Anonymous Referee 1

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" aims to assess the impacts global warming on the thermal characteristics of Lake Erken. Since future projections of global warming are at a daily time step, the authors first analyses the need to disaggregate the input variables to the hourly time step. The manuscript is well written. The topic is scientifically relevant.

Response: We thank the Referee 1 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

Line 70-71: "It is the lake's relatively shallow depth and large surface area, which leads to large inter-annual variability in the timing and patterns of thermal stratification." Why is this? Perhaps explain in one or two sentences how this works and why this is different for relatively deep lakes or lakes with a small surface area.

Response: Mixing and stratification change in response to lake morphometry. Shallow lakes have lower heat storage, responding more directly to short-term variations in the weather conditions and heat can be transferred through the water column by wind mixing (Magee and Wu, 2017). However, deep lakes required greater wind speeds to complete the mix. Large surface areas or fetch increase the effects of mixing and vertical transfer of heat to the bottom (Rueda and Schladow, 2009).

Changes in manuscript: P19 L74-76.

Line 138-139: "More detailed description of the GRNN methods and models are given in the supplementary material to this paper." I was hoping to find equations on how the GRNN model calculates hourly estimations based on daily input, however, I could not find a detailed description of the GRNN methods in the supplementary materials.

Response: GRNN description was added in the supplement material section S1.

Changes in supplement: Section S1.

Line 158: "Schmidt stability", perhaps give a definition or equation of the Schmidt stability.

Response: The following Schmidt stability definition was added: resistance to mechanical mixing due to the potential energy inherent in the density stratification of the water column (Schmidt, 1928; Idso, 1973).

Changes in manuscript: P22 L190-191.

Line 170-174: "Air temperature, short-wave radiation, relative humidity and wind speed were temporarily disaggregated into hourly values from mean daily data, using the GRNN models. A database was constructed using 8 years of measurements. From this whole set of data, the first 5-years of data, that is, from 2008 to 2012, were used for

training, and 3-years of data from 2013 to 2015 were used for validating the results obtained." This sentence was confusing. After reading the methods section I first assumed this was about the calibration/validation of GOTM. However, later I realized it was about the calibration/validation of GRNN. I would expect these sentences in the methods section. Moreover, it would be good to mention clearly that there are two types of calibration/validation: that of GOTM and that of GRNN.

Response*:* It has been moved from Results section 3.1. Hourly meteorological modelling to Material and Methods section 2.5. Temporal disaggregation of meteorological forcing data.

Changes in manuscript: P21 L147-154.

Line 192-193: "Temperature simulations for the validation period were more accurate (average RMSE of 0.66 ºC and NSE of 0.97) than for the calibration period (average RMSE of 0.95 ºC and NSE of 0.94), but in both periods the model performance was considered acceptable." I would expect that the validation period would be less accurate than the calibration period. Therefore, my first thought was then that perhaps the legend was swapped between calibration and validation. Yet, the authors later mention that this is "due the higher variability in observed water temperature during the long calibration period." (Line 284-285). Then the question raises, which data set is more representative? Was the high variability during the calibration period actually quite normal and the validation period exceptionally uniform? And what does this mean for the validity of the output?

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 (ºC)			thermal stratification			Schmidt stability (J m^{-2})
		24h met	1h met	synthetic 1h met	duration (days)	onset	loss	
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: P30 L432-437.

Line 202: “As would be expected the simulations of bottom temperature were slightly less accurate” Why would this be expected?

Response: Higher errors were found at the lowest depth (15 m depth), part of this might have been caused by the presence of internal seiches in lake Erken which cannot be reproduced by 1D models such as GOTM.

Line 349-351: “Combined these results suggest important changes in the factors affecting lake biogeochemistry directly through changes in temperature and indirectly by influencing the availability of light and nutrients.” The presented results only indicate an increase in temperature and stratification period. Since the presented data does not show how this affects biogeochemistry and the availability of light and nutrients, could the authors be a bit more specific on this in the conclusion? How do the authors think/speculate it will change (perhaps refer to the introduction where a short explanation is already given)?

Response: As mentioned in the introduction the projected changes in thermal stratification can influence many aspects of the lake ecosystem. Increases in thermal stability and duration of stratification can intensify hypolimnetic oxygen depletion (Foley et al., 2012; Schwefel et al., 2016) and hence induce enhanced internal phosphorous loading (North et al., 2014), increase the release of dissolved iron and manganese from sediments (Schultze et al., 2017) and also increase methane emissions (Grasset et al., 2018). Warming lake temperature affects biological rates of metabolism, growth and reproduction (Rall et al., 2012) and can promote cyanobacterial blooms (Paerl and Paul, 2012). When coupled to a *reduction* in *oxygen*-rich water, warming water temperature leads to a *lower fish* populations (O'Reilly et al., 2003; Yankova et al., 2017). Increase in evaporation associated with warming can lead to declines in lake water level (Hanrahan et al., 2010) with implications for water security. So these changes are expected in Lake Erken.
The expected changes in the lake ecosystem caused by changes in thermal stratification have been moved from 1. Introduction to 4. Discussion section and so our conclusions can be more understandable.

Changes in manuscript: P34-35 L580-587.

Technical corrections

Figures in general; 1) it would be good to have comparable axes per figure. For example, figure 3a has a y-axis going from 0-1.2 ºC, while figure 3b goes from 0-0.8 ºC. I suggest that the authors uniform the axis and perhaps use the normalized RMSE to compare the different subfigures 2).From the figures caption, it is not always clear if the predicted output is with GOTM or with GRNN. Perhaps include this information in the figure’s caption. General: sometimes I read “wind_factor” and sometimes “wind factor” without “_”. Is there a difference in meaning?

Response:

Figure 3 has been removed because GOTM model performance had been shown in twice (Figure 3 and Table 4).
GRNN and GOTM has been added to the figure captions to indicate if models have been used to disaggregate meteorological forcing data or to simulated water temperature.
Wind factor is the meaning of the parameter wind_factor.

Line 246-247: “Simulated changes were generally slight less for the simulations driven by daily forcing data as shown by the figures in parentheses”. Put a dot after parentheses and change “slight” to “slightly”

Response: Change made.

Line 284: "were more accurate than for the calibration period (2006-2014) due the higher variability in observed water temperature" add "to" after "due".

Response: change made.

Figure 2: 1) the caption says that validation is figure 2a, 2c, 2e and 2g, however, the title of the figures suggest that validation is figure 2b, 2d, 2f and 2h. This is confusing. 2) Perhaps include the words "observations", "daily data", "hourly data", "synthetic hourly data" on the left side of/ or under the figure. It is now quite a puzzle to find which subfigure tells what. 3) Perhaps also include a difference graph where the difference between "observations" and respectively "daily data", "hourly data", "synthetic hourly data" is shown. From figure 2, it is now hard to see the differences. (The same holds for figure 4, where it is hard to see the differences between historical and the rcp's)

Response:

Figure 2 has been renumbered and a subtitle added to each subfigure. **Changes in manuscript:** P54.

Figure S7 has been added to the supplementary material showing the differences between simulated (when the lake model was forced with daily, hourly and synthetic hourly meteorological forcing data) and observed water temperature. **Changes in supplement:** Section S2.

Figure S14 has been added to the supplementary material showing the differences between the historical and RCP 2.6 scenarios, and the historical and RCP 6.0 scenarios for the IPSL-CM5A-LR projection (when the lake model was forced at daily resolutions). **Changes in supplement:** Section S4.

Figure 5 and 6: In figure 5i, the authors indicate the words "deeper" and "shallower" with arrows. This really increases the readability of that specific subfigure and the same would help the reader in all other subfigures.

Response: In Figures 6 and 7 (in the latest version of the manuscript), the arrow and the words "deeper" and "shallower" have been removed for easy viewing of the figure. However, in the figure caption was added the meaning of values greater or less than 0 of each of the thermal indices: changes in thermal metrics greater than 0 show an increase and lower than 0 show a decrease.

Changes in manuscript: P62-63.

Changes in supplement: Sections S3-S4.

References:

Idso, S. B.: On the concept of lake stability, Limnol. Oceanogr., 18, 681–683, 1973.

Magee, M. R., and Wu, C. H.: Response of water temperatures and stratification to changing climate in three lakes with different morphometry, Hydrol. Earth Syst. Sci., 21, 6253-6274, https://doi.org/10.5194/hess-21-6253-2017, 2017.

Rueda, F., and Schladow, G.: Mixing and stratification in lakes of varying horizontal length scales: Scaling arguments and energy partitioning, Limnol. Oceanogr., 54, 2003-2017, https://doi.org/10.4319/lo.2009.54.6.2003, 2009.

Schmidt, W.: Über Temperatur und Stabilitätsverhaltnisse von Seen, Geogr. Ann., 10, 145–177, 1928

Anonymous Referee 2

General comments

The article "Simulations of future changes in thermal structure of Lake Erken: Proof of concept for ISIMIP2b lake sector local simulation strategy" presents impacts of changing climate on lake water temperature. The article is of very high interest and very well written, the work is thoroughly executed and discussion is relevant. The authors used a hydrodynamic lake model GOTM with 4 GCM/RCMs, using Generalized Regression Artificial Neural Network to disaggregate daily climate into hourly data. The GOTM model was able to reproduce observed lake temperature data for current time period (8 years). The model was then executed with climate forcing data from 4 GCM/RCMs.

Response: We thank the Referee 2 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

I would recommend expanding on the Methods section to provide more information that is critical in understanding the study, its aims, and results. It is unclear why the authors chose to consider 2006-2099 as the future even though the period begins 13 years ago. It is also unclear why this full period is evaluated without any consideration of the changes that occur from 2006 to 2099 based on the trend analyses also included in the manuscript. It seems that changes that can occur during this "future" period are considered representative of changes that will occur by 2099. The averages from this 94-year period are compared to averages from a 30-year period of 1975-2005. The variability during a 30-year period and during a 94-year period with a significant trend is expected to differ and this affects the projected changes.

A more typical approach in many climate impact studies is to select two 30-year periods, one that represents a current climate (reference period, e.g. 1981-2010) and one that represents a future climate (e.g., mid-century 2041-2010 or late century 2071-2100). Forcing data from the same climate model would then be used as model inputs for both time periods; the difference between these results would represent the projected impact. It is also not clear from the manuscript how were the reference period values calculated for calculation of anomalies from the respective GCM/RCMs during the reference time period.

The results for the mid-century and late century should be added to the manuscript to evaluate how the change progresses; alternatively, the current results can be replaced with the late century period as that seem to be the focus of the "proof of concept" study.

It is also important to include information on the variability of the simulated thermal indices due to the climate model selection, i.e. present information for all 4 GCM/RCMs for the reference time period. That can give indication to the significance of the projected impact.

Response:

Climate impact studies can be approached in two ways: (1) assessing the difference in mean lake conditions (for example, mean surface temperature) between the reference periods and both mid-century and late-century (Woolway and Merchant, 2019) or (2) long-term trend analysis (O'Reilly et al., 2015; Shatwell et al, 2019; Moras et al., 2019). The use of anomalies or absolute values in trend analysis does not change the value of the slope. The use of anomalies in frequency distribution figures provide an alternative method of comparing the changes simulated by the future climate scenarios. So we consider it appropriate to combine both approaches. We have also added information about the reference period and the same analysis has been made to the meteorological variables in order to understand the variability of the projected thermal metrics derived from GCMs.

Changes in manuscript: Material and methods: section 2.8 Statistical analysis P25 L255-278, Results: section 3.3 Climate data projections P27 L323-342, section 3.4 Long-term modelled changes in thermal stratification P27-30 L343-4219 and section 3.5 Comparison between long-term thermal metrics derived from daily and hourly climate data P30 L420-427. Discussion: P32-34 L499-579.

Changes in supplement: Sections S3-S5.

The information on the GOTM model for Lake Erken is very limited; the methods section should be expanded to include more details on the model structure, e.g. vertical resolution, inflow and outflow from the lake, etc.

Response: The GOTM model version 5.1 was used in this study. The meteorological parameters for running the model were air temperature (ºC), wind speed (m s^{-1}), short-wave radiation (W m^{-2}), cloud cover (dimensionless, 0-1), relative humidity (%), atmospheric pressure (hPa) and precipitation (mm day^{-1}or mm $hour^{-1}$). Inflows and outflows were not included in this study, and water level was considered fixed in the simulations. This version of GOTM 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. The initial conditions for water temperature were derived from a measured vertical profile. GOTM was run at hourly model computational time step, and simulated water temperature was saved as daily mean values each 0.5 m (42 layers).

Changes in manuscript: P21-22 L155-163.

Some relevant parts should be moved from Results to Methods (e.g., the beginning of section 3.1 and 3.3. Also, the periods used in the calibration (training) & validation periods for GOTM and GRNN should be put into the context between these two models. It is not readily apparent from the manuscript.

Response:

The beginning of sections 3.1 and 3.3 have been moved to 2. Materials and Methods section 2.5 Temporal disaggregation of meteorological forcing data 2.8 Statistical analysis respectively.

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 (ºC)			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

Changes in manuscript: P21 L147-154 and P25 L255-263.

It should be emphasized that training the temperature disaggregation algorithm on the current diurnal patterns means those current patterns will be projected to the future time series and any potential changes in diurnal pattern from the changing climate are ignored.

Response: GRNNs proved to be an effective method to disaggregate daily GCM forcing to an hourly temporal resolution for different weather variables such as air temperature, short-wave radiation, etc. However, GRNNs require a training phase, in which the diurnal patterns to be learned are presented to the network from historical meteorological measurements, and therefore if there are future changes in diurnal patterns, these cannot be reproduced. In addition, there is a high computational cost of disaggregating and storing the long-term daily climate data into an hourly data set.

Changes in manuscript: P33 L539-544.

Technical comments

Increases are given to 0.01 ºC – what is the accuracy of the measurement and of the simulations? Is this accuracy adequate?

Response: The accuracy of thermocouple sensor is approximately ± 0.1 ºC and can at times be somewhat better than 0.1. The simulated water temperature is given with 7 decimals. So two decimal places in the GOTM model performance are adequate, and match the best expected performance of our monitoring data.

L 68 –dimictic?

Response: Change made.

L89 Mean sea level (,) pressure (,) relative humidity and precipitation were measured – missing commas?

Response: Change made.

Section 2.6. It would be useful to include model performance for other thermal indices used for evaluation of change, e.g. duration of thermal stratification

Response:

Response: The model performance for the duration, onset and loss of stratification has been added to the section 3.2 Lake Model performance (Table 4).

Changes in manuscript: P26-27 L 318-322.

L 162: Schmidt's stability – needs a reference/ brief explanation

Response: The following Schmidt stability definition was added: resistance to mechanical mixing due to the potential energy inherent in the density stratification of the water column (Schmidt, 1928; Idso, 1973).

Changes in manuscript: P6 L190-191 and P44.

L 231: this model handicap and any other should be described in section 3.2

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

Changes in manuscript: P22 L159-162.

L 312-314 However, the dominant wind (is) along : : : missing word?

Response: Change made.

