GMD responses to Reviewer comments.
Reviewer comments in grey, responses in black.

We thank both reviewers for their effort and insightful comments. These are two of the more keen and constructive reviews that we have received. Reviewers identified several key issues: model validation and performance, model temporal resolution, model vertical resolution, sediment temperatures, and Toolik lake inflow data. We addressed these issues by adding model evaluation metrics, adding new appendices for model resolution sensitivity, correcting errors, and updating manuscript text to address comments. We provide detailed responses to these issues below. As the GMD interactive comments do not allow us to submit a revised manuscript at this stage, we have attached excerpts from our revised manuscript that should be viewed along with our responses.

**Major comments:**

Comment: “I have some issues with the study. Firstly, I think that the model baseline simulations were not correctly validated. I can´t fully evaluate the model performance, and/or compare the model results with other model simulations (e.g. Guo et al., 2021, modeled Toolik lake) without a model evaluation metric such as: mean absolute error (MAE) or root mean square error (RMSE). Furthermore, I don´t understand how the model was calibrated. What function were you trying to minimize in order to optimize the model performance?”

Response: Thank you for your effort and insightful comments. We have updated the manuscript with model evaluation metrics (including MAE and RMSE) as requested. Model performance was similar to Guo et al. 2021, with RMSE ~2C. However, it should be noted Guo et al. 2021 simulated Toolik lake only for the thawed seasons of 1983-1988. The LAKE model was minimally calibrated for each lake, as described in Section 2.2, to initialize water and sediment temperatures. A standard set of model parameters were applied to all lakes to demonstrate the applicability of the LAKE model in simulating Arctic lakes (Table 1).

Comment: “Secondly, why didn´t you show the lakes sediment temperature obtained with the model as a function of water temperature? This kind of data is quite relevant for other researchers.”

Response: We have added results showing sediment temperatures and new figures for water temperature profiles in addition the figures already present showing water temperatures (Appendix C). As the focus of this paper was not directly on lake sediment temperatures we did not attempt to demonstrate sediment temperatures as a function of water temperatures.

**Specific comments:**

Comment: “L25: I think that the word ”completes” is very strong.”

Response: We changed it to ‘is’.

Comment: “L26-L29: This sentence is unclear to me. You say that the model “is not highly sensitive to the weather data perturbations”, and you conclude that “snow depth and lake ice strongly affect water temperatures during the frozen season”?”

Response: We have updated the text to clarify our point. “The sensitivity analysis shows us that lake water temperature is not highly sensitive to small changes in air temperature or precipitation, while changes in shortwave radiation and large changes in precipitation produced larger effects. Snow depth and lake ice strongly affect water temperatures during the frozen season which dominates the annual thermal regime. These findings suggest that reductions in lake ice thickness and duration could lead to more heat storage by lakes and enhanced permafrost degradation.”

Comment: “L31: I suggest the following change to this sentence: “Approximately forty percent…””

Response: We have made this change.

Comment: “L70: **Description of the model:** I think that you need to improve the model description, namely, the multilayer snow and ice modules (Stepanenko and Lykossov, 2005; Stepanenko et al., 2011).”

Response: We have elaborated on this section to include a description of the snow and ice modules, including references.

Comment: “L85: **LAKE model setup:** Please describe the calibration procedure. Which parameters were calibrated in which ranges? Was calibration automatic? Please describe the parameters of the baseline simulation. The table 1 included in Stepanenko et al. (2016) is a very good example.”

Response: Our calibration procedure simply involved the initialization of the soil and water temperature values as described in the Section 2.2. No other parameters were calibrated. The parameters of the baseline scenarios have been added as Table 1.

Comment: “L94: **Input data**: Please describe all meteorological variables. How did you characterize the inflow water temperature to lake Toolik? Please describe the initial water temperature and sediments values, before and after the 10 years simulation.”

Response: We have added text to section 2.3 describing all met variables. Inflow water temperature was measured daily with discharge. Water temperature is included in the inflow input file. Discharge and temperature are described in section 2.7. Initial water temperature was taken from observed water temperature data. Initial and calibrated sediment temperatures are now reported in Table 1.

Comment: “L140: Please replace $Wm^{-1}$ with $Wm^{-2}$.”

Response: Thanks, we made this change.

Comment: “L150: Do you have lake water level values? Do you think that neglecting the lake water level may lead to errors in surface heat flux predictions?”

Response: Interesting point. We do not have observations of lake water level values. The water level change may affect surface fluxes via the thickness of the mixed (or active) layer of a lake, the latter is a layer which total heat capacity interacts with the atmosphere. If not limited by lake depth, the typical summertime ML thickness in mid- and high latitudes is 3-5 m (see e.g. simulated/observed temperature profiles in LakeMIP papers). Thus, there are two situations with respect to the lake level effects on ML depth and thus the surface fluxes. First, the lake is shallower than 3-5 m, then the ML is a lake depth. In this case, the water level may affect fluxes, if it varies significantly retaining the depth below 3-5 m. In the case where the lake depth much exceeds 3-5 m (Toolik lake), the level variations do not change ML depth and thus the fluxes.

