RC1: 'Comment on gmd-2021-75', Shuqi Lin, 13 Apr 2021

General comments:

This study involves tremendous data processing when coupling the river and lake dataset together before conducting the simulations. Could you provide a map or a chart to state the number of river and lakes in different groups you defined and how many systems among them have been processed specifically? I think it will be helpful for the readers to understand the whole dataset and reproduce the framework.

Response: Dear Dr. Lin, thank you for taking the time to review our manuscript. We are pleased your review is supportive of our work and we are happy to respond to the points. We added a table of pre-processing to enhance readers' understanding and uploaded the upscaled river–lake network dataset into zenodo. We also conducted additional validation with the data you recommended and it was really helpful for us to discuss the model applicability and limitation.

Specific comments:

2.1 Harmonization of geographical information

Did you basically implement the lake data from HydroLAKES and river data from MERIT Hydro? I see a lot of preprocessing of lake and river geographical information in the second paragraph. Could you please provide a table or a chart to conclude the results of the preprocessing, like how many lakes are classified into the two groups, respectively, and how many inconsistencies are detected in two datasets and which dataset contained the largest upstream area you chosen in the end, etc.

Response: Yes, the lake and river data were upscaled from HydroLAKES and MERIT Hydro, respectively. To show the variables the pre-processing updated, we added a table to summarize the pre-process as you recommended, which also indicates the manual processing.

In addition, we added the number (369) of lakes resolved in the river–lake network dataset.

We detected the inconsistency between the flow direction of MERIT Hydro and lake distribution of HydroLAKES and filled the gap in the pre-processing. However, those methods are minimal from the perspective of conduct of river–lake coupling simulation. We added a discussion to emphasize the importance of the development of a comprehensive geographical dataset explicitly representing rivers and lakes in the discussion section.

Table 1 Summary of harmonization of geographical information, MERIT Hydro and HydroLAKES. All processes except number 4 are automated.

No.	Process	Updated variable		Reference data
		MERIT Hydro	HydroLAKES	
1	Select lakes to resolve in the river–lake network	-	lake area	lake area
2	Fill in isolated parts of each lake	-	lake area	lake area
3	Select lake outlet	-	-	upstream area calculated in MERIT Hydro
4	Remove outlets from endorheic lakes	flow direction	-	actual geography
5	Change flow direction in each lake	flow direction	-	lake outlet location
6	Recalculate upstream area for all the grids	upstream area	-	flow direction

Line 89: Could you provide the links of these dataset here?

Response: As you recommended, we uploaded the dataset into the zenodo and updated the doi for the new version (the source code was not updated).

3.2 Lake model

Line 166: Any reference of this 1D lake model?

Response: The lake model in this study is not original because it is just a combination of each scheme of the existing models, so we only give an overview and cite corresponding studies.

Line 265: Should have a punctuation after the back bracket.

Response: Thank you for your correction. We added punctuation.

Line 298: How is the shortwave radiation weighted by the area of ice? Could you provide the equation here?

Response: We assume that incoming shortwave radiation and surface heat fluxes are boundary conditions for the energy budget of the water body, and they are different between ice-free and ice-covered lake areas. Shortwave radiation attenuates due to ice thickness in ice-covered areas. The heat flux into the water body is conductive heat beneath ice cover, or heat fluxes from and out of the atmosphere for the ice-free part. So, the model calculates the weighted mean of them with the ice-covered and ice-free areas. We added the explanation instead of equations.

3.3 Implementation of coupling interface

Line 323: For how many river-lake systems in your study you have made the corrections? Are they the minor part of the whole dataset?