L 322, 324 When GOTM was forcing with : : : forced?

Response: Change made.

L 350: it would be good to put the statement into context; what kind of changes can be expected with these increases in temperature?

Response: The expected changes in the lake ecosystem caused by an increase in water temperature have been moved from 1. Introduction to 4. Discussion section.

Changes in manuscript: P34-35 L580-587.

Figure 2 heading: figure shows calibration as plots a, c, e, and g, but the caption says these are validations. I would recommend including a similar plot but with model residuals (perhaps in Supplementary materials); that would make any differences much easier to see especially on the timing.

Response:

Figure 2 has been renumbered and a subtitle added to each subfigure. **Changes in manuscript:** P54.

Figure S7 has been added to the supplementary material showing the differences between simulated (when the lake model was forced with daily, hourly and synthetic hourly meteorological forcing data) and observed water temperature. **Changes in supplement:** Section S2.

Figure 3: it would be helpful if the scale on y axis with the same units had the same range (a-d)

Figure 3 has been removed because GOTM model performance had been shown in twice (Figure 3 and Table 4).

References:

Idso, S. B.: On the concept of lake stability, Limnol. Oceanogr., 18, 681–683, 1973.

Moras, S., Ayala, A. I., and Pierson, D. C.: Historical modelling of changes in Lake Erken thermal conditions, Hydrol. Earth Syst. Sci., 23, 5001–5016, https://doi.org/10.5194/hess-23-5001-2019, 2019.

O'Reilly, C., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowle,y R. J., Schneider, P., Lenters, J. D., McIntyre, P.B., Kraemer, B. M., Weyhenmeyer, G. A., Straile, D., Dong, B., Adrian, R., Allan, M. G., Anneville, O., Arvola, L., Austin, J., Bailey, J. L., Baron, J. S., Brookes, J. D., de Eyto, E., Dokulil, M. T., Hamilton, D. P., Havens, K., Hetherington, A. L., Higgins, S. N., Hook, S., Izmest'eva, L. R., Joehnk, K. D., Kangur, K., Kasprzal, P., Kumagai, M., Kuusisto, E., Leshkevich, 20 G., Livingtone, D. M., McIntyre, S., May, L., Melack, J. M., Mueller-Navarra, D. C, Naumenko, M., Noges, P., Noges, T., North, R. P., Plisnier, P. D., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L. G., Rusak, J. A., Salmaso, N., Samal, N. R., Schindler, D. E., Schladow, S. G., Schmid, M., Schmidt, S. R., Silow, E., Soylu, M. E., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Wiliamson, C. E., and Zhang G.: Rapid and highly variable warming of lake surface waters around the globe, Geophys. Res. Lett., 42, 10773–10781, https://doi.org/10.1002/2015GL066235, 2015.

Schmidt, W.: Über Temperatur und Stabilitätsverhaltnisse von Seen, Geogr. Ann., 10, 145–177, 1928.

Shatwell, T., Thiery, W., and Kirillin, G.: Future projections of temperature and mixing regime of European temperate lakes, Hydrol. Earth Syst. Sci., 23, 1533–1551, https://doi.org/10.5194/hess–23–1533–2019, 2019.

Woolway, R.I., and Merchant, C.J.: Worldwide alteration of lake mixing regimes in response to climate change, Nature Geoscience, 12, 271–276, https://doi.org/10.1038/s41561–019–0322–x, 2019

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 P25 L255-278, Results: section 3.3 Climate data projections P27 L323-342, section 3.4 Long-term modelled changes in thermal stratification P27-30 L343-4219 and section 3.5 Comparison between long-term thermal metrics derived from daily and hourly climate data P30 L420-427. Discussion: P32-34 L499-579.

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 ºC 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 ºC decade-1 (0.11 to 0.19 ºC decade-1). IPSLCM5A- LR projected the largest increase being 0.59 ºC (0.43 ºC) under RCP 2.6 ºC and 2.51 ºC (1.79 ºC) 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 ºC decade^{-1}. The trend analysis has been carefully reviewed and the results rewritten.

Changes in manuscript: P25-27 L281-322, P30-32 L429-498.

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: P33 L529-539.

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 (ºC)	0.44 ºC dec^{-1}	0.33 ºC dec^{-1}	0.43 ºC dec^{-1}	0.32 ºC dec^{-1}
surface temperature (ºC)	0.38 ºC dec^{-1}	0.28 ºC dec^{-1}	0.38 ºC de-1	0.27 ºC dec^{-1}
bottom temperature (ºC)	0.07 ºC dec^{-1}	ns	0.06 ºC dec-1	ns
whole-lake temperature (ºC)	0.25 ºC dec^{-1}	0.17 ºC dec^{-1}	0.25 ºC dec^{-1}	0.16 ºC 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 (ºC)			thermal stratification			Schmidt stability (J m^{-2})
		24h met	1h met	synthetic 1h met	duration (days)	onset	loss	
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: P30 L432-437.

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 (ºC)			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.

Simulations of future changes in thermal structure of Lake Erken: Proof of concept for ISIMIP2b lake sector local simulation strategy

Ana I. Ayala[1, 2], Simone Moras[1], ~~Don~~Donald C. Pierson[1]

[1]Department of Ecology and Genetics, Limnology, Uppsala University, Uppsala, 752 36, Sweden
[2]Department of Applied Physics, Nonlinearity and Climate Group, University of Geneva, Geneva, CH-1211, Switzerland

Correspondence to: Ana I. Ayala (isabel.ayala.zamora@ebc.uu.se)

Abstract. This paper, as a part of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), assesses the impacts of different levels of global warming on the thermal structure of Lake Erken (Sweden). The GOTM one-dimensional hydrodynamic model was used to simulate water temperature when using ISIMIP2b bias-corrected climate model projections as input. These projections have a daily time step, while lake model simulations are often forced at hourly or shorter time steps. Therefore, it was necessary to first test the ability of GOTM to simulate Lake Erken water temperature using daily vs hourly meteorological forcing data. In order to do this three data sets were used to force the model: 1) hourly measured data; 2) daily average data derived from the first data set and; 3) synthetic hourly data created from the daily data set using Generalized Regression Artificial Neural Network methods. This last data set is developed using a method that could also be applied to the daily time step ISIMIP scenarios to obtain hourly model input if needed. The lake model was shown to accurately simulate Lake Erken water temperature when forced with either daily or synthetic hourly data. Long-term simulations forced with daily or synthetic hourly meteorological data suggest that by ~~2099~~late 21st century the lake will undergo clear changes in thermal structure~~, for~~. For RCP 2.6 surface water temperature was projected to increase ~~from 0.87 to~~by 1.~~48~~ 79 ºC and ~~from 0.69 to~~ 1.~~20~~36 ºC when the lake model was forced at daily and hourly resolutions respectively, and for RCP 6.0 these increases were projected to ~~range from 1.58 to~~ be 3.~~58~~08 ºC and ~~from 1.19 to~~ 2.~~65~~31 ºC ~~when the lake model was also forced at daily and hourly resolutions.~~. Changes in lake stability were projected to increase ~~significantly~~ and the stratification duration was projected to be longer by ~~9 to 16 days and from 7 to~~ 13 days and 11 days under RCP 2.6 scenario and ~~from 20 to 33~~22 days and ~~from 17 to 27~~18 under RCP 6.0 scenario for daily and hourly resolutions. Model ~~trends~~changes in thermal indices were very similar when using either the daily or synthetic hourly forcing, suggesting that the original ISIMIP climate model projections at a daily time step can be sufficient for the purpose of simulating lake water temperature ~~in the lake sector in ISIMIP~~.

1 Introduction

The thermal structure of lakes is controlled by heat and energy exchange across the air-water interface, which is in turn determined by meteorological forcing (Woolway et al., 2017). Climate change will affect air-water energy exchanges and alter

the temperature regime and mixing of lakes (~~Mesman et al.,~~Woolway and Merchant, 2019~~ submitted~~). For example, increases in air temperature results in a consequent warming of lake water temperature (Sahoo et al. 2015) causing shorter ice-cover periods (Kainz et al., 2017; Butcher et al. 2015), longer stratified period (Ficker et al., 2017; Woolway et al., 2017; Magee and Wu, 2017) and increased lake stability (Rempfer et al., 2010; Hadley et al., 2014). Decreasing wind speed can induce more stable and long-lasting stratification ~~(Woolway et al. 2017)~~ and increased epilimnetic temperature (~~Stefan~~Woolway et al. 2017; Woolway et al., ~~1996).~~ 2019).

The most direct effect of climate change on lakes is a warming of the lake surface temperature. For example, global average warming rates of 0.34°C decade^{-1} have been observed between 1985 and 2009 by O'Reilly et al. (2015). Hypolimnetic temperature responds less clearly to warming and has been observed to be warming, cooling or not changing significantly with increasing air temperature (Shimoda et al., 2011; Butcher et al., 2015; Winslow et al. 2017). And, these changing water temperatures have also led to an increased stability and duration of stratification (Butcher et al., 2015; Kraemer et al., 2015). A final consequence of warming lake temperature is projected to be the shift in the mixing regime (~~Krillin~~Kirillin, 2010; Shimoda et al., 2011; Shatwell et al., 2019; Woolway and Merchant, 2019). For example, loss of ice cover in deep lakes is likely to turn amictic lakes into cold monomictic lakes, and cold monomictic lakes into dimictic lakes (Nõges et al., 2009).

These changes in lake water temperature and thermal stratification influence lake ecosystem dynamics (MacKay et al., 2009). ~~Increases in stratification stability and duration can intensify hypolimnetic oxygen depletion (Foley et al., 2012; Schwefel et al., 2016) and hence induce enhanced internal phosphorous loading (North et al., 2014), increase the release of dissolved iron and manganese from sediments (Schultze et al., 2017) and also increase methane emissions (Grasset et al., 2018). Warming lake temperature affects biological rates of metabolism, growth and reproduction (Rall et al., 2012) and can promote cyanobacterial blooms (Paerl and Paul, 2012). When coupled to a reduction in oxygen rich water, warming water temperature leads to a lower fish populations (O'Reilly et al., 2003; Yankova et al., 2017).~~

Numerical modeling plays a key role in estimating the sensitivity of the lakes to changes in the climate. One-dimensional lake models are widely used due to their computational efficiency and the realistic temperature profiles they produce. Several studies have investigated the impacts of climate change on lake water temperature under Regional Climatic Model (RCM)/Global Climatic Model (GCM) projections (Persson et al., 2005; Kirillin, 2010; Perroud and Goyette, 2010; Samal et al., 2012; Ladwig et al., 2018; Shatwell et al., 2019; Woolway and Merchant, 2019). Commonly when undertaking climate change impact studies, hydrodynamic lake models are driven by daily resolution RCM/GCM outputs. Bruce et al. (2018) undertook a comparative analysis of model performance using daily and hourly resolution meteorological forcing data, and found a better agreement between observations and predictions of full-profile temperature when the lakes were modelled using hourly meteorological input. This reinforces the importance of diurnal forcing on 1-D model predictive capability.

The purpose of this study is therefore (1) to test the ability of a one dimensional-hydrodynamic model (GOTM) to simulate the water temperature of Lake Erken (Sweden) using daily vs hourly meteorological forcing data for the period 2006-2016, (2) develop a reliable method to disaggregate daily meteorological data to a hourly synthetic product that can be used to force

the GOTM model and ~~(3)~~ convert the daily GCM outputs available from the ISIMIP ~~project~~ into hourly meteorological ~~variables~~data sets and (3) assess the impacts on the thermal structure of Lake Erken at different levels of global warming when GOTM is driven by hourly and daily model projections. In fulfilling these objectives this study provides the first evaluation of modelling methods that will be used by the lake sector within the ISIMIP.

2 Material and Methods

2.1 Study site

Lake Erken (59°51'N, 18°36'E) is a mesotrophic lake located in east central Sweden, with a maximum depth of 21 m, a mean depth of 9 m and a surface area of 23.7 km^2. The lake is ~~dimictive~~dimictic with summer stratification usually occurring beginning in May-June and ending in August-September, while the onset of ice cover occurs ~~from~~between December-February ~~to~~and ice loss is in April-May (Persson and Jones, 2008). It is the lake's relatively shallow depth and large surface area, which leads to large inter-annual variability in the timing and patterns of thermal stratification~~.~~, since heat can be readily transferred through the shallow water column by wind mixing (Magee and Wu, 2017), and since the lake has a relatively low heat storage, and therefore, responds more directly to short-term variations in weather. The lake has a retention time of approximately 7 years and shows annual variations in water level that are less the 1 m (Pierson et al., 1992; Moras et al., 2019 ~~in review~~).

2.2 Lake model

General Ocean Turbulence Model (GOTM) is a one dimensional water column model that simulates the most important hydrodynamic and thermodynamics processes related to vertical mixing in natural waters (Umlauf et al. 2005). GOTM was developed by Burchard et al. (1999) for modelling turbulence in the oceans, but it has been recently adapted for use in hydrodynamic modelling of lakes (Sachse et al., 2014). The strength of GOTM is the vast number of well-tested turbulence models that have been implemented spanning from simple prescribed expressions for the turbulent diffusivities up to complex Reynolds-stress models with several differential transport equations. Typically GOTM is used as a stand-alone model for investigating the dynamics of boundary layers in natural waters but it can also be coupled to a biogeochemical model using the Framework for Aquatic Biogeochemical Models (FABM) (Bruggeman and Bolding, 2014).

2.3 Data sets

~~Meteorological data required to drive GOTM were wind speed (m s^{-1}), atmospheric pressure (hPa), air temperature (ºC), relative humidity (%), cloud cover (dimensionless, 0-1), short wave radiation (W m^{-2}) and precipitation (mm day^{-1}).~~ Local meteorological variables were collected either from a small island 500 m offshore from the Erken Laboratory, or the Swedish Meteorological Hydrological Institute (SMHI) Svanberga Station just behind the laboratory. The Malma Island meteorological Station (59.83909º N, 18.629558º E) measured air temperature at 2 m above water surface, wind speed at 10 m above the water surface and short-wave radiation. These data were measured at one minute intervals and saved as 60 min mean values. Mean

sea level, pressure, relative humidity and precipitation were measured at the Svanberga Meteorological Station at 800 m from the Malma Island Meteorological Station (59.8321º N, 18.6348º E) with a frequency of 60 minutes. Hourly cloud cover was recorder from Svenska Högarna Station (59.4445 N, 19.5059 E) at 69 km south-east of Lake Erken.

The measured hourly meteorological data were used to construct two ~~other~~additional data sets that would replicate the data resolution that could potentially be used to force the GOTM model with ISIMIP scenarios. First to test running the model at a daily resolution, a daily data set was created by averaging the hourly one (except for precipitation which was summed). Secondly, this mean daily data set was disaggregated to form a synthetic hourly data set. Hourly estimations of air temperature wind speed, relative humidity and short wave radiation were estimated using the GRNN methods described below. For atmospheric pressure and cloud cover, mean daily values were assumed to be constant over the day. Precipitation was disaggregated assuming a uniform distribution of the daily total (Waichler and Wigmosta, 2003).

