Note: The comments from the reviewer are in black while the responses from the authors are in blue.

Reviewer 1

The manuscript presents the new CoSWAT-WQ model, which enhances our understanding of global water quality problems related to nutrients, particularly through its high spatial and temporal resolution. This represents an important step in the field of water quality research and paves the way for future model developments. I read the manuscript with great interest. It is generally well-written and presents a new global water quality model with a remarkably high resolution. This model, indeed, holds potential for scientists, policymakers and other stakeholders in the field of water quality. However, currently, the limited methodological detail makes it challenging to fully assess the scientific approach and methods applied. Moreover, the manuscript could highlight the novel insights gained from the model results rather than solely focusing on model evaluation. Improvements in the Methods, Results, and Discussion sections are needed before considering this manuscript for publication. I have explained my concerns in the comments below.

Response: We sincerely thank the reviewer for their thoughtful and encouraging comments, as well as their interest in our work. We appreciate the recognition of the potential and significance of the CoSWAT-WQ model for advancing global water quality research. We acknowledge the concerns raised and will incorporate all suggestions to improve the quality and clarity of the manuscript.

1. One of the main novelties of the new CoSWAT-WQ model is its high spatial and temporal resolution at the global scale. Yet, this also comes with some concerns:

    - The authors present the model as a daily time step model. Yet, the discussion includes the following: "Given the approximations inherent in model structure, uncertainties in input data, and the complexity of nutrient transport and transformation dynamics, the primary objective of global water quality models is not to predict exact daily concentrations (UNEP, 2016). Instead, they aim to identify major spatial and temporal pollution hotspots within global river networks. For this purpose, the CoSWAT-WQ model performs adequately in comparison to other global nutrient models for river nutrient loads but with room for improvement in comparison to the concentration observation data". I would suggest the authors clarify their justification for a daily time-step model and the implications of this choice, especially at the global scale. This particularly holds for tropical and subtropical regions where scheduling using heat units may result in incorrect cropping seasons, as noted in Nkwasa et al., (2022). This raises questions such as "how well does the model capture daily patterns?" and "how should the reader interpret the model outputs?" Another option is to revise the aim of the study and clearly link the purpose of the model and how this fits the daily time step.

Response: Thank you for raising this point. We recognize that the current description may have caused a lack of clarity regarding the model's temporal focus and subsequently, the aim of the study. Although the CoSWAT-WQ is run at a daily timestep, our evaluation has most been done at monthly (with observations) and annual timestep (with other global models). Thus, recommending the current model version for identifying major spatial and temporal pollution hotspots rather than predicting exact daily concentrations. We recognize that this position may appear contradictory to statements in the abstract and conclusion about the model's daily simulation capability. To clarify, we will revise the manuscript to explicitly state that although the model runs on a daily timestep, its current validated use is for broader temporal patterns monthly and annual and spatial hotspot identification. At the same time, we emphasize that the model's daily timestep structure offers significant potential for future improvements, with subsequent versions leveraging daily simulations to provide higher-resolution estimates supported by high-resolution validation.

Regarding cropping season scheduling and concerns about heat unit-based approaches Nkwasa et al. (2022), we actually follow the workflows from Nkwasa et al. (2022) to implement management practices, using global cropping calendars based on actual planting and harvest dates rather than relying on heat units. This approach helps ensure more accurate representation of cropping seasons, particularly in tropical and subtropical regions.

We will revise the manuscript to better align the study's aims with the model's current strengths in hotspot identification and highlight the potential of the daily timestep for future developments.

- I acknowledge that model evaluations at the global scale are challenging. Hence, I appreciate the authors efforts in evaluating the model by comparing the model results with the IMAGE-GNM model (Figures 4-5) and monitoring data from GEMStat (Figure 8). For the comparison with IMAGE-GNM, the authors compared simulations for the world's 30 largest rivers. To enhance the relevance of this comparison, I suggest including information such as the percentage of the global total drainage area that these rivers cover or their share related to total nutrient exports. In Figure 5, it is difficult to interpret the different symbols used (circle for IMAGE-GNM and triangle for CoSWAT-WQ). The use of symbols seems unnecessary as each model is included on a different axis. In addition, I wonder whether the co-authors have considered comparing their model results to other large-scale water quality models (e.g. mQM for Europe and SWAT+ for Africa). Regarding the comparison to monitoring data from GEMStat (the temporal patterns in Figure 8), it becomes clear that the model may underestimate extreme events and shows a slightly advancing pattern (e.g. earlier peaks) compared to the monitoring data. Next to GEMStat, there are other databases available. For example, Jones et al. (2024) compiled a comprehensive dataset of water quality monitoring data. Including this dataset in the evaluation can further strengthen the evaluation efforts.

