Anonymous Referee 3

General comments

The manuscript entitled "Simulations of future changes in thermal structure of Lake Erken: Proof of concept for ISIMIP2b lake sector local simulation strategy" showed the effects of different time-scale forcing data and 4 model forcing and also the 2 RCP future scenario on the simulation with GOTM lake model over Lake Erken. It projected the similar future changing trends of thermal conditions and is helpful for local to understand the effects of climate change and adapt it.

Response: We thank the Referee 3 for the positive comments about the text. The paper was edited very carefully and modifications and improvements were made. Below, we address every comment and explain the corresponding changes in the manuscript.

Specific comments

The work focused on daily characteristics of future thermal contracture in Figure 4-6. The simulated future changing trends are mostly similar with hourly or daily forcing. But lots of work were done to compare the simulation results with different historical data which may be simplified or removed. Then the work could pay more attention to the future changing characteristics.

Response: The purpose of this paper is twofold: (1) evaluate the importance of diurnal forcing in 1D lake model and (2) assess the long-term impacts of climate change on the thermal structure of Lake Erken. Therefore, we do not consider it appropriate to simplify or remove the first purpose. The difference in mean lake conditions between the reference periods and both mid-century and late-century and long-term trend analysis has been analysed for the climate data and thermal metrics. And also the differences of each meteorological variable and thermal metric were evaluated when the lake model was forced at daily and hourly resolutions respectively.

Changes in manuscript: Material and methods: section 2.8 Statistical analysis P9 L255-279, Results: section 3.3 Climate data projections P11 L324-343, section 3.4 Long-term modelled changes in thermal stratification P11-14 L344-420 and section 3.5 Comparison between long-term thermal metrics derived from daily and hourly climate data P13 L421-428. Discussion: P16-18 L500-580.

Changes in supplement: Sections S3-S5.

L244-246 "Rates of change in whole-lake temperature calculated for over the length for RCP2.6 and 6.0 scenarios were projected to 245 increase except in the case of GFDL-ESM2M which showed weaker or non-significant changes for all measures of thermal stratification." did not match with Table 5.

Response: We do not agree with this comment. Table S8 and Table 5 show the trend analysis under RCP 2.6 and 6.0 respectively for the period 2011-2100. For RCP 2.6 the whole-lake temperature projected under GFDL-ESM2M shows a non-significant increase, and for RCP 6.0 the project increase associated with GFDL-ESM2M was the lowest of the GCMs. For RCP 6.0 the increase in whole-lake temperature ranged from 0.26 to 0.14 ^{o}C decade^{-1}.

Some parts were hardly understood, such as "For RCP 6.0, the projected rate of change ranged from 0.15 to 0.27 ^{o}C decade-1 (0.11 to 0.19 ^{o}C decade-1). IPSLCM5A- LR projected the largest increase being 0.59 ^{o}C (0.43 ^{o}C) under RCP 2.6 ^{o}C and 2.51 ^{o}C (1.79 ^{o}C) under RCP 6.0". And IPSL-CM5A-LR did not project the largest temperature increase under RCP 2.6 as showed in Table 5.

Response: We totally agree, sometimes it's hard to understand. The results have been rewritten, reducing the large amount of numbers in the text, making it more readable. All the results can be found in the Figures and Tables of both the manuscript and the supplement material. IPSL-CM5A-LR did not project the largest temperature increase under RCP 2.6, under scenario future RCP 2.6 HadGEM2-ES projected

the largest increase in surface temperature, being 0.15 ^{0}C decade^{-1}. The trend analysis has been carefully reviewed and the results rewritten.
Changes in manuscript: P9-11 L282-323, P14-15 L430-461 and P15-16 L478-499.

Because the lake model parameters are different for different forcing in Table 2. It's hard to know the source of the simulation difference in Table 4 and to evaluate the effects of the time-scale of forcing.
Response: One of the purposes of this study was to test the ability of a 1D lake model (GOTM) to simulate daily water temperature using daily vs hourly meteorological data, i.e. evaluate the importance of diurnal forcing in 1D lake model. In all cases the lake model was ran at hourly model computational time step when the meteorological forcing was provided at either daily or hourly frequencies. In each case a separate calibration was run using the same observed data for comparison, simulated output derived from the models forced at daily and hourly resolution. We felt that this was the fairest and most representative way to test how the model would actually be applied with the different forcing data. When GOTM was forced at daily resolutions, there is no diurnal variability in the input, which leads to changes in heat fluxes. However it became apparent that variations in model parameters resulting from the different calibrations compensated for some of the differences between observations and simulations based on the different time-scale of forcing. We now point this out more clearly in the paper.
Changes in the manuscript: P17 L530-540.

