Responses to Reviewers

Green is quoted comment; Orange is a response, and Blue is quoted text from manuscript.

Reviewer 2

- I think the introduction section needs some work. I found this section to be meandering and in my opinion these many separate paragraphs about the modelling applications are not needed. My recommendation is to squeeze these texts in few paragraphs and then try to elaborate the research gap and SWAT large scale applications as well. If I remember correctly, there were a handful of studies that reported comparing SWAT model in very large scales, are we missing some relevant references here?

  Thank you for the feedback on the introduction and for prompting us to ensure we have included relevant references regarding large-scale SWAT applications. We have condensed the discussion of general GHM applications as requested

  Regarding the specific point on large-scale SWAT studies, our work aims to build upon previous efforts by developing the first high-resolution *global* SWAT+ model. The most relevant precursors in terms of scale are indeed the continental applications. We have cited the key studies in this domain within the introduction (Line 100), specifically:

  - ***SWAT*** : The continental-scale model for Europe by Abbaspour et al. (2015).
  - ***SWAT+:*** The recent continental-scale models developed for Africa (Chawanda et al., 2024; Nkwasa et al., 2024)

  While a limited number of valuable studies have applied SWAT+ to large *river basins* worldwide (in part due to SWAT+ being relatively new), these continental-scale applications represent the largest spatial domains modelled prior to our global study and thus form the most direct context for the research gap our work addresses – the lack of a fully global, high-resolution SWAT+ implementation. We believe these references adequately cover the state-of-the-art in very large-scale SWAT+ modelling relevant to our study's objectives.

- The issue that the authors reported as a potential future improvement with regards to the comparison with SWAT simulated global ET with the GLEAMS dataset are in my opinion somewhat previously known knowledge. Many researchers have reported the issue with GLEAM dataset that the dataset due to partitioning of ET, tends to overestimate transpiration while underestimating soil evaporation. For example see Chen et al. (2022) "Uncertainties in partitioning evapotranspiration by two remote sensing-based models". I also believe that GLEAM v4 is available now, with a higher spatial resolution than the v3. Have the authors considered using this product to do their comparison?

  Thank you for your comment regarding the discussion of the evapotranspiration (ET) comparison and the GLEAM dataset.

  Regarding the potential reasons for discrepancies, we agree that differences between modeled ET and remote sensing products can arise from uncertainties in both the model and the benchmark dataset. We have revised the discussion in Section 4 to reflect this more explicitly. While we maintain that input data limitations (e.g., climate data resolution, lack of lake representation) contribute to some observed differences, we now also acknowledge that inherent characteristics and potential uncertainties within the GLEAM product itself, such as its partitioning methods as you noted, can also play a role in comparison results. We appreciate the reference provided (Chen et al., 2022) and have incorporated this perspective into our revised discussion.

  Line 312:

  For instance, the East African rift valley lake area was all simulated with regular HRUs while implementing lakes would ensure that the land ET and lake ET are not mixed up to improve spatial Pattern (Fig 9). Concurrently, inherent uncertainties within the GLEAM v4 product itself, potentially related to its algorithms for partitioning ET components such as transpiration vs. soil evaporation, as discussed in studies like Chen et al. (2022), can also influence the comparison results. Thus, there is a need to acknowledge these combined uncertainties when interpreting the evaluation of the ET spatial patterns.

Regarding the GLEAM version, we confirm that GLEAM v4 was indeed used for the comparison in this study. We apologize for any confusion caused by an inconsistency in the manuscript text. We have ensured that the manuscript now consistently refers to GLEAM v4 throughout the methods and results sections (including updates to lines 173 and any other mentions).

- I am missing the details about time range for the ET comparison.

Indeed, we now have highlighted the time range used.

Line 263:

A comparison of ET for the effective simulation period (1982 – 1990) shows that the spatial pattern between SWAT+ ET and GLEAM ET is comparable overall.

- I would also like to know was non-availability of spatial data the sole reason for choosing the ASTER GDEM data over SRTM? I think this is important because SRTM does better representation of mountainous regions than ASTER GDEM, so I would like to see what the authors think about this issue.

Thank you for raising the point about the relative qualities of ASTER GDEM and SRTM, particularly concerning mountainous regions. We acknowledge that at their native resolutions, SRTM data is often cited as having advantages in complex terrain compared to earlier ASTER versions.

However, for our specific application – developing a global hydrological model – the primary limiting factor of SRTM is its incomplete spatial coverage, as it does not extend to latitudes beyond 60°N and 56°S. Since our goal was to model the global landmass (excluding Greenland), the comprehensive spatial extent offered by ASTER GDEM was essential.