Response: Thank your for your suggestions. We will include the information such as the percentage of the global total drainage area that these rivers cover or their share related to total nutrient exports as shown in Table R1. Regarding the symbols in Figure 5, we believe that their inclusion helps readers distinguish where each model's data points lie in relation to the 1:1 line. Although the models are plotted on separate axes, the symbols add an extra layer of clarity for comparative interpretation.

With respect to comparisons to other large-scale or regional models, such as SWAT+ for Africa (Nkwasa et al., 2024) and mQM for Europe (Kumar et al., 2020), we agree that these are valuable references. We will include a spatial comparison with these models in the supplementary material to complement our global evaluation.

As for the suggestion to incorporate additional observational datasets beyond GEMStat, we agree that a broader evaluation can enhance robustness. However, GEMStat remains the most extensive individual publicly available global database for water quality. For example, the average time series length per site for total phosphorus (TP) in GEMStat is 6.6 years, compared to 4.9 years in GLORICH. Moreover, many of the sites in GEMStat overlap spatially with other databases such as Waterbase, the Water Quality Portal (WQP), and the Canadian Environmental Sustainability Indicators (CESI), often resulting in duplicate records (Virro et al., 2021). Importantly, large regions in Africa, South America, and Asia remain underrepresented in alternative datasets (Jones et al., 2024), limiting their utility for a truly global-scale evaluation. Given these considerations, we consider GEMStat a sufficient and appropriate choice for the scope of this study.

We will revise the manuscript to enhance the inter-model comparison (including Table R1) and include an evaluation of our global model against regional models in Africa and Europe, both in the main text and the supplementary material.

Table R1: Model intercomparison of CoSWAT-WQ, IMAGE-DGNM (Beusen et al., 2015) and MARINA-Multi (Micella et al., 2024) models for the year 2010

River ID	River basin name	Area (km2)	mouth longitude	mouth latitude	TN (Tg/yr)			TP (Tg/yr)		
					IMAGE-DGNM	MARINA-Multi	CoSWAT-WQ	IMAGE-DGNM	MARINA-Multi	CoSWAT-WQ
1	Amazon	5846870	-51.75	-1.25	3.78	4.10	1.64	0.17	0.29	0.36
2	Nile	3821590	31.25	31.25	0.16	0.36	1.39	0.01	0.04	0.07
3	Congo	3694430	12.75	-5.75	0.58	1.27	0.83	0.00	0.11	0.14
4	Mississippi	3199170	-90.25	29.75	1.38	0.99	0.92	0.11	0.07	0.08
5	Ob	3022320	69.25	66.75	0.29	0.25	1.33	0.03	0.02	0.04
6	Parana	2660890	-58.75	-34.25	0.61	0.59	0.38	0.05	0.07	0.06
7	Yenisey	2575660	82.25	71.25	0.28	0.001	0.001	0.00001	0.02	0.002
8	Lena	2438900	127.25	73.25	0.24	0.001	0.001	0.0002	0.01	0.0001
9	Niger	2237360	6.75	4.75	0.38	0.39	0.10	0.004	0.10	0.05
10	Yangtze	1792120	121.75	31.25	1.13	1.92	1.22	0.01	0.14	0.27
11	Amur	1752600	140.75	53.25	0.42	0.19	1.77	0.03	0.02	0.17
12	Mackenzie	1692900	-134.75	69.25	0.17	0.10	1.63	0.00	0.01	0.00
13	Volga	1474650	48.25	46.25	0.18	0.19	0.91	0.03	0.01	0.14
14	Zambezi	1361960	36.25	-18.75	0.42	0.15	0.11	0.00	0.01	0.03
15	Indus	1141750	67.75	24.25	0.50	0.13	1.36	0.11	0.003	0.09
16	St Lawrence	1052470	-70.75	47.25	0.01	0.16	0.02	0.03	0.02	0.002
17	Orinoco	1038130	-61.25	8.75	0.30	0.95	0.42	0.03	0.09	0.08
18	Murray-Darling	1030290	139.25	-35.25	0.02	0.04	0.001	0.003	0.001	0.002
19	Shatt el Arab	988948	48.25	30.25	0.10	0.004	0.08	0.02	0.002	0.02
20	Yellow	892570	118.25	37.75	0.65	0.08	0.00	0.07	0.02	0.00
21	Yukon	854690	-164.75	62.75	0.04	0.04	0.23	0.01	0.004	0.009
22	Danube	787069	29.25	45.25	0.37	0.26	0.51	0.05	0.02	0.07
23	Mekong	757660	106.25	10.25	0.57	0.60	0.45	0.14	0.03	0.01
24	Columbia	731105	-123.75	46.25	0.14	0.09	0.96	0.01	0.01	0.16
25	Sao Francisco	614418	-36.75	-10.25	0.10	0.13	0.08	0.01	0.01	0.02
26	Pearl	408043	113.25	22.25	0.36	0.82	0.01	0.03	0.05	0.002
27	Irrawaddy	405481	95.25	15.75	0.78	0.43	0.58	0.16	0.04	0.09
28	Salween	273038	97.25	16.75	0.26	0.16	0.01	0.08	0.01	0.09
29	Magdalena	251445	-74.75	10.75	0.30	0.34	0.31	0.05	0.05	0.04
30	Rhine	164864	5.75	52.75	0.15	0.17	0.21	0.01	0.01	0.02