L230 "From these average yearly values were calculated using the months between April and September, due to the fact that the GOTM model was not able to simulate lake ice and winter water temperatures at the same level of accuracy as during the remainder of the year". Does the inaccurate simulation of lake temperature in winter affect the temperature simulation without ice? L68 "The lake is dimictic with summer stratification usually occurring beginning in May-June and ending in August-September, while ice cover occurs from December-February to April-May." Why the average yearly values were calculated including April?
Response: The GOTM model version 5.1 did not have the ability to simulate lake ice, so for this study the inverse stratification period was not analysed. Moras et al., (2019), has shown that despite this limitation, the mode is able to accurately simulate water temperature and the phenology of thermal stratification during the remainder of the year. A new GOTM model version 5.4 with ice-module was released after this project was submitted, allowing to evaluate the effect of the lack of ice module on the onset of the direct stratification. The onset of direct stratification was derived from simulations of water temperature with GOTM version 5.1 and 5.4 from 2006 to 2016. The RMSE between the onset of direct stratification from GOTM version 5.1 and 5.4 was 5.22 days showing a slight impact the lack of ice-module on the onset of the direct stratification.

onset of direct stratification	
GOTM v5.1	GOTM v5.4
2007-04-27	2007-04-16
2008-04-27	2008-04-26
2009-04-27	2009-04-27
2010-05-01	2010-05-13
2011-04-24	2011-04-25
2012-05-03	2012-05-01
2013-05-09	2013-05-09
2014-04-21	2014-04-21
2015-05-22	2015-05-22

2016-05-04	2016-05-03

Annual ice cover observations of the onset and loss of ice cover made at lake Erken since 1941 (Moras et al., 2019) showed a decreased since 1941 by 7.34 day decade^{-1} (57 days from 1941 to 2017), consistent with changes in air temperature. For this reason, we consider relevant in our long-term study to include April in our analysis.

The manuscript was submitted in 2019. It's confused to compare 2006-2099 with 1975-2005 to get the future change.

Response: we totally agree, the choice of reference period is always controversial because the projected impact depends on it. Initially we used as a reference period the last 30 years of the historical scenario (1975-2005) for each GCM, since from 2006 they were already future projections. However, we have decided to slightly update our reference period to 1981-2010.

The table shows the trend analysis for the period 2006-2100 relative to 1975-2005 and for the period 2011-2100 relative to 1981-2010 for HadGEM2-ES under RCP 6.0. The differences are almost unnoticeable, so we do not consider it necessary to update our reference period to 1990-2019.

	HadGEM2-ES RCP 6.0			
	reference period: 1975-2005		reference period: 1981-2010	
	24h met	1h met	24h met	1h met
air temperature (^{o}C)	0.44 ^{o}C dec^{-1}	0.33 ^{o}C dec^{-1}	0.43 ^{o}C dec^{-1}	0.32 ^{o}C dec^{-1}
surface temperature (^{o}C)	0.38 ^{o}C dec^{-1}	0.28 ^{o}C dec^{-1}	0.38 ^{o}C de-1	0.27 ^{o}C dec^{-1}
bottom temperature (^{o}C)	0.07 ^{o}C dec^{-1}	ns	0.06 ^{o}C dec-1	ns
whole-lake temperature (^{o}C)	0.25 ^{o}C dec^{-1}	0.17 ^{o}C dec^{-1}	0.25 ^{o}C dec^{-1}	0.16 ^{o}C dec^{-1}
Schmidt stability (J m^{-2})	7.79 J m^{-2} dec^{-1}	6.22 J m^{-2} dec^{-1}	7.97 J m^{-2} dec^{-1}	6.50 J m^{-2} dec^{-1}
thermocline depth (m)	0.12 m dec^{-1}	0.12 m dec-^{1}	0.13 m dec^{-1}	0.13 m dec^{-1}

Does the lake model need downward longwave radiation drive? What's the usage of the cloud cover when there is the downward shortwave radiation?