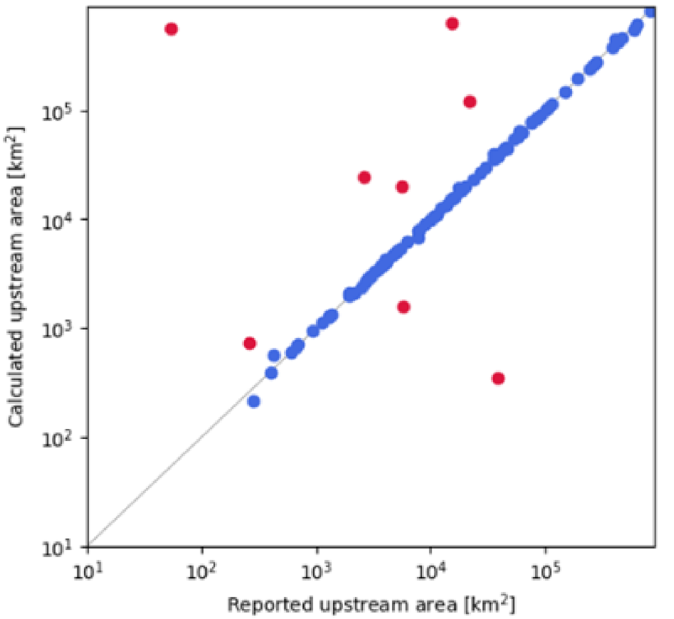
Why don't you leave off these particular systems to avoid the inaccuracy brought by the corrections?

Response: "Correction" here is not related to the dataset, but it limits the river discharge and lake outflow not to exceed the corresponding storage at the previous timestep as the original river model, CaMa-Flood. I'm sorry for the ambiguous explanation and corrected the description.

4 Validation of harmonized geographical information

Table 1: Could you indicate these eight reservoirs in Fig 3 by different colors?

Response: I updated Fig 3 with blue and red colors as you recommended:



5.1 Simulation configuration

Line 384 - 385: Could you mention this information at the beginning of the paper (maybe in section 2?)

Response: We added the number of the lakes resolved into Sect. 2.1. (The information on volume and area are described only here because Sect. 2.1 focuses on the pre-processing.)

Line 407: Where are these initial values from?

Response: The "initial guess" of the initial value for lake water level and temperature were given with the surface elevation of lake water registered in HydroLAKES (attribute name is Elevation) and the air temperature on the first day of the calculation period, respectively. If the preliminary results with those values showed long-term drift, we manually set the new initial value after trial-and-error. In addition, the initial state of the vertical layer thickness was calculated from those values of lake water level, and the initial ice volume was set to be zero. We added the description and separate the paragraph into two; one for the initial condition and the other for model configuration.

5.2 Reference data

Line 444: You can get more vertical observations via https://www.glos.us/

Response: Thank you for letting us know about the portal site of precious observation data. We manually downloaded the part of the data, but it was difficult to proceed with the discussion on the validation because the model does not resolve the spatial variability of water temperature. We would like to utilize the data when we couple (simplified) 3D hydro- and thermodynamics lake models in further study, and added the discussion in the Sect. 7 (Discussion for further development).

5.3 River discharge at downstream areas of lakes

Figure 4: Could you please adjust the y-axis of (a) to integers?

Response: We updated the y-axis of histograms to integers in Fig. 4 and 10 as you recommended.

5.5 Lake water surface elevation

Figure 9: It looks like the "lake-only" simulation simulated much higher water surface elevation than the reality and the "coupled" simulation. Are these results from 20-year spin-up time run? In line 503, you mentioned that is due to the imbalance between precipitation and evaporation. Because I see the increase rate of the elevation in the "lake-only" simulation was not quite sharp. Can you initiate the model with the observations and try the simulation with less spin-up time?

If the imbalance between precipitation and evaporation could induce such a big discrepancy, the upstream rivers must have a big backflows when you change to the "coupled" simulation.

Response: Yes, these results for the "lake-only" experiment are also the results after 20-year spinning up. "Lake-only" experiment solves only the vertical mass budget, i.e., precipitation – evaporation. Because the surface temperature of lake water and evaporative flux is mainly governed by atmospheric conditions, the vertical mass budget does not reach zero even after a very long spin-up period or any initial condition. The vertical mass budget is solved on a sub-daily scale, so the time series of water level has some seasonality and is not quite sharp, but the water level does not reach equilibrium range due to the abovementioned reason.