2. The Method section seems to provide only a limited description of several important aspects of the model. Below, I outline specific areas where further detail would improve clarity and reproducibility.

   - In Lines 114-116, the authors state that the timing of fertilizer and manure applications, as well as irrigation and biomass removal, can be scheduled based on calendar days or heat units, as described in Nkwasa et al. 2022. To fully understand the method, readers may need to consult previous SWAT model documentation. Given that this manuscript emphasizes the novelty of its spatial and temporal resolutions, it would be beneficial to include a brief overview of the downscaling procedures – either directly in the text or as part of a conceptual framework. For example, a brief description on how the timing of manure and fertilizer applications (which are annual datasets, as indicated in Table 1) are determined would enhance clarity. This is particularly important as the timing can be set by the user or is automatically applied by SWAT based on a specified nitrogen stress threshold (according to Neitsch et al., 2005). Additionally, it would be helpful to include a short description of how the 5-year point source input data from Beusen et al. (2022) were downscaled, as well as how the 0.5-degree resolution input data were spatially refined to 2 km resolutions (Table 1).

Response: Thank you your suggestions. We will add the details on how we apply fertilizer in the revised manuscript as this has also been raised in the Community Comment. To give some details here, consider an agricultural grid or hydrologic response unit (HRU), as shown in Figure R1. Above each of these grids, we have several management layers, including crop type, cropping calendar (plant and harvest days), and fertilizer or manure application. For each agricultural grid or HRU, we read the corresponding values from these management layers and apply them accordingly. This process is pre-processed using a Python script, which organizes the information into a decision table (an example is shown in Figure R2), following the workflow developed in Nkwasa et al. (2022). This allows us to automate the assignment of agricultural management practices across all agricultural grids or HRUs.

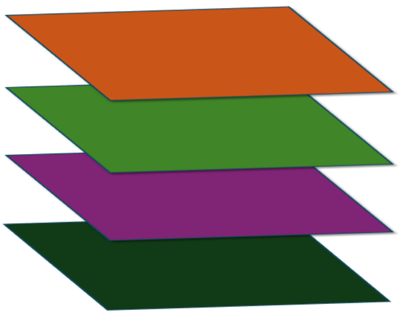

Fig. R1: Schematic representation of an agricultural grid or hydrologic response unit (HRU) and its associated management layers

Figure R2 presents a sample decision table covering planting, harvesting, and fertilizer application. It illustrates different nitrogen (N) and phosphorus (P) practices i.e. amounts, frequencies, and application method. The total annual fertilizer amounts come from existing global datasets; these annual totals are divided into three equal portions (application frequency) (Hu et al., 2021), that are broadcast (application type) after the planting date and before the harvest date, spaced across the cropping season. This preserves the correct yearly total but can still introduce uncertainty, because actual farm practices vary worldwide in timing, frequency, and method of application. Thus, we do not use the nutrient (N and P) stress threshold to trigger fertilization but rather apply the annual fertilizer totals into 3 equal portions across the cropping seasons.

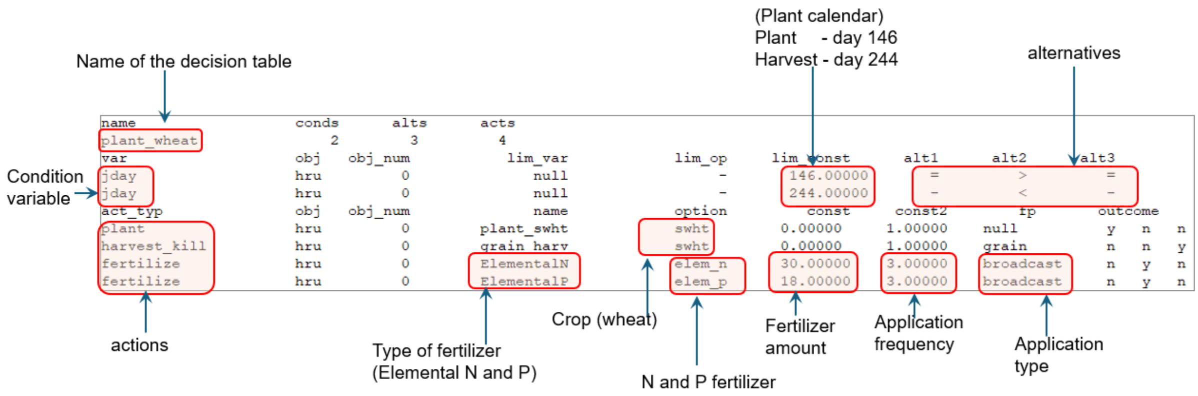

Fig. R2: Sample decision table showing the planting, harvesting and fertilizer application harvesting of a wheat crop