Comment: “L156: I suggest adding a new section, “Evaluation metrics” for the “new” evaluation metrics (e.g. RMSE). The Z-score equation can also be included here. You don’t need to apply the “new” metrics to the sensitivity analysis.”

Response: We have added this section, now section 2.8.

Comment: “L169: “During the frozen season, the modeled temperatures underestimate cooling in the lake.” By how much?”

Response: We have added Table 2 which shows model error (MAE, RMSE, Bias) for the entire time series, and split by frozen and thawed season. For this particular sentence the error for Atqasuk over the frozen period was 5.8 (RMSE).

Comment: “L189-190: “For 2013 and 2014 the modeled shallow (0, 3 m) water temperature was overestimated while for 2015 and 2016 shallow water temperature was underestimated, though it tracked observed temperature.” By how much?”

Response: We have added Table 2 which shows model error (MAE, RMSE, Bias) for the entire time series, and split by frozen and thawed season. The Toolik model simulations have been updated based on corrected discharge data. This sentence and interpretation of the Toolik water temperatures have been changed. Thawed and frozen season errors are presented in Table 2.

Comment: “L192: I can´t see the step-like dip in figures B1 and B2 can this fact be related with inflow water temperature?”

Response: We thank the keen reviewer who caught this error. We were able to trace the ‘dips’ to a formatting error in the inflow data file. This has been corrected. All Toolik simulations have been repeated and figures updated (Section 3.3). The ‘dips’ were an artifact of the erroneous inflow data and are no longer present (Figs. 3 & 4).

Comment: “L200: The datasets length (x values) shown in figures 3 and 4 is smaller than the datasets length shown in figures B1 and B2.”

Response: These have been corrected to show the same length of data.

Comment: “L210: “shallow depth water temperatures (1, 3, and 5 m 210 depth, -0.13 to 0.34)”. I can´t find the value -0.13 in Figure 5.”

Response: This was an error. The text has been updated to reflect the data in the figure. Please note this figure and data have been updated to reflect the new simulations for Fox Den (now hourly) and Toolik (with corrected inflow data)(Fig. 5).

Comment: “L246: “Modeled shallow water (1 m) temperature exceeded the observed temperatures” After the incorporation of inflows/outflows, the water temperature (1 m) in 2013 and 2014, still exceeds observed water temperatures. This kind of analysis would be easier with a model evaluation metric.”

Response: Error metrics have been added and are included in Table 2, B1, & B2 for this sentence.

Comment: “L270: I think that this entire section “Modeling Lake thermal effects in permafrost” must be in the introduction.”

Response: We have moved this section to the Introduction.

Comment: “L286: “The “dips” of water temperature in LAKE model results for Toolik lake down to depths of 10 m prior to ice-off can be explained”. I can see the dip at 19 m (Figure 4, 2014-07).”

Response: We thank the keen reviewer who caught this error. We were able to trace the ‘dips’ to a formatting error in the inflow data file. This has been corrected. All Toolik simulations have been repeated and figures updated (Section 3.3). The ‘dips’ were an artifact of the erroneous inflow data and are no longer present (Figs. 3 & 4).

Comment: “L287: “can be explained by convective instability under the ice, where this instability can be caused by the under-ice penetration of solar radiation” As I said previously, I can´t see the “dips” in figures B1 and B2. Can this be related with the effect of lake inflow?”

Response: We thank the keen reviewer who caught this error. We were able to trace the ‘dips’ to a formatting error in the inflow data file. This has been corrected. All Toolik simulations have been repeated and figures updated (Section 3.3). The ‘dips’ were an artifact of the erroneous inflow data and are no longer present (Figs. 3 & 4).

R2:

Comment: “Modeling of lake thermodynamics in polar regions is a highly relevant topic with regard to the response of the Arctic permafrost to the global change. The model LAKE has been intensively applied in recent studies on lake dynamics and air-lake interaction. Therefore, a study on the LAKE model abilities to simulate thermal properties of lakes in the permafrost zone falls into the scope of the GMD and is of interest for its readership. Comparison of the model performance for three Arctic lakes of different morphometry provides a necessary background for future analysis of the atmosphere-lake-permafrost interaction. Herewith, the study is a valuable contribution to modeling of lakes as components of the climate system. The methods, presentation of results, and discussion are generally adequate to the problem statement, however contain some gaps, related, in particular, to the effects of the spatial and temporal resolution on the modeling results and to the simulation of the water-sediment interaction as a crucial aspect of lake modeling in the permafrost zone. I recommend extending the study with relevant details providing the reader with a necessary overview of the model performance beyond the sensitivity to variations in the input forcing, which is currently the major focus of the manuscript.”

