increases in downstream GPP (Finlay, 2011; Segatto et al., 2021), may also influence ecosystem respiration, such that we would expect faster ER_{WC} with greater GPP due to increases in autochthonous C (Hall et al., 2016; Mejia et al., 2019). Additionally, greater N processing in the water column with increasing stream order (Wang et al., 2022), may suggest that water column biogeochemical processing increases along the stream network.”

Finlay, J. C. (2011). Stream size and human influences on ecosystem production in river networks. *Ecosphere*, 2(8), art87. <https://doi.org/10.1890/es11-00071.1>

Segatto, P.L., Battin, T.J., and Bertuzzo, E. (2021). The Metabolic Regimes at the Scale of an Entire Stream Network Unveiled Through Sensor Data and Machine Learning. *Ecosystems*, 24, 1792 - 1809.

Wang, J., Xia, X., Liu, S., Zhang, S., Zhang, L., Jiang, C., Zhang, Z., Xin, Y., Chen, X., Huang, J., Bao, J., McDowell, W.H., Michalski, G., Yang, Z., and Xia, J. (2022). The Dominant Role of the Water Column in Nitrogen Removal and N₂O Emissions in Large Rivers. *Geophysical Research Letters*, 49 (12).

148: Please write out the abbreviation “DO” first time you mention it.

We added “dissolved oxygen” at the first use of DO in the first paragraph of section 2.3.

200: Can you please clarify how calculating “the distance between each of the replicate samples” helps to identify the outlier.

We added the following information in the second paragraph of section 2.4: “Pairwise differences between NPOC, TDN, and DIC measurements from all replicates were calculated. The sample that had the largest difference from the other samples was removed if the coefficient of variation was greater than 30%. This coefficient of variation

threshold for sample removal is based on inspecting histograms of these data types, and determining the point at which sites likely contain anomalous outlier values.”

255: perform instead of performs

This has been updated.

292: I wonder, whether keeping positive ERWC values below 0.5 in your data biases your results (Fig. 2) towards low respiration rates. Maybe this is the reason why you find on average lower respiration rates compared to studies from the literature. Since you are comparing your data to the literature, it might be reasonable to investigate how studies from literature were dealing with positive ERWC values and subsequently use the same approach. I definitely do not suggest to erase cases with positive ERWC values below 0.5 from the results but I find it more intuitive to change positive ERWC values below 0.5 to zeros, as you rightfully stated that positive values are biologically unrealistic and are not distinguishable from zero. However, if studies from the literature research also keep positive values, I would suggest to keep the current approach to make results comparable.

To the best of our knowledge, other published studies of ER_{wc} do not discuss positive rates, so there is no precedent to base our decision off of. However, we have changed the slightly positive values measured across the Yakima River basin to a value of zero at the reviewer’s suggestion. This did not change the conceptual interpretation of the analysis. This has been updated in section 2.6 in the Methods, Figure 2, Figure 3, Figure 4, and Table 2.

342: I understood that you use LASSO analyses to investigate which variables are selected for your final model. Hence, a rejection of your hypothesis at this point of the manuscript based on the correlation in Fig. 3b seems premature. See also my general comment to this topic

We updated this language in the first paragraph of section 3.2 to include that the relationship “...between ER_{wc} and total drainage area across the Yakima River basin was weak enough that we consider it inconsistent with our hypothesis...” and added text clarifying that multivariate analysis is used in subsequent sections for further evaluation.

347: This part of the sentence is not clear: “... as opposed to a higher order river“, please clarify.

We edited this sentence to clarify that the results of the fastest observed ER_{wc} occurring in a 5th order stream was “... opposed to our hypothesis of ER_{wc} being fastest in the highest stream orders.”

364: Change to “Regression analyses showed ...”

This has been updated in the revised manuscript.

Technical corrections

88: Remove space previous to dot.

This has been updated in the revised manuscript.

216: The reference is missing at the end of the sentence.

This has been updated in the revised manuscript.