Since both of these data sets are based on the same measured hourly data, comparison of model simulations of lake water temperature, allow the importance of hourly vs daily temporal resolution in the forcing data to be evaluated, and also the improvements in model performance that can be obtained from daily data (as in the ISIMIP scenarios) when imposing a diurnal cycle on the mean daily data.

Water temperature data needed to calibrate the model was monitored from an automated floating station (59.84297º N, 18.635433º E). During ice-free conditions measurements were made every 0.5 m from 0.5m to a depth of 15 m. Measurements were made every minute, and a mean of these measurements was stored every 30 minutes.

2.4 Climate scenarios

The ISIMIP climate scenarios are bias-corrected global climate model (GCM) (Hempel et al., 2013) data made available at daily temporal and 0.5º horizontal resolution for the variables listed in Table 1. All data needed as input to the GOTM model are available in these climate scenarios with the exception of cloud cover, which was estimated from shortwave radiation (Martin and McCutcheon, 1999). Data from the grid box overlying Lake Erken were available from the GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 GCM models that were each run under ~~four~~three emission scenarios. These included a scenario having historical levels of atmospheric CO_2 between 1861 and 2005, and two future scenarios (RCP 2.6 and RCP 6.0) from 2006 to ~~2099~~2100. RCP 2.6 is the strongest mitigation pathway that is expected to limit mean global warming to between 1.5 and 2 ºC. RCP 6.0 is ~~a low~~an intermediate mitigation pathway where global warming is projected to rise to between 2.5 and 4 ºC by the end of century compared to the pre-industrial period (Frieler et al., 2017).

2.5 Temporal disaggregation of daily meteorological forcing data

~~The GCM scenarios have a daily time step, while lake model simulations are often forced with meteorological data at hourly or shorter time steps. Therefore, it was necessary to test the ability of the GOTM model to simulate Lake Erken water temperature using daily vs hourly meteorological forcing data, and to evaluate the need to disaggregate the daily GCM scenarios to a shorter time step.~~

Kathib and Elmernreich (2015) proposed a generalized regression artificial neural network (GRNN) model for predicting hourly variations in short-wave radiation from daily average measurements. Using the GRNN model to predict hourly solar radiation required ten geographical and climatic variables as input including hour, day, month, latitude, longitude, daily average short-wave radiation, daily precipitation, the solar elevation associated with the hour, and time of sunrise and sunset. Precipitation was used to define wet and dry status that affected atmospheric attenuation (Waichler and Wigmosta, 2003). There are also empirical models developed for calculating hourly air temperature, wind speed and relative humidity. Parton and Logan (1981) proposed a model for predicting diurnal variations in air temperature. Daylight air temperature was modelled using a sine wave with the minimum value at sunrise, maximum value at solar noon and mean value at sunset. Night-time air temperature was modelled as a linear interpolation between air temperature of the previous day and sunrise air temperature of the following day. Guo et al. (2013) generated hourly values of wind speed by computing a cosine function dependent on the mean daily wind speed, the maximum daily wind speed and the hour of the day when the wind speed is maximum. Waichler and Wigmosta (2003) estimated hourly values of relative humidity from daily maximum and minimum air temperature and daily maximum and minimum relative humidity. Using these studies as guidance, we developed GRNN models to predict hourly ~~a)~~ air temperature, ~~b)~~ wind speed and ~~c)~~ relative humidity. The input parameters for ~~each~~ GRNN ~~model~~models were~~:~~ ~~a) hour, day, month, latitude,~~ geographical variables: longitude, ~~mean daily air temperature, daily maximum and minimum air temperature, daily precipitation, hourly~~ latitude, solar ~~angle, and~~ elevation associated with the hour, time of sunrise and sunset ~~for predicting hourly air temperature; b),~~ hour, day~~, month, latitude, longitude and daily~~ and month; and meteorological variables: average, maximum and minimum daily air temperature, daily wind speed ~~for predicting wind speed; and c) hour, day, month, latitude, longitude, mean daily~~, daily relative humidity~~, daily precipitation, hourly air temperature and hourly short-wave radiation for predicting relative humidity. More detailed description of the GRNN methods and models are given in the supplementary material to this paper~~ and daily precipitation.

The GRNN models were constructed using 8 years of data. From this whole set of data, the first 5-years, from 2008 to 2012, were used for training, and the final 3-years of data from 2013 to 2015 were used for validating the results. The accuracy of the trained network was assessed by comparing the simulated output with actual observed hourly data. The performance index for training and validating sets of GRNN models are given in terms of mean bias error (MBE), root mean squared error (RMSE) and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970). More detailed description of the GRNN methods and models are given in the Supplemental section S1to this paper. The GRNN models were used to disaggregate the mean daily measured data, used to evaluate the necessity of disaggregation (section 3.2) and also for all GCM scenarios (section 3.3) to further evaluate the effects of disaggregation on the results of simulations of future changes in lake thermal structure.

2.6 Model set-up, calibration and validation

The GOTM model version 5.1 was used in this study. The meteorological parameters for running the model were air temperature (ºC), wind speed (m s^{-1}), short-wave radiation (W m^{-2}), cloud cover (dimensionless, 0-1), relative humidity (%), atmospheric pressure (hPa) and precipitation (mm day^{-1}or mm $hour^{-1}$). Inflows and outflows were not included in this study,

and water level was considered fixed in the simulations. This version of GOTM 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. The initial conditions for water temperature were derived from a measured vertical profile. GOTM was run at hourly model computational time step, and simulated water temperature was saved as daily mean values each 0.5 m (42 layers).

Calibration of the GOTM model was conducted to adjust the model parameters within their feasible range in order to minimize the error between measured and modelled temperature (Huang and Liu, 2010). A period of ~~8~~9 years was selected for the calibration of GOTM, 2006-2014 (included 1 year spin-up followed by ~~7~~8 years for calibration). The model parameters that were calibrated were surface heat-flux factor (shf_factor), short-wave radiation factor (swr_factor), wind factor (wind_factor), minimum turbulent kinetic energy (k_min) and e-folding depth for visible fraction of light (g2). The program used to calibrate the model was ACPy (Auto-Calibration Python), developed by Bolding and Bruggeman (https://bolding-bruggeman.com/portfolio/acpy/), it uses a differential evolution algorithm which calculates a log likelihood function based on comparing the modelled and measured water temperature (Storn and Price, 1997). The validation period was 2 years 2015-2016.

For both calibration and validation, daily average water temperatures were simulated when GOTM was forced using the three meteorological data sets described above: measured average daily, measured average hourly and synthetic hourly data. Model simulated profiles of mean daily water temperature were then compared to mean daily measured water temperature. Three separate model calibrations were made, based on simulations forced with the different meteorological data sets. During calibration the model was run approximately 10000 times to obtain a stable solution specifying the optimal parameter set~~, for each meteorological forcing data set.~~ The details of the feasible range of model parameters and ~~calibrated~~the parameters associated with each calibration are given in Table 2. ~~The same calibrated parameters were used to predict the thermal structure under GCM scenarios.~~

Model performance was evaluated by comparing average daily modelled and measured temperature profiles and other metrics describing the lake thermal structure (surface and bottom temperature, volumetrically weighted averaged whole lake temperature, Schmidt stability ~~and~~, thermocline depth~~).~~, duration, onset and loss of thermal stratification). The ~~model efficiency~~ coefficients used to quantify the strength of model fit were ~~mean bias error (~~MBE~~), root mean squared error (~~, RMSE~~)~~ and ~~Nash Sutcliffe efficiency (~~NSE~~) (Nash and Sutcliffe, 1970).~~

2.7 Thermal indices

A range of thermal metrics: surface temperature~~,~~ (shallowest observation), bottom temperature (deepest observation) and thermocline depth (depth of the maximum density gradient) were derived on a daily ~~bases~~basis from the daily simulated lake temperature profiles~~.~~ (temperature data with a vertical resolution of 0.5 m from 0.5 to 15 m depth). Also from these profiles, Schmidt stability (resistance to mechanical mixing due to the potential energy inherent in the density stratification of the water column; Schmidt, 1928; Idso, 1973) and whole-lake temperature (volumetric weighted mean whole lake temperature) were

estimated using Lake Analyzer (Read et al., 2011). ~~were estimated using Lake Analyzer (Read et al., 2011).~~ The duration of thermal stratification was calculated as the longest continuous period when the water column density difference from the bottom to surface of the lake was greater than 0.1 kg m^{-3} (according to ISIMIP2b lake sector protocol). The date of the onset and loss of the thermal stratification was defined as the first time that this density difference persisted for more than 5 days or was absent for at least 5 days (Kraemer et al., 2015).

~~**3 Results**~~

~~**3.1. Hourly meteorological modelling**~~

~~Air temperature, short wave radiation, relative humidity and wind speed were temporarily disaggregated into hourly values from mean daily data, using the GRNN models. A database was constructed using 8 years of measurements. From this whole set of data, the first 5 years of data, that is, from 2008 to 2012, were used for training, and 3 years of data from 2013 to 2015 were used for validating the results obtained. The accuracy of the trained network was assessed by comparing the predicted output with actual measured hourly data. The results are presented in Supplementary Fig. 1,2,3,4. The performance index for training and validating sets of GRNN models are given in terms of MBE, RMSE and NSE (see Table 3).~~

~~There was a close agreement between GRNN model predictions and measured meteorological data as shown in Fig. 1 for a single year data. For air temperature we obtained a NSE of 0.999 and 0.940 and RMSEs of 0.256 and 0.318 ºC for the training and validate data sets. The MBE values indicated a slightly cold temperature bias (MBE of $-1.70\ 10^{-4}$ and -0.057 ºC). Short wave radiation and relative humidity predictions for the training data set also show an accurate model performance with a NSE of 0.999 and 0.998 (RMSEs of 6.345 W m^{-2} and 0.790 %) respectively. For the validation data set the GRNN models performed somewhat worse, NSE of 0.870 and 0.686 (RMSE of 8.196 W m^{-2} and 1.021 %) for the short wave radiation and relative humidity predictions. Wind speed was the variable showing the poorest performance with a NSE of 0.779 and 0.584 (RMSE of 1.060 and 1.370 m s^{-1}) for the training and validate data sets. In general, the calculated MBE values in supplementary Fig. 5 show that the GRNN model tended to overestimate wind speed (MBE of 0.63±0.92 m s^{-1}) when the observed wind speed is lower than or equal to 3.84 m s^{-1}, whereas projected wind speed tends to be underestimated (MBE of −0.78±1.17 m s^{-1}) when the observations are greater than 3.84 m s^{-1}.~~

~~**3.2. Lake model performance**~~

~~Temperature observations and simulations, for the three configurations of meteorological forcing data for both calibration and validation periods, are shown in Fig. 2. Model performance metrics associated with these simulations are provided in Fig. 3 and Table 4.~~

~~These data demonstrate that GOTM was able to accurately reproduce the measured temperature profiles. For an average of all three calibration data sets a RMSE of 0.81 ºC and NSE of 0.96 was obtained. Temperature simulations for the validation period were more accurate (average RMSE of 0.66 ºC and NSE of 0.97) than for the calibration period (average RMSE of 0.95 ºC~~

and NSE of 0.94), but in both periods the model performance was considered acceptable. When comparing the metrics of model fit associated with simulations forced with the three different input data sets the RMSEs for calibration period ranged from 0.88 to 1.04 ºC, with lower error levels associated with simulations driven by hourly meteorological data sets, whereas for the validation period the RMSEs were comparable for all data sets, with slightly lower RMSE 0.63 ºC for the temperature simulations driven by daily meteorological data. The MBE values indicated a slight cold temperature bias (average MBE of -0.05 ºC).

The model performance predicting just surface temperatures (average RMSE of 0.63 ºC and NSE of 0.98) was better that estimations of the full temperature profiles. The MBE, showed that GOTM tended to produce a cold temperature bias (average MBE of -0.10 ºC). As would be expected the simulations of bottom temperature were slightly less accurate having average RMSE of 0.96 ºC and NSE of 0.90, with lower RMSE values for the validation period (average RMSE of 0.67 ºC) than the calibration period (average RMSE of 1.25 ºC), but in contrast to the surface temperature, there was a slight warm temperature bias (average MBE of 0.06 ºC). The best fits in the bottom temperature simulations were those driven by the measured hourly meteorological data set during the validation period (RMSE of 0.59 ºC and NSE of 0.97) and the synthetic hourly meteorological data, during the calibration period (RMSE of 1.16 ºC and NSE of 0.87). When evaluating the simulations of volumetrically weighted averaged whole lake temperatures we found that model errors were of a similar magnitude with an average RMSE of 0.53 ºC and NSE of 0.98, tending to a slight cold temperature bias (average MBE of -0.08 ºC).

The calculation of Schmidt stability, the resistance to mechanical mixing due to the potential energy inherent in the density stratification of the water column (Read et al., 2011), was also well simulated for using all three data sets (average RMSE of 17.24 J m^{-2} and NSE of 0.88). The lowest RMSE values were for the validation period (average RMSE of 13.34 J m^{-2}) whereas during the calibration period values were slightly greater (average RMSE of 21.14 J m^{-2}). Thermocline depth, defined as the depth of the maximum density gradient, was the parameter with the poorest performance (average RMSE of 3 m). The MBE values (average MBE of 0.80 m) indicate a bias towards under prediction of thermocline depth (shallower thermocline depths).

3.3. Long-term modelled changes in thermal stratification

The lake model was forced by four climate model projections and three emissions scenarios (historical, RCP 2.6 and RCP 6.0) available from ISIMIP for GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 using their original daily resolution and also at hourly resolution using meteorological disaggregated data developed using the GRNN models describe above. Simulated water temperatures for the historical, RCP 2.6 and RCP 6.0 scenarios under daily IPSL-CM5A-LR projections are presented as temperature isopleths in Fig. 4. These were created by averaging the daily temperature profiles for all years associated with each of the emission scenarios. These mean scenario temperature isopleths provide a clear visualization of how for future scenarios surface and bottom water temperatures are projected to increase with stronger and deeper stratification, an earlier stratification onset and later fall overturn and consequently longer stratification period.

2.8 Statistical analysis

Anomalies were calculated to further evaluate ~~these~~the impacts on lake water temperature and thermal stratification. The anomalies were computed for each GCM by taking the difference between the annual average of each year (~~2006-2099~~2011-2100) from RCP 2.6 and 6.0 scenarios and the average for the entire period ~~1975-2005~~1981-2010 from the historical scenario. ~~From these~~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~~. The slope of the significant trends were evaluated by least-squares linear regression, except when the residuals did not follow a normal distribution. Then the non-parametric Mann-Kendall test for the significance of trends and the Theil-Sen method (Theil, 1950; Sen, 1968) to estimate the slope of the significant trends were used instead. The t-Student mean difference test was used to compare average values of each of the thermal indices. Distribution normality and variance homoscedasticity were assessed by the Shapiro-Wilk test and Fisher's F test respectively. When thermal indices time series did not follow a normal distribution the non-parametric Mann-Whitney U test (equal variances) or Kolmogorov Smirnov test (different variances) were used instead. The statistical analysis was carried out using R version 3.4.4. The progress of climate-related impacts on the thermal stratification of the lake over the 21st century was assessed as the difference in mean conditions between the reference period (1981-2010) and both mid-century (2041-2070) and late-century (2071-2100). Climate model data followed the same statistical analysis.