The question of downscaling the resolution of the datasets used for the management layers. In our study, we use freely globally available datasets, most of which are at a resolution of 0.5 degrees. Since our model operates at 2 km grid resolution, multiple 2 km grid cells fall within each 0.5-degree cell. This means that all agricultural grid cells within a single 0.5-degree grid share the same management practices. In summary, the degree of heterogeneity in management practices we can capture depends on the resolution of the agricultural input datasets. To our knowledge, there are currently no global management datasets available at a higher resolution than those we have used. However, one strength of our workflow is its flexibility, if higher-resolution data become available in the future, this approach can incorporate and reflect that added heterogeneity.

To downscale the 5-year point source input data from Beusen et al. (2022) to annual resolution, each 5-year data point was assumed to represent the average conditions for the surrounding 5-year period. Specifically, the value for a given year (e.g., 1990) was applied uniformly to the five-year span from 1990 to 1994. Similarly, the 1995 data point was used for the years 1995–1999, and so on. This step-wise temporal downscaling approach assumes constant point source inputs within each 5-year interval.

We will revise the manuscript to include more detailed descriptions of the management workflows in both the main text and supplementary materials, particularly in relation to fertilizer application and point source data integration.

- Section 2.1 refers to the N and P cycles and associated nutrient pools. However, it remains unclear how the model ensures mass balance and closure of these nutrient pools across space (e.g. line 117 refers to basins whereas the model runs on HRU scale) and time (e.g. negative balances).

Response: Thank you for your question. To clarify, in SWAT+, nutrient cycling and mass balance calculations are performed independently at the HRU level, where processes such as plant uptake, leaching, and mineralization are simulated daily. These nutrient fluxes from all HRUs within a subbasin are then aggregated by summing components like surface runoff nitrogen and leached nitrate and passed to the subbasin-level routing unit. Within the subbasin, SWAT+ models in-stream and lateral processes such as nutrient settling, decay, channel transport as nutrients move through the stream network. The resulting subbasin-scale nutrient balance reflects the cumulative contributions from all HRUs along with transformations occurring within the stream system. Mass balance is maintained by explicitly tracking all nutrient inputs, outputs, and transformations across both land and water components. Temporally, SWAT+ updates all nutrient pools on a daily time step and applies the mass balance framework: all inputs, outputs, and transformations are accounted for, ensuring that changes in storage reflect the difference between inflows and outflows. In SWAT+, negative nutrient pools are not directly simulated as physical entities. Instead, negative values in nutrient pool outputs often indicate model behavior where losses (e.g., through denitrification, leaching, or erosion) exceed inputs during a specific time step. A positive nutrient balance value means that the nutrient is being retained in the system and negative nutrient balance means that the nutrient inputs are lower than outputs during a given period, considering the outputs as the losses by all the possible pathways.

We will revise the manuscript to include more details about the nutrient balance and the simulations from HRU to basin scale.

- In contrast to the calibrated SWAT+ model, the new CoSWAT-WQ model is uncalibrated. Yet, it remains unclear which model parameters used to be calibrated and how the uncalibrated alternative works. I suggest the authors to include a table which specifies which model parameters are calibrated in SWAT+ and how the uncalibrated approach works.

Response: Thank you for your suggestion. The current version of the CoSWAT-WQ model, like most global water quality models such as DynQual v1.0 (Jones et al., 2023), MARINA (Micella et al., 2024), and IMAGE-GNM (Beusen et al., 2015) is uncalibrated. We used the default CoSWAT-WQ model parameters, which are based on literature and previous applications of the SWAT and SWAT+ models (Arnold et al., 2012; Arnold et al., 2013). As noted in the discussion (Line 270), future research will focus on model calibration and validation using observational data, especially from data-rich regions, to enhance the model's accuracy and reliability.

We will include the table of the default model parameters (Table R2) in the Supplementary material of the revised manuscript to show which parameters are usually/typically calibrated for hydrology and nutrients.