Reviewer 2:

Laan et al. evaluated water column respiration (ER_{wc}) and water quality parameters at 47 sites in the Yakima River Basin. The goal of the study was to identify factors driving changes in ER_{wc} throughout the river network using LASSO regressions. In addition, the authors collected total and ER_{wc} values from other studies in the continental US and the Amazon River basin. In general, the authors found no clear increase in ER_{wc} over the course of the river network, and ER_{wc} rates were influenced by local factors such as temperature, dissolved organic carbon, total dissolved nitrogen, and suspended sediment, rather than position in the stream network. In addition, the range of ER_{wc} in the Yakima River Basin encompassed the entire range of ER_{wc} that the authors found in the other studies, and ER_{wc} contributed differentially to ER_{tot} from the other studies.

Thank you for the encouraging remarks.

This study is well-focused and addresses a question regarding the processes occurring in the water column of river networks. The research is thorough and directed, leaving little room for criticism from my perspective. I'll leave two comments here that I would have liked the authors to address in a little more detail to see if/how relevant this might be to their study.

One point I am thinking about is the discussion of the importance of water column and sediment processes to overall metabolism. In line 22 of the abstract and several times throughout the manuscript, the authors state that "the relative influence of sediment-associated processes versus water column processes can fluctuate along the river continuum." In my opinion, an important factor in this statement is the greater influence of water column processes due to higher water levels when going downstream, which increases the areal influence of the water column. However, the authors compare volumetric rates, which do not consider the influence of water column height. Why did the authors decide to compare volumetric values? I'm not criticizing the approach, but I think the theory they are testing is largely based on this relationship. This could be a point that could (or should?) be included in the discussion.

One reason we used volumetric rates for analysis is the difficulties in obtaining high quality depth data at all sites across the Yakima River basin. This study includes 7th order main stem sites, where it would be challenging to measure depth. This has been clarified in section 2.6 of the Methods. In addition, some literature estimates of ER_{wc} are in volumetric units and no depth data are provided; the only way to do a direct comparison across all literature rates is via volumetric units. Nonetheless, we acknowledge the importance of considering water column depth for understanding variation in the contribution of water column processes to whole system respiration (i.e., ER_{tot}) in the first paragraph of section 3.2. We also encourage future studies to measure water column depth to obtain areal ER_{wc} rates in the Conclusion.

The authors state in line 395 that "Nitrogen is a key nutrient for microbial growth and is often a limiting nutrient in freshwater rivers (Carroll, 2022)." Another common limiting factor is phosphorus. The authors use a variety of water and catchment parameters to perform the regression. However, phosphorus was not examined. Is there a reason for this? Is this not a potential important factor for ecosystem metabolism in the Yakima River Basin? Including this factor could improve the significance of the regression and significantly influence the conclusion that 40% can be predicted.

We agree that phosphorus is often limiting and probably is important in the Yakima River Basin. Our analysis of phosphorus showed values below detection for more than two-thirds of samples, which is further evidence that it probably is limiting. Because of the analytical limitations, we feel there isn't enough good phosphorus data to include in the analyses. This is unfortunate of course. To address this in the revised manuscript, we acknowledged that phosphorus is likely a key nutrient in the second paragraph of section 3.3 and that while we attempted to measure it, we didn't get data of sufficient quality to include them in the analysis (second paragraph of section 2.4).

Minor comments:

Line 30: "...which explained 40% of ER_{wc} variability across the basin." You could add here the method you used to come to this number as you use LASSO regression, which has certain assumptions.

We described LASSO regression as the method used for regression analyses in the Abstract per the reviewers suggestion. We also added the following description of how R^2 was calculated in section 2.7:

"LASSO does not inherently estimate R^2 , so we calculated it using the total sum of squares and residual sum of squares for each fitted model, as traditionally done with standard multiple regression. The estimation of residual sum of squares used predicted values of ER_{wc} based on the explanatory variables used in the model. The R^2 estimates were used to estimate how much variation in ER_{wc} was explained by each of the LASSO models. Standard deviation of β coefficients were compared to mean values of β

coefficients to confirm that the most important variables were relatively consistent across seeds.”.

Line 216: Reference missing

The missing reference has been added.

We will include the reference in the revised manuscript.

Line 390: Could not find Ochs et al. 2010 in the reference list

The missing reference has been added.