

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-64-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-64

Anonymous Referee #2

Referee comment on "Modeling reservoir surface temperatures for regional and global climate models: a multi-model study on the inflow and level variation effects" by Manuel C. Almeida et al., Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2021-64-RC2>, 2021

In this work, Almeida et al. compared the performance of 2-D lake models with/without accounting for lateral flow, 1-D lake models (Hostetler-based and FLake) and data-based ANN models in simulating the thermal regimes of 24 reservoirs in Portugal. They demonstrated that for reservoirs with short WRT, it is important to represent the effect of lateral flow and water level fluctuation in the lake models of GCMs and RCMs. Although the importance of lateral flow in the thermal regimes of reservoirs has been investigated by previous studies, the work of Almeida et al. is novel in three aspects: 1) the investigation of a large set of reservoirs, 2) the inclusion of ML methods, and 3) the comparison of multiple 1-D lake models. The manuscript is well written and easy to follow. I agree with the comments of the first reviewer and provide additional comments. I recommend the publication of this work after these comments are addressed.

Major comments

First, I possibly misunderstood but it seems that the 1-D Hostetler-based model was developed by the authors for this work. If so, I do not quite understand the rationale because there are many well-tested Hostetler-based models that have already been publicly available, such as WRF-Lake. As a lake modeler myself, I worry that the development of a new model would unavoidably introduce bugs.

Second, it looks that the 1-D lake models were not calibrated in the study but the ANN model because it is based on the 2-D reservoir model was implicitly calibrated. Thus, in my view, the comparison of their performance in the current format is unfair. According to my own experience, by calibration, 1-D lake models can also mimic some effect of lateral flow and water level change. But whether this is physically sound is another story. However, my point is that the current experiment design does not convince me the superiority of ANN over 1-D models in representing lake thermal dynamics for GCMs and

RCMs because when we have data to train ANN we can also use the data to calibrate 1-D models.

Specific comments

L22-24: as indicated above, I do not think the current results can make such a statement. Further, there is another difficulty for ANN models to replace 1-D lake models in GCMs and RCMs. Compared with ANN models, 1-D lake models are much more generalized because they are physically based. For example, due to the limitation of model resolutions, usually the lake grid cells in GCMs and RCMs do not directly correspond to real lakes. We still do not know whether ANN models trained by data from real lakes can be extended to artificial lake grid cells.

L60: please add references to other Hostetler-based models that are involved in the ISIMIP lake sector, such as Tan et al. (2015). "Tan, Z., Zhuang, Q., & Walter Anthony, K. (2015). Modeling methane emissions from arctic lakes: Model development and site-level study. *Journal of Advances in Modeling Earth Systems*, 7, 459-483."

L66: please cite the most recent modeling intercomparison study of Guseva et al. (2020). "Guseva, S., Bleninger, T., Jöhnk, K., Polli, B. A., Tan, Z., Thiery, W., ... & Stepanenko, V. (2020). Multimodel simulation of vertical gas transfer in a temperate lake. *Hydrology and Earth System Sciences*, 24, 697-715."

Table 1: Did you use the bathymetry data of the 24 reservoirs to setup the models? Or did you only use mean depth, maximum depth and surface area to construct ideal bathymetry for these reservoirs? Sometimes, the uncertainty in bathymetry can introduce large uncertainty in 2-D lake modeling.

L152-154: please rewrite this sentence. It is difficult to understand.

Equation 4: What is the definition of Φ ?

Section 3.3: I suggest adding the Kling-Gupta efficiency (KGE) as a model evaluation metric.

L319-320: It is not true for Hostetler-based models. They can account for the wind sheltering effect, as documented in Guo et al. (2021). The difference is that the 2-D models can account for the direction effect of the wind sheltering but the Hostetler-based

models cannot, which may be important for elongated reservoirs. "Guo, M., Zhuang, Q., Yao, H., Golub, M., Leung, L. R., Pierson, D., & Tan, Z. (2021). Validation and Sensitivity Analysis of a 1-D Lake Model across Global Lakes. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033417."

L323: delete "effect"

L325: This sentence is unclear to me. Do you mean the difference of RMSE between W2-reservoir and W2-lake?

L355: please cite Guo et al. (2021). "Guo, M., Zhuang, Q., Yao, H., Golub, M., Leung, L. R., Pierson, D., & Tan, Z. (2021). Validation and Sensitivity Analysis of a 1-D Lake Model across Global Lakes. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033417."

L369-370: I do not think it is true. As shown in Guo et al. (2021), the thermal regimes of deep lakes usually can be better simulated by Hostetler-based models than shallower lakes because deeper lakes usually have larger Wedderburn numbers. Here, the larger errors in these deeper reservoirs may be caused by other factors. For example, the default parameters, such as light attenuation coefficient, may be not suitable for these deeper reservoirs. Also, lateral flow may destabilize the thermal structure of these reservoirs, making them difficult to simulate by 1-D models.

Section 4.2.2: please also add the computational time of 2-D models for reference.

Figure 8: The caption is confusing. I think all models use the same atmospheric forcing. For 2-D models, what this figure presents is the wind stress after accounting for the sheltering effect. Please make it clear.

L483: All reservoirs tested in this study are under the Mediterranean climate. So the conclusion here is too broad. It is better to say "for the same morphometry and under the Mediterranean climate".