

Dear Editor Charles Onyutha,

We sincerely appreciate your constructive comments on our figures. In response, we have thoroughly revised all figures to improve clarity and readability. A brief summary of the revisions is provided below, and detailed, figure-specific responses are presented in the subsequent text.

- For all figures containing spatial maps: we have added scale bars, north arrows, grids, graticules, and legends.
- Figure 1: The colour scheme has been revised to avoid red–green contrast.
- Figure 2: To improve readability, only the map of catchment locations is retained in Figure 2. The remaining spatial maps have been moved to the supplementary materials (Figure S1).
- Figure 3: The layout has been reorganized from a 1 × 3 landscape format to a 2 × 2 arrangement, significantly increasing the size of each subplot. In addition, Table 3 has been added to present multiple evaluation metrics for each catchment.
- Figure 4: Only two representative catchments (W5 and W10) are retained in this figure. Results for all catchments have been moved to the supplementary materials and separated into three figures (Figures S2–S4) to enhance size and readability. A summary table of evaluation metrics for each catchment has also been added (Table 3).
- Figure 5: Similar revisions were applied as in Figure 4. Only W5 and W10 are retained, while results for all catchments are presented in the supplementary materials (Figures S5–S7).
- Figure 6: This figure has been removed as suggested with the key information presented in a new table (Table 4).
- Figure 7: Similar revisions were applied as in Figures 5 and 6, resulting in a more concise Figure 6 and an additional supplementary figure (Figure S8).
- The revision of figures does not change the conclusions of this manuscript.

We hope that these revisions satisfactorily address your comments and improve the clarity and quality of the manuscript.

Best regards,

Songjun Wu

on behalf of all co-authors

Below are suggestions on how the authors should improve the presentations of their results.

Figure 2

Some catchments are too small to appear in Figure 2 given its current size. Re-organize the various components and increase their sizes to ensure that the overall Figure 2 is at least 3/4 of a page with the selected catchments distinctively visible.

** We agree that extending the European map can help distinguish the selected catchments. Therefore, a few revisions were made as suggested:

- Only the main map was retained in **Figure 2** with larger size. We also adjusted the colours of catchment W2, 4, and 17 due to their poor visibility in colourblind mode. The text size was adjusted accordingly.
- All subplots were moved and reconstructed as new **Figure S1**. The color theme was changed to “Viridis” for clear visibility.

Figure 3

Here, I found it difficult to relate the information from the text with the results presented in Figure 3. I noted that there are many sites used to perform calibration. However, the information on the calibration is not clear. Simply present a table comprising information on model quality e.g. KGE, R-squared, Pbias and root mean squared error for each site. Within the text, describe and discuss the spatial variation in the model performance with reference to the said table that you have to create.

** Thank you for your valuable comments. We agree that a table listing performance metrics for each catchment provides a clearer and comprehensive overview of model calibration and validation. Accordingly, we have added **Table 2**, which presents catchment-specific performance for discharge, in-stream isotope, and nitrate, summarized using KGE, RMSE, Pearson correlation coefficient, and percent bias.

Meanwhile, we respectfully argue that a spatial representation of model performance is useful for identifying where the model performs well or poorly. Therefore, the original Figure 2 has been retained but substantially revised as follows:

- The previous 1 × 3 landscape layout resulted in small subplot sizes. We have therefore reorganized the figure into a 2 × 2 layout to improve readability and visibility.
- Scale bars, compass, grids, and graticules have been added to Panels A–C. The colour scheme has also been changed to “Viridis” to enhance visual clarity.
- A scatter plot has been added as Panel D to provide an overall comparison of all simulated and observed values.

These revisions improve both the clarity and interpretability of the figure while maintaining its value as a spatial overview of model performance.

Figure 4

Because the figures are not clear, it is difficult for one to verify some information in the text. For instance, the authors mention that “The results in Figure 4 shows a general good fit between simulation and MODIS records with KGE and r2 above 0.5 in most regions”. To verify this claim, one has to zoom-in the tiny components of Figure 4 to check the KGE and r2. Unfortunately, the poor resolution of the figure leads to blurriness during to the zooming-in process.