~~In general there were significant changes in all the metrics describing thermal stratification evaluated in this study as can be seen from the frequency distribution of yearly anomalies in Fig. 5,6 and the detailed statistics presented in Table 5. The exception to this is the GFDL-ESM2M model which showed lower or non-significant changes. The other three models showed more consistent and larger anomalies as compared to GFDL-ESM2M. Similar trend in the anomaly distributions were seen when the GOTM model was forced with either mean daily or synthetic hourly data. Detailed comparison of the results derived using the two different forcings (Table 5) suggest that the simulated changes are slightly greater when simulations are forced with the mean daily data. However, in both cases the same direction in the trends and the same overall descriptions of change is found.~~

~~Rates of change in whole-lake temperature calculated for over the length for RCP2.6 and 6.0 scenarios~~

3 Results

3.1. Hourly meteorological modelling

There was a close agreement between GRNN model predictions and measured meteorological data as shown in Fig. 1 for a single year and in Supplemental section S1. For air temperature, short wave radiation and humidity the statistics of model performance always suggested a strong model fit in the training data sets and also remained strong, but somewhat lower in the validation data sets (Table 3). NSEs were 1.00 for the training data sets and ranged from to 0.69 to 0.94 for the validation data sets. Estimates of bias were very small. Wind speed was the variable showing the poorest performance with a NSE of 0.78 and

0.58, and RMSE values of 1.06 and 1.37 m s^{-1} for the training and validate data sets. In general, the GRNN model tended to overestimate wind speed (MBE of 0.63±0.92 m s^{-1}) when the observed wind speed is lower than or equal to 3.84 m s^{-1}, whereas projected wind speed tends to be underestimated (MBE of -0.78±1.17 m s^{-1}) when the observations are greater than 3.84 m s^{-1}.

3.2. Lake model performance

Temperature observations and simulations, for the three configurations of meteorological forcing data for both calibration and validation periods, are shown in Fig. 2 and Supplemental section S2. Model performance metrics associated with these simulations are provided in Table 4. These data demonstrate that GOTM was able to accurately reproduce the measured temperature profiles. For an average of all three calibration data sets a RMSE of 0.95 ^{0}C and NSE of 0.94 were obtained. Temperature simulations in the shorter and less variable validation period (RMSE of 0.66 ^{0}C and NSE of 0.97) were more accurate than for the calibration period, but in both periods the model performance was considered strong. For full profile temperature the maximum RMSE value was 1.04 ^{0}C and the minimum NSE was 0.93. Bottom temperature was least accurately simulated with RMSE and NSE values reaching 1.33 ^{0}C and 0.83 respectively.

When comparing the metrics of model fit associated with simulations forced with the three different input data sets the simulations forced with mean daily input were slightly less accurate than those forced with either the measured or synthetic hourly input. As an example, full profile RMSE values for calibration period ranged from 0.88 to 1.04 ºC, with the lower error levels associated with simulations driven by hourly meteorological data sets, whereas for the validation period the RMSEs were comparable for all data sets. The MBE values of the full temperature profiles indicated a slight cold temperature bias (average MBE of -0.05 ºC). The model performance predicting just surface temperatures was similar for all of the three calibrations (average RMSE of 0.67 ºC and NSE of 0.97), and were more accurate than the estimations of the full temperature profiles. The MBE, showed that for surface temperature GOTM also tended to produce a small cold temperature bias (average MBE of -0.10 ºC). The simulations of bottom temperature were slightly less accurate having average RMSE of 0.96 ºC and NSE of 0.90, and also showed a tendency have a slightly higher RMSE values for calibrations forced with daily input. Also the bottom temperature showed lower RMSE values for the validation period (average RMSE of 0.67 ºC) than the calibration period (average RMSE of 1.25 ºC), but in contrast to the surface temperature, there was a slight warm temperature bias (average MBE of 0.06 ºC). When evaluating the simulations of volumetrically weighted averaged whole lake temperatures we found that model errors were of a similar magnitude for all simulations in both the calibration and validation periods with an average RMSE of 0.53 ºC and NSE of 0.98, tending to a slight cold temperature bias (average MBE of -0.08 ºC).

The calculation of Schmidt stability was also well simulated using all three data sets (average RMSE of 17.24 J $m^{-2}$and NSE of 0.88). The lowest RMSE values were for the validation period (average RMSE of 13.34 J m^{-2}) whereas during the calibration period values were slightly greater (average RMSE of 21.14 J m^{-2}). Thermocline depth was the parameter with the poorest performance (average RMSE of 3 m). The MBE values (average MBE of 0.80 m) indicate a bias towards under prediction of thermocline depth (shallower thermocline depths). The RMSE associated with the prediction of the duration of stratification

was, on average, 10.43 days. With lower RMSE values for the validation period (average RMSE of 8.04 days) than the calibration period (average RMSE 12.81 days). The simulations of onset of the stratification were more accurate having average RMSE of 2.64 days, but in contrast predictions of the loss of stratification was less accurate (average RMSE of 7.99 days).

3.3. Climate data projections

The lake model simulations undertaken here were forced by four climate model projections (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5) that were in turn forced with three emissions scenarios (historical, RCP 2.6 and RCP 6.0). Average annual air temperature of the climate model ensemble for the reference period (1981-2010) was 11.88 ºC Disaggregation of the climate input to an hourly time-step resulted in a slightly warmer temperature (12.05 ºC)[1]. Under future scenario RCP 2.6 the average increase was projected to be 2.22 ºC (1.71 ºC) by mid-century (2041-2070) with a negligible change after mid-century, as would be expected from this scenario with the strongest mitigation. During the period up to 2070 air temperature increased at a rate of 0.08 to 0.17 ºC decade^{-1} (0.06 to 0.14 ºC decade^{-1}). In contrast, under RCP 6.0 average air temperature increased by 2.61 ºC (2.01 ºC) by mid-century and continued rising to 3.61 ºC (2.76 ºC) by late-century. For RCP 6.0 the trend in air temperature increased, over the entire 2011-2100 period, on average, by 0.34 ºC decade^{-1} (0.26 ºC decade^{-1}) over all GCMs with the individual trends ranging from 0.18 to 0.43 ºC decade^{-1} (0.14 to 0.33 ºC decade^{-1}). For the remaining meteorological variables there were less distinct changes between the historical and future periods. Under the RCP 2.6 scenario the overall annual mean change in wind speed was negligible, while under RCP 6.0 two options were projected, an increase (GFDL-ESM2M and MIROC5) and decrease (HadGEM2-ES and IPSL-CM5A-LR). Relative humidity was projected to decrease for future scenarios RCP 2.6 and 6.0. For RCP 6.0 significant trends ranged from 0.29 to 0.36 % decade^{-1} in the interval 2011-2100. An increase in short-wave radiation was projected for all RCP scenarios by late-century, with a negligible mean change after mid-century under RCP 6.0. The increase in short-wave radiation is coupled with a decrease in cloud cover. By late-century the mean decrease in cloud cover was projected to be 0.06 for RCP 2.6, and 0.07 for RCP 6.0. More detailed evaluations of the differences in the climate projection based on the original ISIMIP daily time step and the hourly disaggregated data are given in the Supplemental section S3.

3.4. Long-term modelled changes in thermal stratification

Lake model simulations were made using both the original daily resolution of the ISIMIP GCM scenarios and also at hourly resolution using disaggregated data developed using the GRNN models. Simulated water temperatures for the historical, RCP 2.6 and 6.0 scenarios under daily IPSL-CM5A-LR projections are presented as temperature isopleths in Fig. 3 and Supplemental section S4. These were created by averaging the daily temperature profiles for all years associated with each of the emission scenarios. These mean scenario temperature isopleths provide a clear visualization of how for future scenarios

[1] Results based on the hourly disaggregated data are always shown in parenthesis

surface and bottom water temperatures are projected to increase with stronger and shallower stratification, an earlier stratification onset, a later fall overturn and consequently a longer stratification period.

In Figs. 4-5 we show the long-term trends in the anomalies in lake thermal metrics simulated to occur over the RCP 2.6 and 6.0 emission scenarios. Trends in whole-lake temperature calculated for over a period of 90 years (2011-2100) were projected to increase except in the case of GFDL-ESM2M which showed weaker or non-significant changes for all measures of thermal stratification~~. Simulated changes were generally slight less for the simulations driven by daily forcing data as shown by the figures in parentheses For~~ (Table 5 and Supplemental section S4). Under RCP 2.6 ~~rate of change~~ significant trends ranged from 0.~~06~~07 to 0.10 ºC decade^{-1} (0.05 to 0.08 ºC decade^{-1}), but most of the change occurred in the first half of the ~~values in bracket refer to future projections when the lake model was forced at hourly resolutions.~~century. For RCP 6.0, the projected rate of change ranged from 0.~~15~~14 to 0.~~27~~26 ºC decade^{-1} (0.~~11~~10 to 0.19 ºC decade^{-1}). ~~IPSL-CM5A-LR~~By late-century, the mean projected ~~the largest~~ increase ~~being 0.59~~in whole-lake temperature was 1.34 ºC (~~0.43~~1.00 ºC) ~~under~~for RCP 2.6~~ºC~~, and 2.~~51~~39 ºC (1.75 ºC ~~(1.79 ºC) under~~) for RCP 6.0~~.~~, with a negligible change after mid-century under RCP 2.6 (Fig. 6, Table 6 and Supplemental section S4).

~~Surface temperature is warmed significantly, with rates of change ranging from 0.09 to 0.16 ºC decade^{-1} (0.07 to 0.13 ºC decade^{-1}) and from 0.17 to 0.39 ºC decade^{-1} (0.13 to 0.29 ºC decade^{-1}) for RCP 2.6 and 6.0 respectively. These warming rates are consistent with the increase~~ Changes in ~~the air temperature, with rates of change ranging from 0.10 to 0.18 ºC decade^{-1} (0.08 to 0.15 ºC decade^{-1}) for RCP 2.6 and from 0.20 to 0.44 ºC decade^{-1} (0.15 to 0.33 ºC decade^{-1}) for RCP 6.0. For RCP 2.6 HadGEM2-ES projected the largest increases in~~lake surface temperature ~~of 1.48 ºC (1.20 ºC), whereas for RCP 6.0 IPSL-CM5A-LR and HadGEM2-ES both projected similar large increases in~~ were, as expected, greater than for whole lake temperature. For the reference period the mean April-September surface temperature ~~of 3.59 ºC (2.66 ºC) and 3.52 ºC (2.58 ºC) respectively.~~ was on average 13.61 ºC (13.84 ºC) warming up significantly over 21st century, so that by late-century the average projected increase was 1.79 ºC (1.35 ºC) for RCP 2.6, and 3.08 ºC (2.31 ºC) for RCP 6.0. From 2011 to 2100 there was a significant long-term trend for RCP 2.6 surface temperature which increased at a rate of 0.07 to 0.15 ºC decade^{-1} (0.06 to 0.13 ºC decade^{-1}). Under RCP 6.0 the mean surface temperature increase of the ensemble was 0.30 ºC decade^{-1} (0.23 ºC decade^{-1}) ranging between 0.16 to 0.38 ºC decade^{-1} (0.12 to 0.29 ºC decade^{-1}).The projected increase in bottom temperature was not as marked as it was for the other metrics of lake temperature. On average, the bottom temperature increased from 9.20 ºC (9.67 ºC) in the reference period to 9.77 ºC (9.99 ºC) and 10.32 ºC (10.34 ºC) by late-century for RCP 2.6 and 6.0 respectively. Significant rates of change in bottom temperature were not predicted during the RCP 2.6 scenario, but for the RCP 6.0 scenario bottom temperature did undergone significant warming rates ~~in most~~for HadGEM2-ES and MIROC5 projections~~, ranging from~~ being 0.~~07 to~~06 ºC decade^{-1} and 0.11 ºC decade^{-1} (0.09 ºC decade^{-1}) respectively.

There were also projected changes in the resistance to mechanical mixing. ~~For~~ For the reference period, an average of 68.65 J m^{-2} (65.56 J m^{-2}) was required to completely mix the water column, while by late-century it increased by 29.08 J m^{-2} (22.74 J m^{-2}) for RCP 2.6 ~~rates of change in Schmidt stability were significant for IPSL-CMR5A-LR~~ and ~~HadGEM2-ES, corresponding to the same projections that experienced the largest increases in surface temperature, being 2.52 J m^{-2} decade^{-1}~~

49.22 J m^{-2} (38.07 J m^{-2}) for RCP 6.0 (Fig. 4, Table 6 and Supplemental section S4). This greater stability also corresponds to a longer duration of stratification. From 2011 to 2100, a significant increase in the duration stratification was projected for both future scenarios RCP 2.6 and 6.0, ranging from 1~~(~~.13 to 1.70 day decade^{-1} (0.87 ~~J m^{-2}~~to 1.30 day decade^{-1}) for RCP 2.6 and 2.~~60 J m^{-2}~~45 to 3.56 day decade^{-1} (2.~~21 J m^{-2}~~00 to 3.08 day decade^{-1}) for RCP 6.0 (Fig. 5, Table 5 and Supplemental section S4.),which led to a mean change of was 13 days (11 days) and 22 days (18 days) for RCP 2.6 and 6.0 respectively~~. For RCP 6.0 a significant rate of change was projected in all projections, ranging from 2.50 to~~ (Fig. 7~~.93 J m^{-2} decade^{-1} (1.89 to 6.14 J m^{-2} decade^{-1}). For RCP 2.6 a significant rate of change in the duration~~, Table 6 and Supplemental section S4.). The longer period of stratification ~~was projected to be 1.73 and 1.01 days decade^{-1} (1.37 and 0.80 day decade^{-1}), which increased the stratification period by 16 and 9 days (13 and 7 days) for IPSL-CM5A-LR and HadGEM2-ES projections respectively. This~~ resulted from ~~changes in both the onset and loss of stratification (Table 5). For RCP 6.0, there was a further rate of change in the duration of~~ both an earlier onset of thermal stratification ~~from 2.19 to 3.55 day decade^{-1} (1.78 to 2.94 day decade^{-1}), which resulted in a 20 to 33 days (17 to 27 days) longer~~ and a later loss of thermal stratification ~~period. Thermocline depth~~(Figs 5, 7, Table 5-6 and Supplemental section S4). Mean annual thermocline depth ~~is expected~~ was simulated to be shallower under future conditions. ~~However, for RCP 2.6 significant rates of change were not simulated, but for RCP 6.0 significant rates of change were found to be 0.08 and 0.12 m decade^{-1} (0.08 and 0.12 m decade^{-1}) or rather 0.74 and 1.13 m shallower (~~By late-century the reduction in thermocline depth was projected to be 0.38 m (0.41 m) for RCP 2.6 and 0.49 m (0.57 m) for RCP 6.0, although a significant trend in the decline were only found for the later scenario~~0.78 and 1.14 m) for IPSL-CM5A-LR and HadGEM2-ES~~.