Table R2: Values of SWAT+ parameters that are typically subject to calibration for hydrology and nutrients

Parameter	Description	Common calibration range
	Flow	
CN2	Curve number (percent)	±10% of default value
ESCO	Soil evaporation compensation coefficient (replace)	0-1
AWC	Soil available water capacity (relative)	±0.04 of default value
perco	Percolation coefficient (replace)	0-1
Surlag	Surface runoff lag coefficient (relative)	0.05-24
Latq_co	Lateral flow coefficient (replace)	0-1
	Phosphorus	
P_updis	P uptake distribution parameter	0-100
Phoskd	Phosphorus soil partitioning coefficient	100-400
biomix	Biological mixing efficiency	0-1
psp	Phosphorus availability index	0.01-0.7
Lat_orgp	lateral organic phosphorus	0-200
	Nitrogen	
cdn	Denitrification exponential rate coefficient	0-3
cmn	Rate factor for humus mineralization	0.001-0.003
sdnco	Denitrification threshold water content	0-1
N_perco	Nitrate percolation coefficient	0-1
N_updis	N uptake distribution parameter	0-100

The Results section provides a comparison of the model results with other models and monitoring data. However, it lacks a presentation of novel insights from the newly developed model. For example, what new understanding do the high-spatial-temporal-resolution results offer regarding global and local water quality assessments? From my point of view, highlighting such novelties would strengthen the value of the study and fit the aim of the study.

Response: We thank the reviewer for the valuable suggestion. We agree that the manuscript would benefit from a clearer articulation of the novel insights provided by the CoSWAT-WQ model. In the revised manuscript, we will strengthen the purpose of the study by emphasizing several key aspects e.g. (i) Although the current version of CoSWAT-WQ has not yet been validated at high temporal resolution, the model is capable of simulating intra-annual and even daily simulations in nutrient and pollutant loading. This temporal resolution provides the potential to better understand the timing of peak pollution events and align management actions accordingly, especially under global changes. (ii) By leveraging globally available datasets within a consistent modelling framework, CoSWAT-WQ supports scalable applications that can be used for both local catchment-level studies and global water quality assessments. This dual capacity enhances its utility for comparative analysis across regions. (iii) We highlight the community-based nature of CoSWAT-WQ. Users can download and apply individual subbasin models for their own research, including further validation and calibration where local data is available. Importantly, users also have the option to contribute improved or calibrated parameter sets for specific subbasins back to the model database. This participatory structure is designed to enhance model accuracy over time and foster a collaborative global modelling effort.

These points will be included in the revised manuscript to better convey the novelty and value of the CoSWAT-WQ model.

3. The Discussion section could be strengthened by a more thorough examination of certain modeling choices and their implications on the study's findings.

    - Line 162, "The ratios are conservative and will be updated in future versions…". I suggest reflecting on the implications of the conservative ratios on the model outputs.

Response: We thank the reviewer for this suggestion. Applying the same speciation ratios for all subbasins globally introduces a degree of simplification that may not fully capture spatial variability in wastewater discharge characteristics. This may lead to over- or underestimation of both bioavailable (e.g., NH_4^+, PO_4^{3-}) and less bioavailable (e.g., organic N and P) nutrient fractions. For example, the model could underestimate bioavailable loads in regions with low treatment efficiency, or overestimate them where advanced treatment reduces reactive forms. Similarly, it could misrepresent the quantity of organic-bound forms that contribute to long-term nutrient cycling. While our current model version applies the same speciation fractions globally, future versions should incorporate region-specific speciation ratios to better capture spatial variability in nutrient forms as this would improve the accuracy of regional and local water quality assessments.

These points will be included in the revised manuscript to better reflect on the speciation ratios/fractions.

- The study used and downscaled several global input datasets. In Section 4.1, the uncertainties in the input data are acknowledged in the Discussion section of the manuscript. Yet, the authors could elaborate more on the implications of these uncertainties on the results.

Response: We thank the reviewer for this comment. As noted, the model relies on several global input datasets, which inherently introduce uncertainties into our simulations. While these datasets are essential for achieving global coverage, they come with limitations that can affect model accuracy in multiple ways. For example, crop management data at a 0.5-degree resolution may fail to capture the heterogeneity of agricultural practices, including crop types, fertilizer application rates, and management intensity. This can lead to misrepresentation of nutrient source areas at the local level. Similarly, using gridded climate data (e.g., precipitation) at the same resolution can significantly influence hydrological processes such as runoff generation, which are key drivers of nutrient mobilization and transport. Point source nutrient loads, which are derived from global gridded datasets, also carry substantial uncertainty, especially because such data distribute emissions over grid cells, whereas in reality, wastewater discharges occur at specific point locations. This spatial generalization may lead to mismatches between modelled and actual nutrient input locations, particularly in urban or industrial regions. These issues are further compounded in regions with sparse observational data, where global datasets rely more heavily on interpolation or modelling assumptions. As a result, model outputs in such areas may be less reliable in absolute terms, although they may still be useful for identifying relative patterns, trends, or hotspots. Thus, to improve the robustness of future simulations, regionally validated input datasets should be integrated wherever available.

