Response to the comments of anonymous referee #1

We would like to thank anonymous referee #1 for the comments that help to improve the manuscript. Below are our responses to the concerns raised. The comments of anonymous referee #1 are shown in black. Authors' responses are shown in blue.

Summary
The authors present multi-objective calibration and evaluation for high-resolution hydrologic simulations over France produced by the ORCHIDEE land surface model (LSM). They conduct a comprehensive evaluation for the model performance, considering both classic goodness-of-fit indicators including KGE and bias and trends in streamflow and ET. The comparison is promising. Overall, the manuscript is well-written with high-quality figures.
However, I still have the following concerns.

Response: We thank this positive evaluation of our work. We are also thankful for the constructive comments and suggestions provided which have certainly helped improve the paper.

1. In the abstract, the authors claim that they present a strategy to obtain a realistic hydrological simulation over France, but in fact, the ORCHIDEE land surface model does not consider human impacts, leading to poor performance in some regions in France. Therefore, it would be more precise to state "a reliable hydrological simulation".

Response: We thank anonymous referee #1 for this suggestion. Following this suggestion, we will change "realistic" to "reliable".

2. Through the comparison of streamflow and ET, the model always performs better against GLEAM than FLUXCOM. Is that because GLEAM reanalysis data is still modeling data and FLUXCOM is generated based on observational datasets? If the ORCHIDEE land surface model uses similar physics equations to those of GLEAM, we may expect the results from ORCHIDEE and GLEAM are in good agreement.

Response: There are no perfect evapotranspiration products given the inherent uncertainties (e.g., Liu et al., 2023; Xie et al., 2024). Both GLEAM and FLUXCOM products are based on sound methodologies: GLEAM generates ET at 0.25° resolution from 1980 to 2020 based on the Priestley-Taylor potential evapotranspiration formula with satellite-based products (net radiation, precipitation, surface soil moisture, skin and air temperatures, vegetation optical depth, and snow water equivalent) as inputs; FLUXCOM relies on machine learning algorithms to spatially interpolate in situ FLUXNET measurements at 0.5° resolution from 2001 to 2015, using constraints from remote sensing and meteorological observations.

ORCHIDEE calculates ET as the sum of plant transpiration, evaporation of intercepted water, soil evaporation and snow sublimation based on water and energy budgets.

We can consider, however, that ORCHIDEE, GLEAM and FLUXCOM provide independent estimates of ET, all with their own uncertainties (Liu et al., 2023). In this framework, our guideline was to compare the ORCHIDEE simulation with several products, which offers plausible range. Eventually, in our study, we do not conclude that “the model always performs better against GLEAM than FLUXCOM”: the bias of the simulated ET to GLEAM over the entire study domain is better than that to FLUXCOM, but the simulated ET is more spatially consistent with FLUXCOM than GLEAM (L302-309 and Figure 4).

To be clear, we did not stop at EXP4 because it has a good bias value compared with GLEAM (-0.5%). Instead, ET is underestimated compared with FLUXCOM (-4.3%) and Q is overestimated compared with observations (6.3%), which means that there is probably a physical consistency between these two datasets. We simulated the natural behavior of the French water system without considering human perturbations, such as pumping and irrigation, which could result in an underestimation of ET and an overestimation of Q if we consider FLUXCOM and Q observations at the same time. Anyway, we did not seek to obtain perfect bais values against these datasets but we try to make some compromises to make our manual calibration more reasonable and reliable.

Following this comment, we will add a sentence in L203: “GLEAM and FLUXCOM provide independent ET estimates, both of them with large uncertainties (Liu et al., 2023). They are used in combination to approach the plausible range of observed ET.”

3. In the Introduction section (Line 77-87, Page 3), the authors introduce the first distributed LSM at the nationwide scale of France, SIM. SIM has shown very good performance in generating hydrologic simulations. Why do the authors decide to use another LSM, ORCHIDEE for France? What are the limitations of SIM?

Response: There are no perfect models for hydrological simulations and each model has its own strengths and limitations. It is always encouraging to have several models that are capable of obtaining reliable hydrological simulations at the nationwide scale of France, especially when the goal is to test the response to changing conditions (e.g. climate change, land use change, etc.) . Multi-model assessment is more robust than single one model for hydrological simulation and projection. Both SIM and ORCHIDEE LSMs have contributed to the national EXPLORE2 project (https://professionnels.ofb.fr/fr/node/1244) for climate change impact analysis.

4. In Line 264-265, Page 11, the authors state that "The timelag criterion of the simulated Q is also greatly improved from a range of -11 to 27 days to a range of -3 to 5 days.". In my opinion, the zero timelag is the best, right?

Response: Yes, and we will underline this in L211 of the paper for the sake of clarity.

5. Specific comments: Line 102, Page 4: "(revision 7738)" should be deleted.

Response: Thanks for the suggestion. This will be deleted from the title, but for clear documentation of the code (refer to EGU guidelines), we will change L103-105 to "The ORCHIDEE model is a physically-based LSM developed at the Institut Pierre Simon Laplace (IPSL) as the land component of the IPSL climate model, which is used for all the past and future climate simulation exercises carried out for the IPCC reports as part of the Coupled Model Intercomparison Project (CMIP) (IPCC,2023). Here, we use ORCHIDEE version 2.2 (with revision 7738), which is very close to the version used as the land component of the IPSL-CM6 climate model (Boucher et al., 2020; Cheruy et al., 2020)."

6. Line 219, Page 8: "STD" first appears. What is "STD"?

Response: We will reconstruct the sentence in L219 to "The starting point experiment of calibration design is called STD and uses the "standard" parameter set sourced from CMIP6…".

7. Table 2: Please explain the meanings of the labels in the caption, such as "PPV".

Response: Thank you for this suggestion and for the sake of clarity, we will explain the meanings of the labels in section 2.3.

Final note: We will also correct the style of some sentences so that they are more readable, and orthographic and grammar errors.

References

Boucher et al. (2020), Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, J. Adv. Model. Earth Syst., 12, e2019MS002010, doi: 10.1029/2019MS002010.

Cheruy et al. (2020), Improved Near-Surface Continental Climate in IPSL-CM6A-LR by Combined Evolutions of Atmospheric and Land Surface Physics, J. Adv. Model. Earth Syst., 12, e2019MS002005, doi: 10.1029/2019MS002005.

IPCC (2023), Climate Change 2021 - The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, doi: 10.1017/9781009157896.

Liu et al. (2023), Intercomparison and evaluation of ten global ET products at site and basin scale, J. Hydrol., 617, 128887, doi:10.1016/j.jhydrol.2022.128887.

Xie et al. (2024), Evaluation of seven satellite-based and two reanalysis global terrestrial evapotranspiration products, J. Hydrol., 630, 130649, doi:10.1016/j.jhydrol.2024.130649.