~~Extreme changes also showed a pronounced increase during the future scenarios. For the RCP 6.0 scenario, the 95th percentile from the distribution of surface temperature anomalies ranged from 2.46 ºC (1.87 ºC) for GFDL-ESM2M to 5.02 ºC (3.76 ºC) for HadGEM2-ES (Fig. 5). The HadGEM2-ES projection also showed the highest increase in stability, 95th percentile of 112.51 J m^{-2} (89.05 J m^{-2}), and the shallowest thermocline depth, 95th percentile of 1.99 m (1.78 m). Extreme changes in the duration of stratification were least for GFDL-ESM2 and greatest for HadGEM2-ES. Longer stratification periods were projected for both models with the 95th percentile of change in duration increasing 51 days (38 days) for HadGEM2-ES and 22 days (21 days) GFDL-ESM2M projection. These increases in duration were dominated by earlier onset of stratification for GFDL-ESM2M and by both earlier onset and later fall overturn for HadGEM2-ES (Table 5).~~

The trends in Figs. 4-5 are quite variable from year to year, and as would be expected there is no direct correspondence in the temporal variations of one GCM relative to another. To provide an alternative method of comparing the changes simulated by the future climate scenarios shown in Figs. 4-5, the daily anomalies for each trend line are also presented as frequency distributions in the Figs. 6-7 for the simulations made under the RPC 6.0 scenario. These show that for all metrics there is a clear shift in the lake thermal conditions that are consistent with a warmer climate, but that also there are extremes in the distributions that can lead to unrepresentative results, when for example future conditions briefly return to historical levels, or when the effects of warming are much greater that would be expected on average. This later case can cause important changes in lake ecology if the extreme conditions result in a change in lake state by the passing of a tipping point. Figs. 6-7 also clearly

shows the differences in simulations forced by different GCMs. Most obvious is the difference in the results from GFDL-ESM2M, which consistently simulated smaller changes in lake thermal structure during the mid and late century periods, despite having a data distribution that was similar to the other models during the historical period.

3.5. Comparison between long-term thermal metrics derived from daily and hourly climate data

Future changes in thermal metrics based on both RCP 2.6 and RCP 6.0 were slightly greater when the GOTM model was forced at daily resolutions (Tables 5-6 and Supplemental section S4) than at an hourly resolution. This included changes in mean surface temperatures and also in annual average whole-lake temperature (Supplemental section S5). However, under RCP 2.6 non-significant differences were found for bottom temperature, Schmidt stability, thermocline depth, or the duration, onset and loss of stratification. In all cases where differences were found between the simulations forced with daily vs. hourly data there were no change in direction and only minor changes in the magnitude of the change suggested by the simulations (Supplemental sections S4-S5).

4. Discussion

The ~~GOTM model was able to produce a good~~simulated water temperature and related metrics of thermal stratification were in excellent agreement ~~between the model output and observed data during both the~~ with the extensive set of measured water temperature data that were available for model calibration ~~and the validation period (~~at Lake Erken (Moras 2019 Fig. 3, Table ~~4~~3). Water temperature simulations were apparently more accurate for the validation period (2015-2016) ~~were more accurate~~ than for the calibration period (2006-2014~~) due~~), which may appear unusual, but is due to the higher variability in observed water temperature during the ~~long~~longer calibration period. ~~Bottom water temperature RMSEs showed better agreement between simulations~~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 ~~observations when~~ weaker stability. The thermocline depth was the ~~lake model~~ thermal metric that was ~~forced using measured average hourly meteorological data instead of measured average daily meteorological data.~~predicted with the greatest uncertainty. This ~~could be attributed to a more confident prediction of the diurnal heating and cooling cycles, and hence~~ is in part caused by the ~~downward flow of heat into the hypolimnion. Thermocline depth was harder to predict, simulated thermocline depths differed from~~ presence of internal seiches in Lake Erken, which result in the measured temperatures in the ~~observed values contributing to additional errors in water temperature for depths near~~ region of the thermocline~~.~~ having a level of variability that cannot be reproduced by 1D models such as GOTM. Bruce et al. (2018) detected a strong correlation between accuracy of the extinction coefficient and model simulations of full-profile temperature and thermocline depth, ~~and thus~~which shows the importance of light extinction in the prediction of thermocline depth. ~~A~~Since a single ~~calibrated~~fixed value of e-folding depth (Table 2) for the visible fraction of the light (the inverse of the extinction coefficient) ~~were~~was used in the GOTM ~~which prevented~~simulations, the ~~evaluation~~effects of seasonal ~~effects~~ variations in light extinction (Perroud et al., 2009~~). Also it should be noted that the internal seiche movement observed in the measured data~~

~~(Fig. 2) is not simulated by a 1D model such as GOTM. Thus errors in ~~), on thermocline depth ~~are at least in part due to limitations in the 1D model framework~~were not evaluated.

~~The performance of the GOTM model are comparable to those reported in other 1-D modelling studies. Moras et al. (2019, in review) ran GOTM using hourly measured meteorology for a 56-year period and RMSE for daily full-profile water temperature was 1.12 ºC, Magee and Wu (2017) reported RMSEs of 0.30 and 0.53ºC (lake Mendota) and 1.45 and 1.94 ºC (Fish lake) for DYRESM temperature simulations in the epilimnion and hypolimnion respectively, and Perroud et al. (2009) simulated water temperature profiles of lake Geneva for a 10-year period with RMSEs of < 2 ºC for DYRESM and 3 ºC for SIMSTRAT. For our simulations with GOTM, model performance was slightly more accurate for the calibration data set when GOTM was forced with synthetic hourly meteorological input, rather than measured hourly meteorological input. Similar levels of performance using the two different data sets was in part caused by changes in the calibrated parameters used to characterize the lake thermal structure. Apparently calibration can in part compensate for the lack of diurnal variability in the daily forcing data.~~

The model parameters adjusted during the calibration processes were ~~nondimensional~~non-dimensional scaling factors (shf_factor, swr_factor and wind_factor) and physical parameters ~~with strong~~which strongly influence in the vertical distribution of light and temperature (k_min and g2). ~~These parameters are key for the determination of the heat budget in the water column.~~ Wind is the dominant driver of mixing in lakes, and increases or decreases of wind speed (wind_factor) changes the amount of ~~turbulence~~turbulent kinetic energy available for mixing. The wind scaling factor is often important when wind ~~station is located~~measurements occur some distance from the lake and/or to ~~consider~~account for wind sheltering effects (Markfort et al., 2010). One would not expect ~~that these factors would be important for~~the wind factor to deviate greatly from 1.0 at Lake Erken where wind ~~was~~is measured on an island in the lake. However, the dominant wind direction is along the lake's longest east-west fetch (Yang et al., 2014), ~~this~~which could explain the need to scale up the unidirectional wind speed measurements that were used as an input to GOTM. Furthermore, since it is the cube of wind speed that affects lake mixing, use of longer averaging periods will underestimate the effects of gusting and variable winds. This could explain why we obtain higher calibrated values of the wind~~-~~_factor when forcing the model with measured daily rather than hourly data (Table 2). Higher values of the wind_factor were also obtained when GOTM was forcing with synthetic hourly meteorological drivers. This is due to an underestimation of the faster wind speed predictions from the GRNN model (Fig.-1~~, supplementary Fig. 5).~~ and Supplemental section S1.). During the ACPy calibration each of these parameters were calibrated while simultaneously influencing each other; shf_factor, swr_factor, wind_factor and g2 have a strong influence on heat and energy exchange across the air-water interface. There is to some extent an unavoidable tendency for ~~some error in one parameter to be cancelled out by opposite errors in another parameter. When GOTM was forcing with measured daily average and synthetic hourly meteorological drivers a large wind_factor led to surface cooling, but on the other hand a lower shf_factor and smaller values of g2 (equivalent to higher extinction coefficient) promoted an increase in surface water temperature. When GOTM was forcing with measured hourly average meteorological drivers, g2 had the larger values (equivalent to lower extinction coefficient) and shf_factor was closer to 1 showing that a deeper penetration of energy entering into the lake provided more~~

~~realistic warming of the surface, a lower wind factor was also found, which means that less surface cooling is to be offset~~the error in one parameter to be cancelled out by opposite errors in another parameter. This also illustrates how, to some extent, the calibration process can compensate for some of the limitations related to the temporal resolution of the model forcing data.

The performance of the GOTM model obtained in this study is comparable with results reported by others. Moras et al. (2019) who ran GOTM using hourly measured meteorology for a 57-year period obtained a RMSE for daily full-profile water temperature of 1.09 ºC. Using the DYRESM model Magee and Wu (2017) reported RMSE values of 0.30 and 0.53ºC for lake Mendota and 1.45 and 1.94 ºC for Fish lake for temperature estimates of the epilimnion and hypolimnion respectively. Perroud et al. ~~In general, surface water temperature was projected to increase at a rate that is 83-93 % of that of the air temperature increase. This conclusion is in close agreement with other modelling studies which found a relationship between the surface water and air warming rates of 75-90% (Schmidt et al., 2014) and 70-85% (Shatwell et al., 2019). However, one exception was observed for the IPSL-CM5A-LR projection under RCP 2.6 scenario using daily resolution in the forcing inputs when increase of 0.109 ºC decade^{-1} in surface water temperature slightly exceeded the increase of 0.105 ºC decade^{-1} in air temperature. The reasons for that this scenario shows a somewhat different behaviour is probably related to some inconsistencies in the GCM models and also bias correction that was applied to the ISIMIP data.~~

(2009) simulated water temperature profiles of lake Geneva over a 10-year period and obtained RMSE values of < 2 ºC for the DYRESM model and < 3 ºC for the SIMSTRAT model.

The projected changes in lake thermal metrics depends on the selected GCM model and the future scenario or representative concentration pathway (RCPs) that was simulated. The range of greenhouse gas (GHG) emissions include in this study were a stringent mitigation scenario (RCP 2.6) and an intermediate scenario (RCP 6.0). Consistent with the ISIMIP2b simulation strategy that is intended to evaluate RCP 2.6 as a scenario that aims to keep global warming below 2°C above pre-industrial temperatures by 2100. In contrast for RCP 6.0 increased levels of GHG emissions suggest that the global mean temperature will continually increase by 2.5 and 4 ºC by the end of the century. The effects of the mitigation measures that were adopted in RCP 2.6 on lake thermal structure become most apparent in the late century. For example, for MIROC5 (when the lake model was forced at daily resolutions) the projected surface temperature change for mid-century was similar for the two RCPs (2.10 ºC for RCP 2.6 and 1.98 ºC for RCP 6.0), but for the late-century period the projected change in surface temperature diverge among RCPs. Under RCP 2.6 the surface temperature change declines from 2.10 to 1.80 ºC, while under RCP 6.0 the change in surface temperature was projected to further increase from 1.98 to 2.97 ºC. Similar changes were projected for all thermal metrics. Under RCP 6.0 there was also an increase in bottom temperature but at rates that were slower than surface temperature, changes in lake stability increased from 38.67 J m^{-2} by mid-century to 64.62 J m^{-2} by late-century, increasing the duration of stratification (from 16 to 22 days). While the there was a general agreement among the models in the direction and relative magnitude of change in many of the metrics of lake thermal structure there were also some differences among GCMs (Figs. 4-7 and Supplemental section S4) especially in relation to the GFDL-ESM2M model which consistently estimated lower levels of change. For example, by late century largest changes in surface temperature were projected for HadGEM2-ES (4.04 ºC) and the lowest for GFDL-ESM2M (1.67 ºC) under future scenario RCP 6.0 when the lake model was forced at daily

resolutions. However, the increase in the projected bottom temperature for GFDL-ESM2M (1.24 ºC) was greater than for HadGEM2-ES (0.91 ºC). This could be in part due to the projected changes in wind speed. The wind speed was projected to increase by 0.18 m s^{-1} for GFDL-ESM2M transferring heat to the lake bottom, but for HadGEM2-ES the wind speed decrease by 0.15 m s^{-1} (atmospheric stilling; Woolway et al, 2017; Woolway et al., 2019) reducing the magnitude of vertical mixing. Resulting in a greater gradient between surface and bottom temperatures and higher increases in the Schmidt stability (77.57 J m^{-2}). This increased thermal gradient for HadGEM2-ES promoted shallower thermocline depth (1.26 m), but for GFDL-ESM2M a lower change in lake stability was projected (12.26 J m^{-2}) thereby a deeper thermocline depth (0.22 m). Higher surface water temperature and stronger Schmidt stability can explain why the increased duration of stratification was projected to be longer for HadGEM2-ES (34 days) than for GFDL-ESM2M (6 days). The small change in thermal stability also explains why no change in loss of stratification was projected for GFDL-ESM2M. This illustrates the complexity of climate model – lake model interaction, and clearly shows the importance of ensemble model simulations, as adopted by ISIMIP, to evaluate the effects of climate change on lakes.

When calibrating the GOTM model we found that model errors between simulated and measured water temperature were similar when GOTM was forced with either measured hourly or synthetic hourly meteorological data, and that the results obtained from the calibrations forced with mean daily metrological input were also similar to those obtained from the calibrations based on hourly input. This suggests that the daily time step of the ISIMIP climate scenarios is sufficient for forcing the GOTM model and that for most studies within the ISIMIP lake sector disaggregation to hourly time step will not be necessary. For example, changes in surface water temperature was on the order of 0.29 ºC $decade^{-1}$, with simulations forced with daily inputs 0.22 ºC $decade^{-1}$ degrees with hourly input data for MIROC5 under RCP 6.0. These differences are of the same magnitude as the differences simulated using different GCM models. Similar levels of model performance using daily or hourly forcing data were obtained in part because of separate calibrations when the GOTM model was forced with the different data sets. Changes in the calibrated parameters used to characterize the lake thermal structure (Table 2), apparently can compensate for the lack of diurnal variability in the daily forcing data. GRNNs proved to be an effective method to disaggregate daily GCM forcing to an hourly temporal resolution for different weather variables such as air temperature, short-wave radiation, etc. However, GRNNs require a training phase, in which the diurnal patterns to be learned are presented to the network from historical meteorological measurements, and therefore if there are future changes in diurnal patterns, these cannot be reproduced. In addition, there is a high computational cost of disaggregating and storing the long-term daily climate data into an hourly data set.