We intend to include this discussion in the revised manuscript.

- To my understanding, the SWAT+ model accounts for five crops that represent different croplands (Table 2 and Figure 2, Nkwasa et al., (2022)). However, globally, many different crops exist, with each having a characteristic cropping pattern. Hence, this raises the question of whether the representative crops as used in SWAT+ are also representative for a global application of the model (e.g. rice does not seem to be included). I suggest the authors justify their choice and reflect on uncertainties associated with the crop selection.

Response: We thank the reviewer for highlighting this important point. To give context, the global crop dataset we use classifies crops into broad functional types: C3 annual crops (C3ann), C3 perennial crops (C3per), C4 annual crops (C4ann), C4 perennial crops (C4per), and C3 nitrogen-fixing crops (C3nfx). For each cropland category, a single representative crop was selected based on the global crop distribution estimates by Leff et al. (2004), as shown in Table 2 (Nkwasa et al., 2022). For example, Both wheat and rice are classified as C3 annual crops; however, wheat was selected as the representative crop for this category because it occupies the largest (22 %) global cultivation area among the major crops (Leff et al., 2004),. However, we acknowledge that in certain regions, particularly South and Southeast Asia, rice is the dominant crop and plays a significant role in local hydrology and nutrient dynamics. This generalization introduces uncertainties in regional simulations, especially in areas where the selected representative crop differs significantly from the actual dominant crop in terms of phenology, rooting depth, water requirements, and nutrient uptake. In the case of rice, for example, flooded field conditions can lead to different nutrient transformation and transport processes compared to dryland wheat systems (Zhao et al., 2012). As a result, runoff and leaching estimates in rice-growing areas may be misrepresented in the current model setup. We recognize this as a limitation of using simplified, globally representative crop types for large-scale applications. Future model development will explore incorporating region-specific dominant crops, particularly in hotspot regions, to better reflect local agricultural practices and reduce uncertainty in nutrient loss estimates.

We intend to include this discussion in the revised manuscript.

- Lines 282-284, "…but also to improve processes such as lakes and reservoir implementations as they have a big influence on nutrient enrichment and residence times". Could you provide some more insights on this? For example, include a reference or define 'a big influence'.

Response: We thank the review for this suggestion. To provide further context, lakes and reservoirs play a critical role in nutrient dynamics by through retention or transformation within a watershed. Globally reservoirs act as "sinks" for nitrogen and phosphorus, although the removal rate of phosphorus exceeds that of nitrogen (Gan et al., 2025). They influence nutrient enrichment and residence times through processes such as sedimentation, biological uptake, and nutrient recycling, which can significantly alter nutrient concentrations downstream. For example, reservoirs often increase water residence time, promoting nutrient settling and uptake by aquatic organisms, thereby reducing nutrient loads reaching downstream ecosystems (Yin et al., 2024). Similarly, lakes can retain nutrients for extended periods, affecting timing and magnitude of nutrient export. Recent studies in the Amazon, the world's largest river basin have highlighted the substantial impact of damming, showing how individual tributary basins experience reductions in downstream sediment and nutrient supply (Best, 2019). Thus, including more detailed lake and reservoir process representations in the model would therefore improve the accuracy of nutrient load estimates and better capture the spatial and temporal variability in nutrient enrichment within river basins, especially rivers that are dammed.

We intend to include this discussion in the revised manuscript.

4. I suggest the authors to check the citations used.

    - In Section 2.3, several global water quality models are mentioned, including DynQual, MARINA, and IMAGE-GNM. However, I recommend the authors to check the citations. For IMAGE-GNM, the model version of Beusen et al., (2015) is cited, while a more updated version exist as described in Beusen et al., (2022). For the MARINA model, the calibrated model version of Strokal et al., (2016) is cited, while the authors refer to the uncalibrated version (e.g. Micella et al., (2024).

Response: Thank you for your suggestions. Actually, for the IMAGE-GNM model, with suggestions from the authors, we will revise IMAGE-GNM to IMAGE-DGNM and use both citations i.e. (Beusen et al., 2015; Beusen et al., 2022).

So we will correct this and align the citations in the revised manuscript.

    - In Section 3.1, the authors compare the global TN export estimates of CoSWAT-WQ with other global models. Here, I also suggest the co-authors to check the values and the citations. For example, according to the manuscript estimates for total N export include "39.1 Tg/yr from IMAGE-GNM", whereas Table 1 in Beusen et al. (2022) reports 41 Tg/yr. Perhaps the authors have excluded aquaculture for comparison purposes. If this is the case, I suggest the authors specify this. Next to the IMAGE-GNM model, estimates of the MARINA model are provided "34.8 Tg/yr from the MARINA model (Strokal et al., 2021)" whereas Strokal et al., (2021) report only inputs to rivers from urban sources.