On the other hand, the model also considers riverine inflows and outflows, so a higher water level in lakes is consumed by the increase in outflows (increase in inflows from rivers to lakes could occur as you pointed out, but the increase in outflows is dominant). Consequently, the water level does not get too high but reaches the equilibrium range in the "coupled" simulation.

5.7 Vertical profile of lake water temperature

Line 598: Can you manually correct the lake depth? Because, especially in the Great Lakes, the incorrect lake depth may induce a big error in the thermal structure.

Response: Yes, we can correct the lake depth by editing a configuration file. We agree that the simulated thermal regime is sensitive to the lake depth. However, in the field of vertical 1D lake temperature model, there is still no consensus on how to set an appropriate lake depth (e.g. mean depth or the maximum depth) according to a model intercomparison project (Stepanenko et al., 2013). In this study, we consistently used the mean depth (attribute name is Depth_avg) contained in HydroLAKES dataset for all the lakes.

Line 582: Have you validated the ice simulation in the Great Lakes during early spring? The assumption of ice thickness in this model may affect the temperature simulation in the Great Lakes.

Response: As you recommended, we newly validated (1) ice cover period in each year and (2) monthly maximum of ice cover fraction in the Great Lakes and Lake St. Clair with dataset provided by GLERL (Assel, 2003; Wang et al., 2012). The model underestimates the period for all the lakes except for Lake St. Clair, which suggests that the vertical 1D lake model does not resolve the spatial distribution of the cover ice. Tuning of the ice shape parameterization could improve the bias as you discussed, and we think the implementation of a horizontal 2D (3D including vertical 1D) model is also a solution. On the other hand, we found that incorporation of river flows improved the underestimation of the monthly maximum of ice cover fraction in Lake Erie and St. Clair. This result suggests that cooler riverine inflow from the Northern area has an impact on the ice formation, and we confirmed that it has a similar effect on lake surface temperature. So, in Fig. 11 (d) showing the time series of surface water temperature in Lake Champlain, we replaced it with that in Lake St. Clair. We added those discussions into Appendix A1.

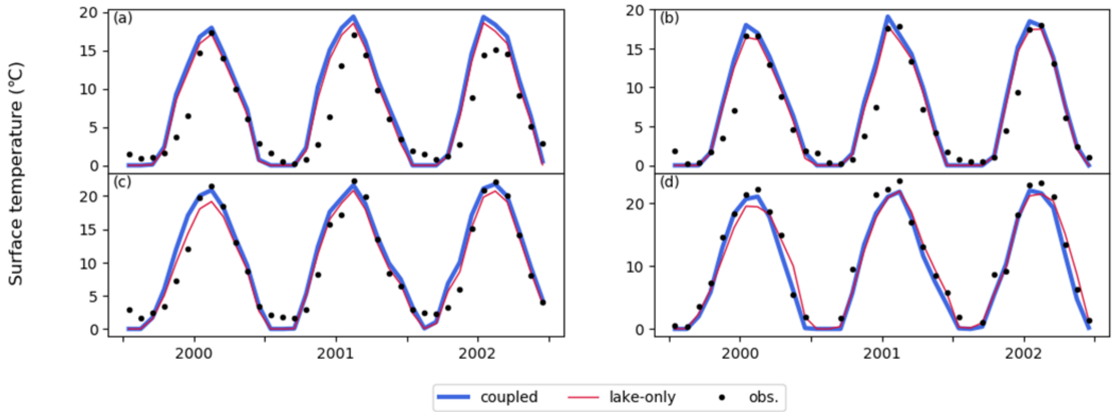


Figure 11 (d) is updated from Lake Champlain to Lake St. Clair.