The projected changes in thermal metrics were strongly influenced by the GMCs used to drive the water temperature simulations. Due to the high interannual variability long periods of simulation were needed to ensure that the uncertainty is fully represented (Figs 4-7 and Supplemental sections S3-S4). Under RCP 6.0 trends in surface temperature calculated for the period 2011-2100 were projected to increase 0.38 ºC $decade^{-1}$ for both HadGEM2-ES and IPSL-CM5A-LR when the lake model was forced at daily resolutions. However, 5^{th}, 50^{th} and 95^{th} percentiles for surface temperature anomalies differ between models, being 0.84, 2.93 and 4.86 ºC for HadGEM2-ES and 0.33, 2.56 and 4.37 ºC for IPSL-CM5A-LR. Placing the probability

density function (PDF) for HadGEM2-ES to the right of the PDF for IPSL-CM5A-LR, illustrating that more extreme increases in surface temperature were projected by HadGEM2-ES. Projected bottom temperatures differed between HadGEM2-ES and IPSL-CM5A-LR. HadGEM2-ES PDF was left-skewed and the median was 0.58 ºC while IPSL-CM5A-LR PDF was right-skewed and the median was 1.16 ºC, as a consequence lake stability was stronger for HadGEM2-ES (5^{th} and 95^{th} percentiles were 5.22 and 110.55 J m^{-2}) than for IPSL-CM5A-LR (5^{th} and 95^{th} percentiles were -18.77 and 90.64 J m^{-2}), even though the Schmidt stability medians were similar for both GCMs. Similarly, when projecting longer duration of stratification for HadGEM2-ES (5^{th}, 50^{th} and 95 percentiles were -0.63, 26.37 and 49.42 days) than IPSL-CM5A-LR (5^{th}, 50^{th} and 95 percentiles were -10.33, 17.67 and 40.92 days). GCMs are useful for assessing climate change impacts on lakes, but GCMs configurations vary from one to another. Therefore, it is crucial to assess different GCMs in a probabilistic manor (Figs. 4-7) to encapsulate the uncertainty in the lake thermal metrics without compromising the variability

The study carried out by Moras et al. (2019) found changes in the phenology of Lake Erken thermal stratification from 1961 to 2017. A significant increase in summer epilimnetic and whole-lake temperature of 0.35 ºC $decade^{-1}$and 0.24 ºC $decade^{-1}$ occurred over the entire study period. While in spring and autumn larger significant positive trends were detected over the subinterval 1989-2017. In the present work future changes under the RCP 6.0 emission scenario found trends that were of a similar magnitude. The summer trends for the period 2011-2100 projected a significant increase in epilimnetic and whole-lake temperature ranging from 0.19 to 0.45 ºC $decade^{-1}$ and 0.15 to 0.26 ºC $decade^{-1}$, respectively, when the lake model was forced at daily resolutions, while changes in summer hypolimnetic temperature was non-significant. During the spring and autumn significant increases in epilimnetic whole-lake temperature were also projected under RCP 6.0 when the lake model was forced at daily resolutions, but they were somewhat lower than the ones detected by Moras et al. (2019). The increase in spring epilimnetic and whole-lake temperature ranged from 0.15 to 0.38 ºC $decade^{-1}$ and 0.15 to 0.30 ºC $decade^{-1}$, while those in Moras et al. (2019) showed a higher rate of warming (0.44 ºC $decade^{-1}$and 0.40 ºC $decade^{-1}$ for epilimnetic and whole-lake temperature, respectively), and the GCM simulations promoted shorter duration in stratification. The projected increase in spring and autumn hypolimnetic temperature were similar in magnitude and in summer non-significant trends were detected either in this study or in Moras et al. (2019). The simulations made here and by Moras et al (2019) are for the same lake using the same lake model. The fact that the simulations presented here using the RCP 6.0 emission scenarios showed similar or slightly lower rates of change compared to the simulations made by Moras et al (2019) when the model was forced with measured historical data, are unexpected given that the RCP 6.0 scenario would project an accelerated rate of climate change compared to the historical period. This suggests that, at least for Lake Erken, future changes in lake thermal structure based on the ISIMIP2b GCM projections may to some extent underestimate the actual changes that will occur.

The projected changes in thermal stratification can influence many aspects of the lake ecosystem. Increases in thermal stability and duration of stratification can intensify hypolimnetic oxygen depletion (Foley et al., 2012; Schwefel et al., 2016) and hence induce enhanced internal phosphorous loading (North et al., 2014), increase the release of dissolved iron and manganese from sediments (Schultze et al., 2017) and also increase methane emissions (Grasset et al., 2018). Warming lake temperature affects biological rates of metabolism, growth and reproduction (Rall et al., 2012) and can promote cyanobacterial blooms (Paerl and

Paul, 2012). When coupled to a reduction in oxygen-rich water, warming water temperature leads to a lower fish populations (O'Reilly et al., 2003; Yankova et al., 2017). Increase in evaporation associated with warming can lead to declines in lake water level (Hanrahan et al., 2010) with implications for water security.

4. Conclusion

~~This study showed the ability of the GOTM model to simulate accurately Lake Erken water temperature when the model was forced using either daily or hourly temporal resolution inputs. Neuronal networks were shown to be an effective method to disaggregate different weather variables such as air temperature, and short-wave radiation, in order to generate synthetic hourly meteorological data from the daily data that is typically available from GCM models. Model performance was slightly improved when using the synthetic hourly data, and climate change effects were somewhat greater when using such data to drive future climate simulations. However, it is not clear that data disaggregation is needed given the computational costs of developing such data sets and running long-term simulations at an hourly time step. Future climate simulations showed similar trends in the anomaly distributions when the GOTM model was forced with both mean daily or synthetic hourly meteorological data, and we also found evidence that the calibration procedure partly compensates for differences in the temporal resolution of the model input.~~

In this study, which was the first test simulating lake hydrothermal structure following ISIMIP2b protocols, ensemble simulations show that changes in Lake Erken's surface temperature was projected to increase ~~from 0.87 to~~on average by 1.~~48~~79 ^{0}C for RCP 2.6 and ~~from 1.59 to~~by 3.~~59~~08 ^{0}C for RCP 6.0, and the length of the stratification also was projected to be longer ~~from 9 to 16~~by 13 days for RCP 2.6 and ~~from 20 to 33~~by 22 days for RCP 6.0 by the end of the 21st century. Most changes in the RCP 2.6 scenario occurred during the first half of the century suggesting that the aggressive mitigation methods represented in this scenario would be effective at reducing future changes in lake thermal structure. We also extensively document coinciding changes in water column temperatures, water column stability and ~~mixed layer~~thermocline depth both in this paper and the supplementary material. Combined these results suggest important changes in the factors affecting lake biogeochemistry directly through changes in temperature and indirectly by influencing the availability of light and nutrients. By providing an initial test for the simulations that will be carried out by the ISIMIP lake sector this paper sets the stage for more extensive world-wide evaluation of the effects of climate change on lake thermal structure.

This study showed the ability of the GOTM model to simulate accurately Lake Erken water temperature when the model was forced using either daily or hourly temporal resolution inputs. Neural networks were shown to be an effective method to disaggregate different weather variables such as air temperature and short-wave radiation, in order to generate synthetic hourly meteorological data from the daily data that are typically available from GCM models. Model performance was slightly improved when using the synthetic hourly data, and climate change effects were somewhat lower when using such data to drive future climate simulations. However, it is not clear that data disaggregation is needed given the computational costs of developing such data sets and running long-term simulations at an hourly time step. Future climate simulations showed similar trends in the anomaly distributions when the GOTM model was forced with mean daily or synthetic hourly meteorological

data, and we also found evidence that the calibration procedure partly compensates for differences in the temporal resolution of the model input.

Code and data availability

The source code of the model GOTM is freely available online at https://gotm.net/. The input data, model configuration, output and observed data for calibration are stored in HydroShare at https://doi.org/10.4211/hs.ace98c3bc72b44f1834a58ec8b3af310. The ISIMIP climate scenarios are available online at https://www.isimip.org/. Future projections of simulated water temperature derived from both the original ISIMIP input data and synthetic hourly projections are stored in HydroShare at https://doi.org/10.4211/hs.2b4cfe0f02bf4375bcd0b62e45c61b19. Matlab codes, R codes and all the datasets produced during this study are available upon request from the corresponding author.

Author contributions

DCP and AIA designed the study. SM provided meteorological data. AIA performed GRNN models and GOTM simulations and analysed the results. AIA wrote the paper with contribution from DCP.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We are grateful to ISIMIP for their roles in producing, coordinating, and making available the ISIMIP climate scenarios, we acknowledge the support of the ISIMIP cross sectoral science team. We acknowledge funding from the EU and FORMAS in the frame of the collaborative international Consortium PROGNOS financed under the ERA-NET WaterWorks2014 Cofunded Call. This ERA-NET is an integral part of the 2015 Joint Activities developed by the Water Challenges for a Changing World Joint Programme Initiative (Water JPI). We also acknowledge funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement H2020-MSCA-ITN-2016 No 722518 for the Project MANTEL, and the project WATExR which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by MINECO (ES),FORMAS (SE), BMBF (DE), EPA (IE), RCN (NO), and IFD (DK),with co-funding by the European Union (Grant 690462) and FORMAS grant 2017-01738.

References

Bruce, L. C., Frassl, M. A., Arhonditsis, G. B., Gal, G., Hamil-ton, D. P., Hanson, P. C., Hetherington, A. L., Melack, J. M.,Read, J. S., Rinke, K., Rigosi, A., Trolle, D., Winslow, L., Adrian, R., Ayala, A. I., Bocaniov, S. A., Boehrer, B., Boon, C.,Brookes, J. D., Bueche, T., Busch, B. D., Copetti, D., Cortés,A., de Eyto, E., Elliott, J.A., Gallina, N., Gilboa, Y., Guyen-

non, N., Huang, L., Kerimoglu, O., Lenters, J.D., MacIntyre, S.,Makler-Pick, V., McBride, C. G., Moreira, S., Özkundakci, D.,Pilotti, M., Rueda, F. J., Rusak, J. A., Samal, N. R., Schmid, M.,Shatwell, T., Snorthheim, C., Soulignac, F., Valerio, G., van derLinden, L., Vetter, M., Vinçon-Leite, B., Wang, J., Weber, M.,Wickramaratne, C., Woolway, R. I., Yao, H., and Hipsey, M. R.: A multi-lake comparative analysis of the General Lake Model(GLM): Stress-testing across a global observatory network, Environ. Modell. Softw., 102, 274–291, https://doi.org/10.1016/j.envsoft.2017.11.016, 2018.

Bruggeman, J., and Bolding, K.: A general framework for aquatic biogeochemical models, Environ. Model Softw., 61, 249–265, https://doi.org/10.1016/j.envsoft.2014.04.002, 2014.

Burchard, H., Bolding, K., and Ruiz-Villarreal, M.: GOTM, a General Ocean Turbulence Model. Theory, implementation and test cases, Technical Report, 1999.

Butcher, J. B., Nover, D., Johnson, T. E., and Clark, C. M.: Sensitivity of lake thermal and mixing dynamics to climate change, Clim. Change, 129, 295-305, https://doi.org/10.1007/s10584-015-1326-1, 2015.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geosci. Model Dev., 10, 4321-4345, https://doi.org/10.5194/gmd-10-4321-2017, 2017.

Ficker, H., Luger, M., and Gassner, H.: From dimictic to monomictic: Empirical evidence of thermal regime transitions in three deep alpine lakes in Austria induced by climate change, Fresh. Biol., 62, 1335-1345, https://doi.org/10.1111/fwb.12946, 2017.

Foley, B., Jones, I.D., Maberly, S.C., and Rippey, B.: Long-term changes in oxygen depletion in a small temperate lake: effects of climate change and eutrophication., Freshwater Biol., 57, 278-289, https://doi.org/10.1111/j.1365-2427.2011.02662.x, 2012.

Grasset, C., Mendonça, R., Villamor Saucedo, G., Bastviken, D., Roland, F., and Sobek, S.: Large but variable methane production in anoxic freshwater sediment upon addition of allochthonous and autochthonous organic matter, Limnol. Oceanogr., 63, 1488-1501, https://doi.org/10.1002/lno.10786, 2018.

Guo, Z., Chang, C., Wang, R.: A Novel Method to Downscale Daily Wind Statistics to Hourly Wind Data for Wind Erosion Modelling, in: Geo-Informatics in Resource Management and Sustainable Ecosystem, GRMSE 2015, Wuhan, China, 16-18 October 2015, 611-619, 2015.

Hadley, K. R., Paterson, A. M. , Stainsby, E. A., Michelutti , N. , Yao, H. , Rusak, J. A., Ingram, R., McConnell, C., and Smol, J. P.: Climate warming alters thermal stability but not stratification phenology in a small north-temperature lake, Hydrol. Process., 28, 6309-6319, https://doi.org/10.1002/hyp.10120, 2013.

Hanrahan, J. L., Kravtsov, S. V., and Roebber, P. J.: Connecting past and present climate variability to the water levels of Lakes Michigan and Huron, Geophys. Res. Lett., 37, L01701, https://doi.org/10.1029/2009GL041707, 2010.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, Earth Syst. Dynam., 4, 219-236, https://doi.org/10.5194/esd-4-219-2013, 2013.

Huang, Y. T., and Liu, L.: Multiobjective water quality model calibration using a hybrid genetic algorithm and neural network-based approach, J. Environ. Eng., 136, 1020-1031, https://doi.org/10.1061/(ASCE)EE.1943-7870.0000237, 2010.

Idso, S. B.: On the concept of lake stability, Limnol. Oceanogr., 18, 681–683, 1973.

Kainz, M. J., Ptacnik, R., Rasconi, S., and Hager, H. H.: Irregular changes in lake surface water temperature and ice cover in subalpine Lake Lunz, Austria, Inland Waters, 7, 27-33, http://dx.doi.org/10.1080/20442041.2017.1294332, 2017.

Khatib, T., and Elmenreich, W.: A Model for Hourly Solar Radiation Data Generation from Daily Solar Radiation Data Using a Generalized Regression Artificial Neural Network, Int. J. Photoenergy, 968024,1–13, https://doi.org/10.1155/2015/968024, 2015.

Kirillin, G.: Modeling the impact of global warming on water temperature and seasonal mixing regimes in small temperate lakes~~.~~, Boreal Enviro. Res., 15, 279-293, 2010.

Kraemer, B. M., Anneville, O., Chandra, S., Dix, M., Kuusisto, E., Livingstone, D. M., Rimmer, A., Schladow, S. G., Silow, E., Sitoki, L. M., Tamatamah, R., Vadeboncoeur, Y., and McIntyre, P. B.: Morphometry and average temperature affect lake stratification responses to climate change, Geophys. Res. Lett., 42, 4981–4988, https://doi.org/10.1002/2015GL064097, 2011.

Ladwig, R., Furusato, E., Kirillin, G., Hinkelmann, R., and Hupfer, M.: Climate change demands adaptative management of urban lakes: model-based assessment of management scenarios for lake Tegel (Berlin, Germany), Water, 10, 186, https://doi.org/10.3390/w10020186, 2018.

MacKay, M. D., Neale, P. J., Arp, C. D., De Senarpont Domus, L. N., Fang, X., Gal, G., Jöhnk, K. D., Kirillin, G., Lenters, J. D., Litchman, E., MacIntyre, S., Marsh, P., Melack, J., Mooij, W. M., Peeters, F., Quesada, A., Schladow, S. G., Schmid, M., Spence, C., and Stokes, S. L.: Modeling lakes and reservoirs in the climate system, Limnol. Oceanogr., 54, 2315–2329, https://doi.org/10.4319/lo.2009.54.6_part_2.2315, 2009.