Therefore, pick two sizable catchments e.g. W5 and W10 from among those selected for analysis and show how the model performance varies in space. Then in one table, summarize the key information

such as KGE, R2, root mean squared error, and Pbias for each catchment. Accordingly, revise the relevant text on Figure 4 based on the revised figure and newly created table.

** Thank you for the comments. We agree that placing three continental maps into one figure is difficult for readers to identify the detailed spatial pattern. Therefore, three revisions were implemented as suggested.

- Only two representative catchments (W5 and W10) were shown in **Figure 4**. Scale, compass, grids, graticules, and legends were also added.
- A new **Table 3** was added to show the evaluation metrics in each catchment.
- We respectfully argue that the continental maps can still provide spatial information at boarder scale. Therefore, they were moved to **Figure S2-4** with each individual map showing the validation map of a specific variable (evapotranspiration, water storage anomaly, and snow depth) to ensure the readability.

Figure 5

The authors can select just two catchments, say, W5 and W10 to demonstrate the spatial variation of the results. This means that for Figure 5 we need a total of six sizable sub-plots of good maps each having scale, compass, grids and graticules (or longitudes and latitudes), and legend. For clarity, the size of Figure 5 could be at least 75% of a page. The various sub-plots should include (a) stream water age (days) for W5, (b) stream water age (days) for W10, (c) soil water age (days) for W5, (d) soil water age (days) for W10, (e) travel time in soils (days) for W5, and (f) travel time in soils (days) for W10.

Key information on the other catchments apart from W5 and W10 can be highlighted in the text. Further details of the results for the rest of the catchments can be appended or put in the supplementary material.

** Thank you for the comments. We agree that presenting the representative catchment will increase the readability of the figure. Therefore, Similar revisions as Figure 4 were implemented as follows:

- Only two catchments (W5 and W10) were shown in **Figure 5**. Scale, compass, grids, graticules, and legend were also added.
- The continental maps were separated presented in **Figure S5-7** to maximum visibility of the spatial pattern at continental scale.

Similar revisions were adopted for Figure 7, which resulted in a more concise **Figure 6** showing spatial maps of two catchments, and **Figure S8** showing the continental map of Damköhler number.

Figure 6

The current Figure 6 should be replaced with a table of results for the regression analysis. The authors can, for each selected catchment, present the regression slope and intercept as well as (R-squared instead of the correlation coefficient), and p-value.

** We agree that presenting the results in a table can more clearly and explicitly show the relationship between soil travel time and denitrification. Accordingly, we have replaced the original Figure 6 with a new table (**Table 4**). This table reports the mean soil travel time and denitrification over the period

1982–2024 in each selected catchment, together with the slope of the fitted relationship and the p-values of the Spearman test.

We did not include the coefficient of determination (R square) because a linear relationship between soil travel time and denitrification is not necessarily expected. Soil denitrification is controlled not only by travel time but also by fertilization, plant uptake, etc. Therefore, R^2 yields very low values (ranging from -800 to $-2e11$), which are not meaningful for interpretation in this context. Such situation differs from model performance evaluation, where a strong agreement between simulated and observed values is expected and R square serves as an appropriate metric.

Finally,

The authors need to ensure that the figures' color schemes permit readers with color vision deficiencies to correctly interpret the research findings. See guidelines for figures & tables via <https://www.geoscientific-model-development.net/submission.html>. Where necessary, the authors should use the color blindness simulator (check relevant information via <https://www.color-blindness.com/coblis-color-blindness-simulator/>).

** Thank you for the suggestions. To ensure the readability, we have changed the colour codes of all figures to “Viridis” and “Red-blue”, the widely used colour themes that are colourblind friendly due to their high contrasts. We have checked all figures using the colour blindness simulator.