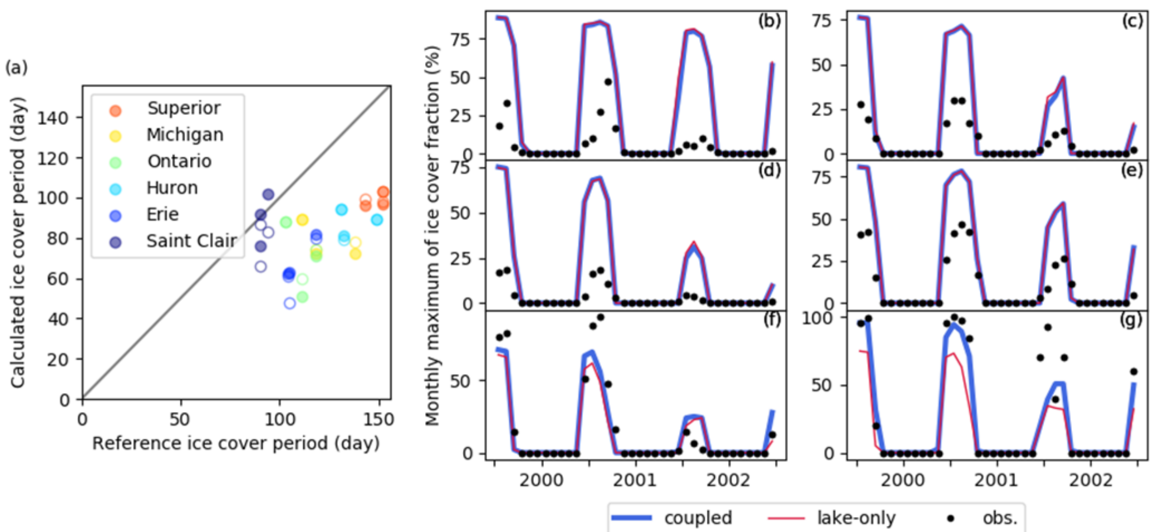


Figure A1: The comparison of (a) ice cover period in each year (day) and (b)-(g) monthly maximum of cover-ice fraction (%) in the Great Lakes region between the simulations and reference dataset. (a) The colored (white) face shows the results of the "coupled" ("lake-only") simulation. (b)-(g) Black dots show observed values, and the blue (red) line shows value simulated by the "coupled" ("lake-only") simulation. (b) Lake Superior, (c) Michigan, (d) Ontario, (e) Huron, (f) Erie, and (g) St. Clair (the HydroLAKES ID is 5, 6, 7, 8, 9, and 66, respectively).

7 sensitivity to meteorological forcing dataset

Is this section necessary in the main body of this manuscript if the different meteo forcing did not generate obvious discrepancy?

Response: We move the section to Appendix A2 as you recommended.

9 Conclusion

Line 672: Please list some metrics here to show how much the "coupled" simulation is better than "river-only" and "lake-only" simulations.

Response: We summarize the reproducibility indices in a table as you recommended.

Table 2 Summary of comparison of reproducibility indices between coupled and uncoupled (“river-only” for riverine and “lake-only” for lacustrine variables) simulations.

Variable	Statistical Index (unit)	Coupled	Uncoupled (“river-only” or “lake-only”)
	CORR (-)	0.562 (0.462)	0.482 (0.440)
River discharge	pBIAS (-)	-0.094 (0.106)	-0.080 (0.146)
	pRMSE (-)	1.009 (1.276)	1.093 (1.387)
	CORR (-)	0.894 (0.808)	0.871 (0.807)
River water temperature	BIAS (°C)	0.723 (0.753)	1.067 (0.914)
	RMSE (°C)	2.493 (2.792)	2.478 (2.814)
	CORR (-)	3.314 (0.274)	0.343 (0.289)
Lake water surface elevation	BIAS (m)	-1.594 (-0.818)	- (-)
	RMSE (m)	3.479 (6.864)	- (-)
	CORR	0.969 (0.922)	0.961 (0.928)
Lake surface temperature	BIAS (°C)	-1.165 (-1.533)	-1.364 (-1.730)
	RMSE (°C)	2.197 (2.879)	2.404 (2.955)

Technical corrections:

Figure 9: the unit of lake surface elevation should be (m) in the caption.

Response: Thank you for your kind correction. We corrected the unit in the caption.