Magee, M. R., and Wu, C. H.: Response of water temperatures and stratification to changing climate in three lakes with different morphometry, Hydrol. Earth Syst. Sci., 21, 6253-6274, https://doi.org/10.5194/hess-21-6253-2017, 2017.

Markfort, C. D., Perez, A. L. S., Thill, J. W., Jaster, D. A., Porte-Agel, F., and Stefan, H. G.: Wind sheltering of a lake by a tree canopy or bluff topography~~.~~, Water Resour. Res. 46, 1-13, https://doi.org/10.1029/2009WR007759, 2010.

~~Mesman, J. P., Stelzer, J. A. A., Dakos,~~Martin, J., and McCutcheon, M.: Hydrodynamics and Transport for Water Quality Modeling, Lewis Publishers, US, 1999.

~~V., Goyette, S., Jones, I. D., Kasparian, J., McGinnis, D. F., and Ibelings, B. W.: The role of internal feedbacks in shifting deep lake mixing regimes under a warming climate, submitted to Glob. Chang. Biol., 2019.~~

Moras, S., Ayala, A. I., and Pierson, D. C.: Historical modelling of changes in Lake Erken thermal conditions, Hydrol. Earth Syst. Sci~~. Discuss.~~., 23, 5001–5016, https://doi.org/10.5194/hess-23-5001-2019~~-199, in review~~, 2019.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I - A discussion of principles~~.~~, J. Hydrol., 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nõges, T., Nõges, P., Jolma, A., and Kaitaranta, J.: Impacts of climate change on physical characteristics of lakes in Europe. European Commission Joint Research Centre Report EUR 24064 EN, Office for Official Publications of the European Communities, Luxembourg, 2009.

North, R. P., North, R. L., Livingstone, D. M., Köster, O., and Kipfer, R.: Long-term changes in hypoxia and soluble reactive phosphorus in the hypolimnion of a large temperate lake: consequences of a climate regime shift~~.~~, Glob. Chang. Biol., 20, 811-823, https://doi.org/10.1111/gcb.12371, 2013.

O'Reilly, C., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowle,y R. J., Schneider, P., Lenters, J. D., McIntyre, P.B., Kraemer, B. M., Weyhenmeyer, G. A., Straile, D., Dong, B., Adrian, R., Allan, M. G., Anneville, O., Arvola, L., Austin, J., Bailey, J. L., Baron, J. S., Brookes, J. D., de Eyto, E., Dokulil, M. T., Hamilton, D. P., Havens, K., Hetherington, A. L., Higgins, S. N., Hook, S., Izmest'eva, L. R., Joehnk, K. D., Kangur, K., Kasprzal, P., Kumagai, M., Kuusisto, E., Leshkevich, 20 G., Livingtone, D. M., McIntyre, S., May, L., Melack, J. M., Mueller-Navarra, D. C, Naumenko, M., Noges, P., Noges, T., North, R. P., Plisnier, P. D., Rigosi, A., Rimmer, A., Rogora, M., Rudstam, L. G., Rusak, J. A., Salmaso, N., Samal, N. R., Schindler, D. E., Schladow, S. G., Schmid, M., Schmidt, S. R., Silow, E., Soylu, M. E., Teubner, K., Verburg, P., Voutilainen, A., Watkinson, A., Wiliamson, C. E., and Zhang G.: Rapid and highly variable warming of lake surface waters around the globe, Geophys. Res. Lett., 42, 10773–10781, https://doi.org/10.1002/2015GL066235, 2015.

~~Shimoda, Y., Azim, M.E., Perhar, G., Ramin, M., Kenney, M.A., Sadraddini, S., Gudimov, A., and Arhonditsis, G.B.: Our current understanding of lake ecosystem response to climate change: What have we really learned from the north temperate deep lakes?, J. Great Lakes Res., 37, 173 193, https://doi.org/10.1016/j.jglr.2010.10.004, 2011.~~

O'Reilly, C. M., Alin, S. R., Plisnier, P. D., Cohen, A. S., and McKee, B. A.: Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa, Nature, 424, 766-768, https://doi.org/10.1038/nature01833, 2003.

Parton, W. J., and Logan, J. A.: A model for diurnal variation in soil and air temperature, J. Agric. Meteorol., 23, 205-2016, 1981.

Pearl, H.W. and Paul, V.J.: Climate change: links to global expansion of harmful cyanobacteria~~.~~, Water Res., 46, 1349-1363, https://doi.org/10.1111/j.1758-2229.2008.00004.x, 2012.

Perroud, M., and Goyette, S.: Impact of warmer climate on Lake Geneva water-temperature profiles, Boreal Environ. Res., 15, 255-278, 2010.

Perroud, M., Goyette, S., Martynov, A., Beniston, M., and Anneville, O.: Simulation of multiannual thermal profiles in deep Lake Geneva: A comparison of one-dimensional lake models, Limnol. Oceanogr., 54, 1574-1594, https://doi.org/10.4319/lo.2009.54.5.1574, 2009.

Persson,I., Jones,I., Sahlberg,J., Dokulil, M., Hewitt, D., Leppäranta, M., and Blenckner, T.: Modeled thermal response of three European lakes to a probable future climate, Verh Internat Verein Limnol., 29, 667–671, https://doi.org/10.1080/03680770.2005.11902762, 2005.

Persson, I., and Jones, I.D.: The effect of water colour on lake hydrodynamics: A modelling study, Freshwater Biol., 53, 2345-2355, https://doi.org/10.1111/j.1365-2427.2008.02049.x, 2008.

Pierson, D. C., Petterson, K., and Istvanovics, V.: Temporal changes in biomass specific photosynthesis during the summer: regulation by environmental factors and the importance of phytoplankton succession, Hydrobiologia, 243, 119-135, 1992.

Rall, B. C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O., and Petchey, O. L.: Universal temperature and body-mass scaling of feeding rates, Phil. Trans. R. Soc. B, 367, 2923–2934, https://doi.org/10.1098/rstb.2012.0242, 2012.

Read, J. S., Hamilton, D. P., Jones, I. D., Muraoka, K., Winslow, L. A., Kroiss, R., Wu, C. H., and Gaiser, E.: Derivation of lake mixing and stratification indices from high-resolution lake buoy data, Environ. Model Softw., 26, 1325-1336, https://doi.org/10.1016/j.envsoft.2011.05.006, 2011.

Rempfer, J., Livingstone D. M., Blodau, C., Niederhauser, P., Forster R., and Kipfer, R.: The effect of the exceptionally mild European winter of 2006–2007 on temperature and oxygen profiles in lakes in Switzerland: a foretaste of the future?, Limnol. Oceanogr., 55, 2170–2180, https://doi.org/10.4319/lo.2010.55.5.2170, 2010.

Sachse, R., Petzoldt, T., Blumstock, M., Moreira Martinez, S., Pätzig, M., Rücker, J., Janse, J., Mooij, W. M., and Hilt, S.: Extending one-dimensional models for deep lakes to simulate the impact of submerged macrophytes on water quality, Environ. Model Softw., 61, 410–423, https://doi.org/10.1016/j.envsoft.2014.05.023, 2014.

Sahoo, G. B., Forrest, A. L., Schladow, S. G., Reuter, J. E., Coats, R., and Dettinger, M.: Climate change impacts on lake thermal dynamics and ecosystem vulnerabilities, Limnol. Oceanogr., 61, 496–507, https://doi.org/10.1002/lno.10228, 2019.

Samal, N. R., Pierson, D. C., Schneiderman, E., Huang, Y., Read, J. S., Anandhi, A., and Owens, E. M.: Impact of climate change on Cannonsville Reservoir thermal structure in the New York City water supply, Water Qual. Res. J. Can., 47, 389–405, https://doi.org/10.2166/wqrjc.2012.020, 2010.

~~Schmid, M., Hunziker, S. and Wüest, A.: Lake surface temperatures in a changing climate: a global sensitivity analysis, Clim. Change, 124, 301-315. https://doi.org/10.1007/s10584-014-1087-2, 2014.~~

Schmidt, W.: Über Temperatur und Stabilitätsverhaltnisse von Seen, Geogr. Ann., 10, 145–177, 1928.

Schultze, M., Boehrer, B., Wendt-Potthoff, K., Katsev, S., and Brown, E. T.: Chemical Setting and Biogeochemical Reactions in Meromictic Lakes., Ecology of Meromictic Lakes, Springer, 35–59 pp., 2017.

Schwefel, R., Gaudard, A., Wüest, A., and Bouffard, D.: Effects of climate change on deepwater oxygen and winter mixing in a deep lake (Lake Geneva): Comparing observational findings and ~~modeling.~~modelling, Water Resour. Res., 52, 8811–8826, https://doi.org/10.1002/2016WR019194, 2016.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's Tau, J. Am. Stat. 5 Assoc., 63, 1379–1389, https://doi.org/10.1080/01621459.1968.10480934, 1968.

Shatwell, T., Thiery, W., and Kirillin, G.: Future projections of temperature and mixing regime of European temperate lakes, Hydrol. Earth Syst. Sci., 23, 1533–1551, https://doi.org/10.5194/hess-23-1533-2019, 2019.

Shimoda, Y., Azim, M.E., Perhar, G., Ramin, M., Kenney, M.A., Sadraddini, S., Gudimov, A., and Arhonditsis, G.B.: Our current understanding of lake ecosystem response to climate change: What have we really learned from the north temperate deep lakes?, J. Great Lakes Res., 37, 173–193, https://doi.org/10.1016/j.jglr.2010.10.004, 2011.

Storn, R., and Price K.: Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous 10 Spaces, J. Global Optim., 11, 341–359, https://doi.org/10.1023/A:1008202821328, 1997.

Theil, H.: A rank invariant method of linear and polynomial regression analysis, I, II, III, Proc. K. Ned. Akad. Wet., Ser. A Math. Sci., 53, 386–392, 1950.

Umlauf, L., and Burchard, H.: Second-order turbulence closure models for geophysical boundary layers. A review of recent work, Continental Shelf Research, 25,795–827. https://doi.org/10.1016/j.csr.2004.08.004, 2005.

Waichler, S. R., and Wigmosta, M. S.: Development of hourly meteorological values from daily data and significance to Hydrological Modeling at H. J. Andrews Experimental Forest, J. Hydrometeorol., 4, 251–263, https://doi.org/10.1175/1525-7541(2003)4<251:DOHMVF>2.0.CO;2, 2003.

Winslow, L. A., Hansen, G. J. A, Read J. S., and Notaro, M.: Large-scale modeled contemporary and future water temperature estimates for 10774 Midwestern U.S. Lakes, Sci. Data, https://doi.org/10.1038/sdata.2017.53, 2017.

Woolway, R.I., and Merchant, C.J.: Worldwide alteration of lake mixing regimes in response to climate change, Nature Geoscience, 12, 271–276, https://doi.org/10.1038/s41561-019-0322-x, 2019.

Woolway, R. I., Meinson, P., Nõges, P., Jones, I. D., and Laas, A.: Atmospheric stilling leads to prolonged thermal stratification in a large shallow polymictic lake, Clim. Change, 141, 759–773, https://doi.org/10.1007/s10584-017-1909-0, 2017.

Yang, Y., Colom, W., Pierson, D. C, and Pettersson, K.: Water column stability and summer phytoplankton dynamics in a temperate lake (Lake Erken, Sweden), Inland Waters, 6, 499–508, https://doi.org/10.1080/IW-6.4.874, 2016.

Yankova, Y., Neuenschwander, S., Köster, O., and Posch, T.: Abrupt stop of deep water turnover with lake warming: Drastic consequences for algal primary producers, Sci. Rep., 7, 13770, https://doi.org/10.1038/s41598-017-13159-9, 2017.

Tables

Table 1: Bias-corrected variables in the ISIMIP dataset

variable name	abbreviation	units
precipitation	pr	kg m^{-2} s^{-1}
surface pressure	ps	Pa
surface downwelling shortwave radiation	rsds	W m^{-2}
near-surface wind speed	sfcWind	m s^{-1}
near-surface air temperature	tas	K
daily maximum near-surface air temperature	tasmax	K
daily minimum near-surface air temperature	tasmin	K
relative humidity	hurs	%

Table 2: GOTM ~~l~~Lake model parameters and calibrated values

model parameter	feasible range	calibrated values		
		24 h met	1h met	synthetic 1h met
shf_factor	0.5–1.5	0.69	0.81	0.77
swr_factor	0.8–1.2	1.15	0.90	0.91
wind_factor	0.5–2.0	1.55	1.37	1.51
k_min	$1.4\ 10^{-7}$–$1.0\ 10^{-5}$	$1.47\ 10^{-6}$	$1.40\ 10^{-6}$	$1.29\ 10^{-6}$
g2	0.5–3.5	1.99	2.30	1.62

Table 3: GRNN models performance evaluation.

	training			validation		
GRNN model	MBE	RMSE	NSE	MBE	RMSE	NSE
Air temperature (ºC)	$-1.70\ 10^{-4}$	0.256	0.999	-0.057	0.318	0.940
Short wave radiation (W m^{-2})	$5.76\ 10^{-4}$	6.345	0.999	-0.037	8.196	0.870
Relative humidity (%)	$-7.94\ 10^{-4}$	0.790	0.998	0.341	1.021	0.686
Wind speed (m s^{-1})	$-5.67\ 10^{-3}$	1.060	0.779	-0.009	1.370	0.584

	MBE		RMSE		NSE	
	training	validation	training	validation	training	validation
air temperature (ºC)	$-1.70\ 10^{-4}$	-0.06	0.26	0.32	1.00	0.94
short wave radiation (W m^{-2})	$5.76\ 10^{-4}$	-0.04	6.35	8.20	1.00	0.87
relative humidity (%)	$-7.94\ 10^{-4}$	0.34	0.79	1.02	1.00	0.69
wind speed (m s^{-1})	$-5.67\ 10^{-3}$	-0.01	1.06	1.37	0.78	0.58

Table 4: LakeGOTM lake model performance evaluation: MBE, RMSE and NSE for full profiles temperature, surface temperature, bottom temperature, volumetrically weighted averaged whole lake temperatures, Schmidt stability and, thermocline, and duration, onset and loss of stratification using simulated results from running GOTM driven by daily (24h met), hourly (1h met) and synthetic hourly (synthetic 1h met) meteorological data sets.

	calibration								
	24h met			1h met			synthetic 1h met		
	MBE	RMSE	NSE	MBE	RMSE	NSE	MBE	RMSE	NSE
full-profile temp (ºC)	-0.08	1.04	0.93	-0.02	0.94	0.94	-0.02	0.88	0.95
surface temp (ºC)	-0.04	0.69	0.97	0.04	0.72	0.97	-0.01	0.61	0.98
bottom temp (ºC)	-0.06	1.33	0.83	0.07	1.24	0.85	-0.11	1.16	0.87
whole lake temp (ºC)	-0.07	0.57	0.98	-0.03	0.52	0.98	-0.01	0.49	0.98
Schmidt stability (J m^{-2})	0.53	22.09	0.85	0.59	21.69	0.85	0.76	19.64	0.88
thermocline depth (m)	0.58	2.77	0.32	0.84	3.07	0.22	0.43	2.84	0.32
	validation								
	24h met			1h met			synthetic 1h met		
	MBE	RMSE	NSE	MBE	RMSE	NSE	MBE	RMSE	NSE