Response: Thank you for spotting this. We will rectify the citations and cross-check the values.

5. Lastly, I would like to suggest some textual changes:

    - The title refers to a "community global SWAT+ water quality model". Yet, the word "community" does not appear in the rest of the manuscript. In the Conclusion, the authors refer to the free accessibility and highly customizable aspects of the model. I suggest the authors to integrate the word 'community model' somewhere to strengthen this message.

Response: Thank you for spotting this. We will incorporate more the word "community model" in the manuscript text to strengthen the message.

- The abbreviations for N and P are introduced multiple times in the Introduction section; the same applies to HRU. I suggest only introducing them once.

Response: Thank you for the suggestion. We will rectify the introductions of N, P and HRUs.

- Line 68, "This makes physically based model approaches…", could be rephrased to make the connection between process-based models and physically based models more clear.

Response: Thank you for the suggestion. We will rephrase the sentence to connect process-based and physically based models more clearly.

- Line 82, "Despite being data-intensive, SWAT(+) models can now benefit from…", I would start the sentence with "SWAT(+) models can now benefit from….", as the first part may cause confusion due to the aforementioned lack of data and the computational intensity already becomes clear from the rest of the sentence.

Response: Thank you for the suggestion. We will rephrase the sentence as suggested.

- Line 80, "application" to "applications"

Response: Thank you for the suggestion. We will rectify this in the revised manuscript.

- Lines 90-91, "to identify hotspots and trends", does this refer to seasonal trends or future trends? Or both?

Response: Thank your questions. It refers to both. We will make this clear in the revised manuscript.

- "CoSWAT-WQ" or "COSWAT-WQ" both spellings are used

Response: Thank you for spotting this. We will harmonize the wording in the revised manuscript.

- "nonpoint" and "non-point" are both used, choose one to stay consistent

Response: Thank you for spotting this. We will harmonize the wording in the revised manuscript.

- Section 2.1 is titled "SWAT+ model description", change to "CoSWAT-WQ model description"? As I would expect a description of the newly developed model.

Response: Thank you for the suggestion. We will rephrase the sentence as suggested.

- Line 156, "was extracted" to "were extracted"

Response: Thank you for the suggestion. We will rephrase the sentence as suggested.

- Lines 259-261, "Hydrology plays a particularly important role in model results, though this discussion focuses on nutrient-related processes. Particularly, there is a need to improve river flow modelling and to incorporate water management features such as lakes and reservoirs, as highlighted in Chawanda et al. (2025)." The use of the word 'particularly' feels a bit odd here, as the previous sentence highlights the focus on nutrient-related processes.

Response: Thank you for the suggestion. We will rephrase the sentence as suggested in the revised manuscript and remove the word “particularly”.

References included in this review [all references suggested here are already included in the manuscript or are linked to water quality models that have already been referred to in the manuscript]

- Nkwasa, A., Chawanda, C. J., Jägermeyr, J., & Van Griensven, A. (2022). Improved representation of agricultural land use and crop management for large-scale hydrological impact simulation in Africa using SWAT+. Hydrology and Earth System Sciences, 26(1), 71-89.
- Jones, E. R., Graham, D. J., van Griensven, A., Flörke, M., & van Vliet, M. T. (2024). Blind spots in global water quality monitoring. Environmental Research Letters, 19(9), 091001.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., King, K.W., 2005. SWAT theoretical documentation. Soil Water Res. Lab. Grassl. 494, 234–235.
- Beusen, A. H., & Bouwman, A. F. (2022). Future projections of river nutrient export to the global coastal ocean show persisting nitrogen and phosphorus distortion. Frontiers in Water, 4, 893585.
- Micella, I., Kroeze, C., Bak, M. P., Tang, T., Wada, Y., & Strokal, M. (2024). Future scenarios for river exports of multiple pollutants by sources and sub-basins worldwide: Rising pollution for the Indian Ocean. Earth's Future, 12(11), e2024EF004712.