Furthermore, while differences in data quality (e.g., noise, artifacts) between the two DEMs are more pronounced at their native ~30-90m resolutions, our global model utilizes these data resampled to a 2km resolution. At this coarser scale, used for watershed delineation and topographic parameterization in a global context, the finer-scale differences between ASTER and SRTM become less significant for capturing the overall large-scale topographic features driving hydrological processes. Therefore, given the necessity of complete spatial coverage for our global domain and the mitigating effect of our 2km working resolution, ASTER GDEM was deemed the most suitable choice for this study. We have added a sentence to the manuscript in Section 2.1.1 to briefly clarify this consideration.

Line 144:

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global DEM (Abrams, 2016) was preferred over the Shuttle Radar Topography Mission (SRTM) global DEM (Farr et al., 2007) primarily due to its more complete global spatial coverage (Fig. 1), which is essential for this study's domain. While potential differences in DEM quality exist between the datasets, particularly in mountainous regions at finer native resolutions, these differences are considered less critical at the 2km resolution used for deriving topographic parameters in this global model setup.

- Could the authors provide any details about the actual runtime for basins of different spatial scale?

Thank you for asking for details on computational runtime. This is indeed an important practical aspect of large-scale modeling. We have added specific examples to the Methodology section (Section 2.2) based on our experience setting up and running the model using the CoSWAT framework on the 64-core HPC environment described in the paper.

We clarified that runtime for both model setup (data processing, watershed delineation, HRU generation, file writing) and simulation execution depends significantly on the specific hardware (CPU cores, clock speed, I/O performance) and the parallel processing configuration employed within the framework

Line 206:

The CoSWAT framework was optimised by iteratively implementing parallel processing wherever possible and feasible. This reduces the time required for data processing and model setup, making large-scale simulations feasible by leveraging High Performance Computing (HPC) environments which often allow highly parallelised workflows (Chawanda et al., 2020). The efficiency gains from parallel processing significantly reduce computation time, though actual runtimes depend heavily on the specific HPC hardware (CPU cores, clock speed, Input/Output (I/O) speed) and parallel configurations used. For context, using the 64-core, 3.00 GHz, 128 GB RAM Linux environment described below (Section 2.3), the CoSWAT framework setup phase (including data preprocessing, watershed delineation, HRU generation, and file writing – not including data download times) required approximately 12 minutes for a moderately sized region such as Save Basin in Africa, and about 1 hour 49 minutes for a large, complex region such as the Nile Basin. Executing a 10-year SWAT+ simulation for very large basins like the Amazon could take over 24 hours, with runtime strongly influenced by the requested output frequency (e.g., daily outputs requiring significantly more time due to I/O demands).

- Line 67 : Explain CC

  We now write CC in full upon first mention as suggested.

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., and Kløve, B.: A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model, J. Hydrol., 524, 733–752, https://doi.org/10.1016/j.jhydrol.2015.03.027, 2015.

Abrams, M.: ASTER GLOBAL DEM VERSION 3, AND NEW ASTER WATER BODY DATASET, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., XLI-B4, 107–110, https://doi.org/10.5194/isprs-archives-XLI-B4-107-2016, 2016.

Chawanda, C. J., George, C., Thiery, W., Griensven, A. V., Tech, J., Arnold, J., and Srinivasan, R.: User-friendly workflows for catchment modelling: Towards reproducible SWAT+ model studies, Environ. Model. Softw., 134, 104812, https://doi.org/10.1016/j.envsoft.2020.104812, 2020.

Chawanda, C. J., Nkwasa, A., Thiery, W., and Van Griensven, A.: Combined impacts of climate and land-use change on future water resources in Africa, Hydrol. Earth Syst. Sci., 28, 117–138, https://doi.org/10.5194/hess-28-117-2024, 2024.

Chen, H., Zhu, G., Shang, S., Qin, W., Zhang, Y., Su, Y., Zhang, K., Zhu, Y., and Xu, C.: Uncertainties in partitioning evapotranspiration by two remote sensing-based models, J. Hydrol., 604, 127223, https://doi.org/10.1016/j.jhydrol.2021.127223, 2022.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Rev. Geophys., 45, 2005RG000183, https://doi.org/10.1029/2005RG000183, 2007.

Nkwasa, A., James Chawanda, C., Theresa Nakkazi, M., Tang, T., Eisenreich, S. J., Warner, S., and Van Griensven, A.: 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, 2024.