Response: Thank you for your comments. We have added several new sections to the manuscript and to the Appendix that we believe add more detail and aid in understanding the spatial resolution an temporal resolution on modeling results as well as results of the water-sediment interaction.

Comment: “As it was pointed out by the previous reviewer, the model validation is presented in a rather qualitative way, and some numerical scores of the model performance, like bias, absolute error, RMSE etc., will be useful here.”

Response: We have updated the manuscript with model evaluation metrics (including MAE and RMSE) as requested. Please see Table 2 and the Appendix.

Comment: “The temporal resolution of the model input was different for three different lakes: 1 day for one of them and 1 hour for the two others. It is unclear how the diurnal cycle of the atmospheric forcing and radiation was treated in the model. Were the daily data interpolated on sub-diurnal scales? If yes, how the interpolation was performed? How does the neglect of the sub-daily variations in the input data affect the model output? The question could be answered by comparison of model runs with daily and hourly inputs for the lakes where sub-diurnal data on forcing are available.”

Response: We have updated the simulations to use the same temporal resolution (1 hour) for all lakes. Additionally we have added a section to the Appendix that shows the effect of temporal resolution on model performance (Appendix F). Daily data are linearly interpolated to finer temporal scales within LAKE.

Comment: “The vertical resolution for both water column and sediment was set to 1 meter and did not vary between lakes. What were the criteria for the choice of the resolution? One can assume that for the vertical diffusion rates within the sediment of 10^{-6} m^2 s^{-1}, the vertical resolution of 1 m will capture the processes with typical time scales of >10 days. Is it sufficient? How many vertical grid points did Fox Den have, whose depth is 1.5 m? Can you perform sensitivity runs demonstrating the effect of the vertical resolution on the model output?”

Response: We regret that the text incorrectly stated the vertical resolution for the water column. We have corrected the text. We used 40 layers for the water layer for all lakes which results in a different vertical resolution for each lake (Atqasuk=0.065m, Fox Den=0.04m, and Toolik=0.65m, see Table 1). Our experience has shown that 40 layers is sufficient (Stepanenko et al. 2010, 2013, & 2016). We have performed the additional sensitivity runs demonstrating the effect of vertical resolution on model performance and have included the results in Appendix D.

Comment: “L316, Section 5.4 The details on the sediment layer modeling results are crucial for discussion on the model applicability to permafrost lakes. The information is missing in the ms. How did the soil temperatures under the lake bottom vary during the modeling period? What are the values of the bottom heat flux and how do they depend on the model configuration, initial and boundary conditions?”

Response: We have added new figures and a new section in the results to show the sediment temperatures and heat flux during simulations (Appendix C). Soil temperatures responded differently in each lake. In general, shallow sediment showed warming in the thaw period and deeper sediments remained constant over the simulation period.

Some minor remarks:

Comment: “ “It is a large lake (2,732,050 m2 )...” why 2 km^2 area is large for a lake?”

Response: It is large relative to our study lakes. We have updated the text reflect this comparison.

Comment: “ “ 30 cm and 250 cm” better use meters here for consistency.”

Response: We have made this change.

Comment: “In Fox Den the model calculated up to 1.0 m thick ice cover in a 1.5 m deep lake. Was the water volume/depth adjusted during the ice-covered period? Was 1 m vertical resolution sufficient for simulation?”

Response: All frozen water (which formed the ice layer) is subtracted from the lake water volume. The water depth is adjusted accordingly. As to resolution, the grid spacing in water and ice is automatically adjusted in the model to keep the predefined number of numerical layers in each physical layer. In the manuscript you reviewed, we misstated the vertical resolution used for the simulations (Table 1). We have corrected these errors. For Fox Den the vertical resolution was 0.0375m which we believe was sufficient for simulation.

Comment: “L286: “The “dips” of water temperature in the LAKE model results for Toolik lake…” How did the vertical model resolution affect the representation of free convection? The 1 m resolution seems to be crude for the typical values of the convective layer entrainment rates of < 1 m/day (e.g. Kirillin et al. 2012)."

Response: The statement of 1m resolution was incorrect. We have corrected the text to reflect the vertical resolution used in each Lake (Table 1). For Toolik the resolution was 0.65m. We have added an appendix to look at the effects of increasing model water vertical resolution. Using 1m, 0.5m, and 0.25m vertical resolutions we found minimal effects on lake water temperatures and model performance. Kirillin et al. 2012 report rates of 0.5 m per day increasing to several meters per day in deep lakes. We simulated lakes with vertical resolutions of 0.0635m, 0.0375m, and 0.65m (for Atqasuk, Fox Den, and Toolik respectively) and tested vertical resolutions down to 0.25m for Toolik and 0.025m for Atqasuk (Appendix D). We did not see evidence that the vertical resolutions used in the manuscript was too coarse.