Citation: https://doi.org/10.5194/egusphere-2025-703-RC1

References:

Arnold, J.G., Kiniry, J.R., Srinivasan, R., Williams, J.R., Haney, E.B., Neitsch, S.L., 2013. SWAT 2012 input/output documentation. Texas Water Resources Institute.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Van Griensven, A., Van Liew, M.W., 2012. SWAT: Model use, calibration, and validation. Trans. ASABE 55, 1491–1508. https://doi.org/doi:10.13031/2013.42256

Best, J., 2019. Anthropogenic stresses on the world’s big rivers. Nat. Geosci. 12, 7–21. https://doi.org/10.1038/s41561-018-0262-x

Beusen, A.H.W., Doelman, J.C., Van Beek, L.P.H., Van Puijenbroek, P.J.T.M., Mogollón, J.M., Van Grinsven, H.J.M., Stehfest, E., Van Vuuren, D.P., Bouwman, A.F., 2022. Exploring river nitrogen and phosphorus loading and export to global coastal waters in the Shared Socio-economic pathways. Glob. Environ. Change 72, 102426. https://doi.org/10.1016/j.gloenvcha.2021.102426

Beusen, A.H.W., Van Beek, L.P.H., Bouwman, L., Mogollón, J.M., Middelburg, J.B.M., 2015. Coupling global models for hydrology and nutrient loading to simulate nitrogen and phosphorus retention in surface water–description of IMAGE–GNM and analysis of performance. Geosci. Model Dev. 8, 4045–4067.

Gan, J., Wang, X., Yuan, Q., Xing, X., Liu, S., Du, C., Zheng, Y., Liu, Y., 2025. Impact of damming on nutrient transport and transformation in river systems: A review. Water Sci. Eng. 18, 209–220. https://doi.org/10.1016/j.wse.2024.11.001

Hu, C., Sadras, V.O., Lu, G., Zhang, P., Han, Y., Liu, L., Xie, J., Yang, X., Zhang, S., 2021. A global meta-analysis of split nitrogen application for improved wheat yield and grain protein content. Soil Tillage Res. 213, 105111. https://doi.org/10.1016/j.still.2021.105111

Jones, E.R., Bierkens, M.F.P., Wanders, N., Sutanudjaja, E.H., van Beek, L.P.H., van Vliet, M.T.H., 2023. DynQual v1.0: a high-resolution global surface water quality model. Geosci. Model Dev. 16, 4481–4500. https://doi.org/10.5194/gmd-16-4481-2023

Jones, E.R., Graham, D.J., Griensven, A. van, Flörke, M., Vliet, M.T.H. van, 2024. Blind spots in global water quality monitoring. Environ. Res. Lett. 19, 091001. https://doi.org/10.1088/1748-9326/ad6919

Kumar, R., Heße, F., Rao, P.S.C., Musolff, A., Jawitz, J.W., Sarrazin, F., Samaniego, L., Fleckenstein, J.H., Rakovec, O., Thober, S., Attinger, S., 2020. Strong hydroclimatic controls on vulnerability to subsurface nitrate contamination across Europe. Nat. Commun. 11, 6302. https://doi.org/10.1038/s41467-020-19955-8

Leff, B., Ramankutty, N., Foley, J.A., 2004. Geographic distribution of major crops across the world. Glob. Biogeochem. Cycles 18. https://doi.org/10.1029/2003GB002108

Micella, I., Kroeze, C., Bak, M.P., Strokal, M., 2024. Causes of coastal waters pollution with nutrients, chemicals and plastics worldwide. Mar. Pollut. Bull. 198, 115902. https://doi.org/10.1016/j.marpolbul.2023.115902

Nkwasa, A., Chawanda, C.J., Jägermeyr, J., van Griensven, A., 2022. Improved representation of agricultural land use and crop management for large-scale hydrological impact simulation in Africa using SWAT+. Hydrol. Earth Syst. Sci. 26, 71–89. https://doi.org/10.5194/hess-26-71-2022

Nkwasa, A., James Chawanda, C., Theresa Nakkazi, M., Tang, T., Eisenreich, S.J., Warner, S., van Griensven, A., 2024. One third of African rivers fail to meet the 'good ambient water quality' nutrient targets. Ecol. Indic. 166, 112544. https://doi.org/10.1016/j.ecolind.2024.112544

Virro, H., Amatulli, G., Kmoch, A., Shen, L., Uuemaa, E., 2021. GRQA: Global River Water Quality Archive. Earth Syst. Sci. Data 13, 5483–5507. https://doi.org/10.5194/essd-13-5483-2021

Yin, Y., Yang, K., Gao, M., Wei, J., Zhong, X., Jiang, K., Gao, J., Cai, Y., 2024. Declined nutrients stability shaped by water residence times in lakes and reservoirs under climate change. Sci. Total Environ. 953, 176098. https://doi.org/10.1016/j.scitotenv.2024.176098

Zhao, X., Zhou, Y., Min, J., Wang, S., Shi, W., Xing, G., 2012. Nitrogen runoff dominates water nitrogen pollution from rice-wheat rotation in the Taihu Lake region of China. Agric. Ecosyst. Environ. 156, 1–11. https://doi.org/10.1016/j.agee.2012.04.024