GOTM internally calculates net long-wave radiation from cloud cover according to Clark et al. (1974). Cloud cover for long-term water temperature simulations was estimated from bias-corrected model data according to Martin and McCutcheon (1999):

$$H_{SW} = H_0 \cdot a_t \cdot (1 - R_s) \cdot C_a$$

where H_{sw} is the short-wave solar radiation (W · m^{-2}), H_0 is the amount of radiation reaching the earth's outer atmosphere (W · m^{-2}), a_t is an atmospheric transmission term, R_s albedo or reflection coefficient, and C_a is the fraction of solar radiation not absorbed by clouds.

$$C_a = 1 - 0.65 \cdot C_l^2$$

where C_l is the fraction of the sky covered by clouds.

Cloud cover would be:

$$C_l = \sqrt{\frac{1 - \frac{H_{SW}}{H_0 \cdot a_t \cdot (1 - -R_s)}}{0.65}}$$

Usually the simulation in the calibration period is better. Why temperature simulations in the validation period were more accurate in the manuscript?

Response: Water temperature simulations were apparently more accurate for the validation period (2015-2016) than for the calibration period (2006-2014), which may appear unusual, but is due to the higher variability in observed water temperature during the longer calibration period. Years with a longer duration of stratification and stronger stability, generally had higher simulation errors. Half of the eight-year calibration period exhibited these conditions, while the two-years used for validation both exhibited shorter duration of stratification and weaker stability.

	year	RMSE (°C)			thermal stratification			Schmidt
		24h met	1h met	synthetic 1h met	duration (days)	onset	loss	stability (J m^{-2})
Calibration	2007	0.58	0.59	0.83	23	176	230	17.42
	2008	**1.42**	**1.13**	**1.04**	**103**	**124**	**227**	**31.52**
	2009	0.75	0.68	0.63	69	122	242	35.17
	2010	**1.10**	**0.92**	**0.99**	**111**	**139**	**254**	**80.77**
	2011	**0.92**	**0.79**	**0.81**	**90**	**152**	**252**	**43.77**
	2012	0.71	0.66	0.77	38	141	244	32.98
	2013	**1.42**	**1.52**	**1.08**	**124**	**129**	**259**	**79.48**
	2014	0.83	0.73	0.79	55	137	263	52.40
Validation	2015	0.59	0.66	0.65	71	162	240	17.60
	2016	0.69	0.73	0.71	67	173	239	47.25

Changes in manuscript: P14 L433-438.

L 110 "under four emission scenarios". As shown in the manuscript, there were only 2 emission scenarios.
Response: Change made.

If the years for calibration and validation match the years for training and validating, it may be better.
The GRNN training and validation periods do not fit into GOTM calibration periods. Putting these periods in context does not produce significant changes in the GRNN models performance (see table below) but it would entail a high computational cost since changing the GRNN models would require all the GCM scenarios to be disaggregated a second time and all GCM scenarios to be run again using these alternative data.

	Air temperature (°C)			Relative humidity (%)			Wind speed (m s^{-1})			Short-wave rad (W m^{-2})		
	BIAS	RMSE	NSE	BIAS	RMSE	NSE	BIAS	RMSE	NSE	BIAS	RMSE	NSE
Training: 2006-2014	0.00	0.32	1.00	0.00	0.96	1.00	-0.01	1.15	0.74	0.00	8.39	1.00
Validation: 2015-2016	0.03	0.70	0.95	0.44	2.09	0.69	-0.07	2.50	0.60	0.08	18.15	0.86
Training: 2008-2012	0.00	0.26	1.00	0.00	0.79	1.00	-0.01	1.06	0.78	0.00	6.35	1.00
Validation: 2013-2015	-0.06	0.32	0.94	0.34	1.02	0.69	-0.01	1.37	0.58	-0.04	8.20	0.87

References:

Martin, J., and McCutcheon, M.: Hydrodynamics and Transport for Water Quality Modeling, Lewis Publishers, US, 1999.

Clark, N. E., Eber, L., Laurs, R. M., Renner, J. A., and Saur, J. F. T.: Heat exchange between ocean and atmosphere in the Eastern North Pacific for 1961-71, NOAA Technical Report NMFS SSRF-682, U. S. Dept. of Commerce, Washington, DC, 72 pp., 1974.