Second review

Reviewer 1

I would like to thank the authors for their revised manuscript, which provides answers to questions raised during the initial review process.

We want to thank reviewer 1 to revise the first and our second version of our manuscript. We follow your last technical comments to do the last correction.

Technical Comments:

L68: replace (Stepanenko et al., 2010, 2013) by Stepanenko et al. (2010, 2013)

L127: replace we use the period 2000-2016 by the period 2000-2016 was used

L130&131: replace day time by daytime and night time by nighttime

L195: replace and that the water balance is not solved by and not the water balance

L197: I suggest to remove the parenthesis and add comas because albedo and extinction coefficient are as important as fetch

L262: 2.29 W m^{-1} K^{-1}

L312: remove the extra space after soils

L316: it is not clear if the spin up was run before 1979 or if it corresponds to the period 1979-1999 (which seems to be the case). If it is so then you could replace "A spin up procedure of 20 years in the previous years was applied to bring…" by "The first 20 years (1979-1999) were used as spin up to bring..."

L327: replace succeed by succeeded

L328: replace are by were

L330: replace correspond and assign by corresponded and assigned

L408: replace Erie and Superior lakes by Lake Erie and Lake Superior

L410: replace Baikal lake by Lake Baikal

L428: replace It should be noted also by It should also be noted

L470: add unit after -0.7

L473&474: show is used 4 times, please rephrase

L463: the meaning of the curves (solid lines) in the description of Figure 7 (histogram) is missing

L589: evaluation can be removed (Choulga et al. showed the impact of depth on ice phenology)

L593: replace do LSWT by does LSWT (meaning surface temperature in general)

Reviewer 2

Review of "Modeling subgrid lake energy balance in ORCHIDEE terrestrial scheme using the FLake lake model" by Bernus Anthony and Ottlé Catherine Article shows results of testing 1-D lake parametrization scheme FLake incorporated in the ORCHIDEE land surface scheme. Experiments are performed at different horizontal resolutions and temporal scales with different lake mean depth aggregation techniques. Experiment results are compared with observed lake surface temperatures and ice start, end, duration dates.

We want to thank reviewer 2 for its useful comments.

GENERAL COMMENTS

In the paper lake depth from HydroLAKES is used as ground truth, yet over some latitudes and regions this data plays a crucial role. It would be good if data was crosschecked/verified at least for categories where only few lakes were used (small number of lakes was used).

Depth validation has been done on a large amount of lakes (5101 lakes) during the elaboration of the HYDROLAKES database and the authors also provide an estimation of the error on the depth. For the 8 large lakes highlighted in the local scale study, we have checked that the HYDROLAKES depth as well as the tile depth simulated are well in accordance with the published literature on the subject. To account for depth uncertainty, we could generate different depth maps for our modeled tiles and look at the impact on the model results but this is another study and could be also the subject of a next paper.

For different lake categories errors are provided, but it is obvious that not all categories have vast amount of lakes inside, so it would be good to see if numbers presented have statistical significance. For example errors can be high for some latitudes but are based only on two lakes, so this is statistically not significant at all.

You are right, and it is for this reason that we do not compare the RMSEs of each category of lake type. In our comparisons (impacts of depth parameterization of atmospheric forcings) , we always compare statistics performed on the same number of lakes. In Figure 3, since indeed the number of lakes vary with latitude and longitude, we have plotted the number of lakes which have been included in the comparison, to put in perspective our results.

Authors use different lake mean depth aggregation techniques – it would be a great advantage if they could also add MODE aggregation technique, or at least say about it in discussion.

The mode value of the lake depths within a tile could be also a way to define the tile depth and will give more weight to the main peak of the distribution if any. We have seen that in some cases where the lake distribution is very heterogeneous, this could lead to significant differences compared to the mean or median values. Furthermore, a large peak mode could mix up shallow and intermediate lakes presenting different physical behaviors and it is preferable to differentiate them. In this study, we wanted to define tiles with depths representative of the 0.5° grid cell, so we think that the median or the mean is a better approach than the mode, and this is why we did not include the mode approach in this work.

Paper has huge amount of abbreviations and some are not explained (e.g., line 74 – LM4, GFDL, etc.) – it would be a big help for the reader to have as little abbreviations as possible, e.g. to keep only the ones that are used all the time, like LSWT, add remove all unnecessary ones, e.g. write them in full and add abbreviation in brackets if necessary.

We removed all the unnecessary abbreviations and checked that all were well defined. We agree that in the presentation of the state of the art, there are a huge amount of acronyms but we think that it is important to leave them because the models in our community are more known by their acronyms than by their full names. Therefore, we don't see any other satisfactory solution than leaving the model names, but we entirely write the institute names instead of their acronyms to limit them at best and ease the reading.

In Section 4.2.1 model is evaluated based on 8 lakes, but all these lakes are huge and their surface temperature can be also due to internal currents, wind fetch, etc., so these lakes can't represent lakes globally even though they are located in different climate zones. Most of lakes globally are quite small in area and around 10 m deep (mean depth). Results are interesting and well presented but it should be noted that they are not representative globally.

We added a sentence to notice this fact in the corresponding section.

Some sections, like Section 4.3, have a very interesting results, for better understanding it would be great to put it in a table – all numbers easy to compare and understand.

We added a table which gives the 5th and 9th decile of the depth error distribution and which is complementary to figure 7. Thanks for your suggestion.

TECHNICAL COMMENTS

l.31 "Lake distribution is spatially unequal all other the world with two regions …" – should instead of "other" be "around"?

Done

l.69 "the UK Met Office Unified Model and its JULES Land simulator (Rooney and Bornemann, 2013)" – I guess FLake there is not couples operationally yet.

l.114 (same as l.31) – "all other the world" – should it be "around"?

Done

l.183 "the free water bottom temperature" – what is meant here exactly?

Done

l.243 "to adjust the time split factor which has been finally set to a value of 50" – could you explain with more details?

We added in this paragraph that we made some tests for different values of this split factor and explained that we had to increase this factor to a value of 50 to remove all the instabilities. We hope that this part is better explained now.

l.387 (same as l.31) – "all other the world" – should it be "around"?

Done

Figure 4 – explain please column plots – what different shading represent?

Done

Figure 7 – explain please left plot – what columns and lines represent exactly?

We added in the label of the figure that the lines represent the smoothed distributions.

l.589 "GHG fluxes like carbon and methane…" – did you mean carbon dioxide?

Done