~~full-profile temp (ºC)~~	~~-0.07~~	~~0.63~~	~~0.98~~	~~-0.12~~	~~0.69~~	~~0.97~~	~~0.00~~	~~0.68~~	~~0.97~~
~~surface temp (ºC)~~	~~-0.24~~	~~0.54~~	~~0.99~~	~~-0.19~~	~~0.64~~	~~0.98~~	~~-0.15~~	~~0.54~~	~~0.99~~
~~bottom temp (ºC)~~	~~0.16~~	~~0.68~~	~~0.96~~	~~0.09~~	~~0.59~~	~~0.97~~	~~0.23~~	~~0.74~~	~~0.95~~
~~whole lake temp (ºC)~~	~~-0.13~~	~~0.48~~	~~0.99~~	~~-0.17~~	~~0.59~~	~~0.98~~	~~-0.06~~	~~0.51~~	~~0.98~~
~~Schmidt stability (J m^{-2})~~	~~-5.26~~	~~13.27~~	~~0.90~~	~~-3.26~~	~~13.50~~	~~0.90~~	~~-4.47~~	~~13.26~~	~~0.90~~
~~thermocline depth (m)~~	~~0.89~~	~~2.86~~	~~0.07~~	~~1.07~~	~~3.27~~	~~-0.07~~	~~0.98~~	~~3.18~~	~~-0.14~~

calibration

	MBE			RMSE			NSE		
	24h met	1h met	synthetic 1h met	24h met	1h met	synthetic 1h met	24h met	1h met	synthetic 1h met
full-profile temp (ºC)	-0.08	-0.02	-0.02	1.04	0.94	0.88	0.93	0.94	0.95
surface temp (ºC)	-0.04	0.04	-0.01	0.69	0.72	0.61	0.97	0.97	0.98
bottom temp (ºC)	-0.06	0.07	-0.11	1.33	1.24	1.16	0.83	0.85	0.87
whole lake temp (ºC)	-0.07	-0.03	-0.01	0.57	0.52	0.49	0.98	0.98	0.98
Schmidt stability (J m^{-2})	0.53	0.59	0.76	22.09	21.69	19.64	0.85	0.85	0.88
thermocline depth (m)	0.58	0.84	0.43	2.77	3.07	2.84	0.32	0.22	0.32
duration (day)	0.25	3.75	6.63	9.25	14.25	14.94	-	-	-
onset (day)	-0.63	-0.50	-0.13	1.54	1.12	1.17	-	-	-
loss (day)	-0.63	-2.00	-2.88	1.54	5.87	9.13	-	-	-

validation

	MBE			RMSE			NSE		
	24h met	1h met	synthetic 1h met	24h met	1h met	synthetic 1h met	24h met	1h met	synthetic 1h met
full-profile temp (ºC)	-0.07	-0.12	0.00	0.63	0.69	0.68	0.98	0.97	0.97
surface temp (ºC)	-0.24	-0.19	-0.15	0.54	0.64	0.54	0.99	0.98	0.99
bottom temp (ºC)	0.16	0.09	0.23	0.68	0.59	0.74	0.96	0.97	0.95
whole lake temp (ºC)	-0.13	-0.17	-0.06	0.48	0.59	0.51	0.99	0.98	0.98
Schmidt stability (J m^{-2})	-5.26	-3.26	-4.47	13.27	13.50	13.26	0.90	0.90	0.90
thermocline depth (m)	0.89	1.07	0.98	2.86	3.27	3.18	0.07	-0.07	-0.14
duration (day)	-4.50	-3.50	-4.50	8.75	8.28	7.11	-	-	-
onset (day)	0.50	-7.50	0.50	0.71	10.61	0.71	-	-	-
loss (day)	-4.00	13.00	-4.00	8.94	15.26	7.21	-	-	-

~~Table 5. Projected trends of change (2006–2099) in air temperature,~~Table 5.Trend analysis from 2011-2100 for surface temperature, bottom temperature, whole-lake temperature, Schmidt stability, thermocline depth, duration, onset and loss of stratification (ns: not significant~~).) for RCP 6.0.

	~~RCP 2.6~~	
	~~GFDL-ESM2M~~	
	~~24-h met~~	~~1h met~~
~~air temperature~~	~~ns~~	~~ns~~
~~surface temperature~~	~~ns~~	~~ns~~
~~bottom temperature~~	~~ns~~	~~ns~~
~~whole-lake temperature~~	~~ns~~	~~ns~~
~~Schmidt stability~~	~~ns~~	~~ns~~
~~thermocline depth~~	~~ns~~	~~ns~~
~~duration~~	~~ns~~	~~ns~~
~~onset~~	~~ns~~	~~ns~~
~~loss~~	~~ns~~	~~ns~~
	~~HadGEM2-ES~~	
	~~24-h met~~	~~1h met~~
~~air temperature~~	~~0.18 ºC decade^{-1} (p-value < 0.001)~~	~~0.15 ºC decade^{-1} (p-value < 0.001)~~
~~surface temperature~~	~~0.16 ºC decade^{-1} (p-value < 0.001)~~	~~0.13 ºC decade^{-1} (p-value < 0.001)~~
~~bottom temperature~~	~~ns~~	~~ns~~
~~whole-lake temperature~~	~~0.10 ºC decade^{-1} (p-value < 0.001)~~	~~0.08 ºC decade^{-1} (p-value < 0.001)~~
~~Schmidt stability~~	~~2.60 J m^{-2} decade^{-1} (p-value < 0.001)~~	~~2.21 J m^{-2} decade^{-1} (p-value < 0.001)~~
~~thermocline depth~~	~~ns~~	~~ns~~
~~duration~~	~~1.01 day decade^{-1} (p-value = 0.014)~~	~~0.80 day decade^{-1} (p-value = 0.034)~~
~~onset~~	~~0.63 day decade^{-1} (p-value = 0.047)~~	~~ns~~
~~loss~~	~~ns~~	~~0.43 day decade^{-1} (p-value = 0.015)~~
	~~IPSL-CM5A-LR~~	
	~~24-h met~~	~~1h met~~
~~air temperature~~	~~0.11 ºC decade^{-1} (p-value = 0.005)~~	~~0.09 ºC decade^{-1} (p-value = 0.004)~~
~~surface temperature~~	~~0.11 ºC decade^{-1} (p-value = 0.003)~~	~~0.08 ºC decade^{-1} (p-value = 0.004)~~
~~bottom temperature~~	~~ns~~	~~ns~~
~~whole-lake temperature~~	~~0.06 ºC decade^{-1} (p-value = 0.038)~~	~~0.05 ºC decade^{-1} (p-value = 0.049)~~
~~Schmidt stability~~	~~2.52 J m^{-2} decade^{-1} (p-value = 0.012)~~	~~1.87 J m^{-2} decade^{-1} (p-value = 0.024)~~
~~thermocline depth~~	~~ns~~	~~ns~~

duration	1.73 day decade^{-1} (p-value < 0.001)	1.37 day decade-1 (p-value = 0.002)
onset	ns	0.93 day decade^{-1} (p-value = 0.004)
loss	0.85 day decade^{-1} (p-value = 0.005)	0.45 day decade^{-1} (p-value = 0.049)
	MIROC5	
	24 h met	1h met
air temperature	0.10 ºC decade^{-1} (p-value = 0.007)	0.08 ºC decade^{-1} (p-value = 0.008)
surface temperature	0.09 ºC decade^{-1} (p-value = 0.001)	0.07 ºC decade^{-1} (p-value = 0.002)
bottom temperature	ns	ns
whole-lake temperature	0.08 ºC decade^{-1} (p-value < 0.001)	0.06 ºC decade^{-1} (p-value < 0.001)
Schmidt stability	ns	ns
thermocline depth	ns	ns
duration	ns	ns
onset	ns	0.60 day decade^{-1} (p-value = 0.018)
loss	ns	ns
	RCP 6.0	
	GFDL-ESM2M	
	24 h met	1h met
air temperature	0.20 ºC decade^{-1} (p-value < 0.001)	0.15 ºC decade ^{1} (p-value < 0.001)
surface temperature	0.17 ºC decade^{-1} (p-value < 0.001)	0.13 ºC decade^{-1} (p-value < 0.001)
bottom temperature	ns	ns
whole-lake temperature	0.15 ºC decade^{-1} (p-value < 0.001)	0.11 ºC decade^{-1} (p-value < 0.001)
Schmidt stability	2.50 J m^{-2} decade^{-1} (p-value = 0.003)	1.89 J m^{-2} decade^{-1} (p-value = 0.008)
thermocline depth	ns	ns
duration	ns	0.81 day decade^{-1} (p-value = 0.031)
onset	ns	ns
loss	ns	ns
	HadGEM2-ES	
	24 h met	1h met
air temperature	0.44 ºC decade^{-1} (p-value < 0.001)	0.33 ºC decade^{-1} (p-value < 0.001)
surface temperature	0.38 ºC decade^{-1} (p-value < 0.001)	0.28ºC decade^{-1} (p-value < 0.001)
bottom temperature	0.07 ºC decade^{-1} (p-value = 0.010)	ns
whole-lake temperature	0.25 ºC decade^{-1} (p-value < 0.001)	0.17 ºC decade^{-1} (p-value < 0.001)
Schmidt stability	7.79 J m^{-2} decade^{-1} (p-value < 0.001)	6.22 J m^{-2} decade^{-1} (p-value < 0.001)

thermocline depth	0.12 m decade^{-1} (p-value < 0.001)	0.12 m decade^{-1} (p-value < 0.001)
duration	3.55 day decade^{-1} (p-value < 0.001)	2.94 day decade^{-1} (p-value < 0.001)
onset	1.90 day decade^{-1} (p-value < 0.001)	1.41 day decade^{-1} (p-value < 0.001)
loss	1.80 day decade^{-1} (p-value < 0.001)	1.38 day decade^{-1} (p-value < 0.001)
	IPSL-CM5A-LR	
	24 h met	1h met
air temperature	0.43 ºC decade^{-1} (p-value < 0.001)	0.33 ºC decade^{-1} (p-value < 0.001)
surface temperature	0.39 ºC decade^{-1} (p-value < 0.001)	0.29 ºC decade^{-1} (p-value < 0.001)
bottom temperature	0.08 ºC decade^{-1} (p-value = 0.033)	ns
whole-lake temperature	0.27 ºC decade^{-1} (p-value < 0.001)	0.19 ºC decade^{-1} (p-value < 0.001)
Schmidt stability	7.93 J m^{-2} decade^{-1} (p-value < 0.001)	6.14 J m^{-2} decade^{-1} (p-value < 0.001)
thermocline depth	0.08 m decade^{-1} (p-value = 0.003)	0.08 m decade^{-1} (p-value < 0.001)
duration	3.06 day decade^{-1} (p-value < 0.001)	2.39 day decade^{-1} (p-value < 0.001)
onset	1.87 day decade^{-1} (p-value < 0.001)	1.38 day decade^{-1} (p-value < 0.001)
loss	1.26 day decade^{-1} (p-value < 0.001)	1.02 day decade^{-1} (p-value < 0.001)
	MIROC5	
	24 h met	1h met
air temperature	0.32 ºC decade^{-1} (p-value < 0.001)	0.25 ºC decade^{-1} (p-value < 0.001)
surface temperature	0.28 ºC decade^{-1} (p-value < 0.001)	0.21ºC decade^{-1} (p-value < 0.001)
bottom temperature	0.11 ºC decade^{-1} (p-value < 0.001)	0.09 ºC decade^{-1} (p-value = 0.002)
whole-lake temperature	0.22 ºC decade^{-1} (p-value < 0.001)	0.17 ºC decade^{-1} (p-value < 0.001)
Schmidt stability	3.99 J m^{-2} decade^{-1} (p-value < 0.001)	2.77 J m^{-2} decade^{-1} (p-value = 0.005)
thermocline depth	ns	ns
duration	2.19 day decade^{-1} (p-value < 0.001)	1.78 day decade^{-1} (p-value < 0.001)
onset	1.71 day decade^{-1} (p-value = 0.002)	1.39 day decade^{-1} (p-value = 0.017)
loss	0.69 day decade^{-1} (p-value < 0.001)	0.43 day decade^{-1} (p-value < 0.001)

		RCP 6.0			
		24h met		1h met	
		rate (decade^{-1})	p-value	rate (decade^{-1})	p-value
surface temperature (ºC)	GFDL-ESM2M	0.16	< 0.001	0.12	< 0.001
	HadGEM2-ES	0.38	< 0.001	0.27	< 0.001
	IPSL-CM5A-LR	0.38	< 0.001	0.29	< 0.001

	MIROC5	0.29	< 0.001	0.22	< 0.001
bottom temperature (ºC)	GFDL-ESM2M		ns		ns
	HadGEM2-ES	0.06	< 0.05		ns
	IPSL-CM5A-LR		ns		ns
	MIROC5	0.11	< 0.001	0.09	< 0.01
whole-lake temperature (ºC)	GFDL-ESM2M	0.14	< 0.001	0.10	< 0.001
	HadGEM2-ES	0.25	< 0.001	0.16	< 0.001
	IPSL-CM5A-LR	0.26	< 0.001	0.19	< 0.001
	MIROC5	0.23	< 0.001	0.18	< 0.001
Schmidt stability (J m^{-2})	GFDL-ESM2M	2.69	< 0.01	1.92	< 0.01
	HadGEM2-ES	7.97	< 0.001	6.50	< 0.001
	IPSL-CM5A-LR	8.15	< 0.001	6.36	< 0.001
	MIROC5	4.23	< 0.01	2.93	< 0.01
thermocline depth (m)	GFDL-ESM2M	0.07	< 0.05		ns
	HadGEM2-ES	0.13	< 0.001	0.13	< 0.001
	IPSL-CM5A-LR	*0.05*	*0.06*	0.09	< 0.01
	MIROC5		ns		ns
duration (days)	GFDL-ESM2M		ns		ns
	HadGEM2-ES	3.56	< 0.001	3.08	< 0.001
	IPSL-CM5A-LR	3.16	< 0.001	2.50	< 0.001
	MIROC5	2.45	< 0.001	2.00	< 0.001
onset (day)	GFDL-ESM2M		ns		ns
	HadGEM2-ES	-1.95	< 0.001	-1.43	< 0.001
	IPSL-CM5A-LR	-1.98	< 0.001	-1.48	< 0.001
	MIROC5	-1.80	< 0.001	-1.45	< 0.001
loss (day)	GFDL-ESM2M		ns		ns
	HadGEM2-ES	1.83	< 0.001	1.42	< 0.001
	IPSL-CM5A-LR	1.31	< 0.001	1.06	< 0.001
	MIROC5	0.83	< 0.001	0.52	< 0.01

825

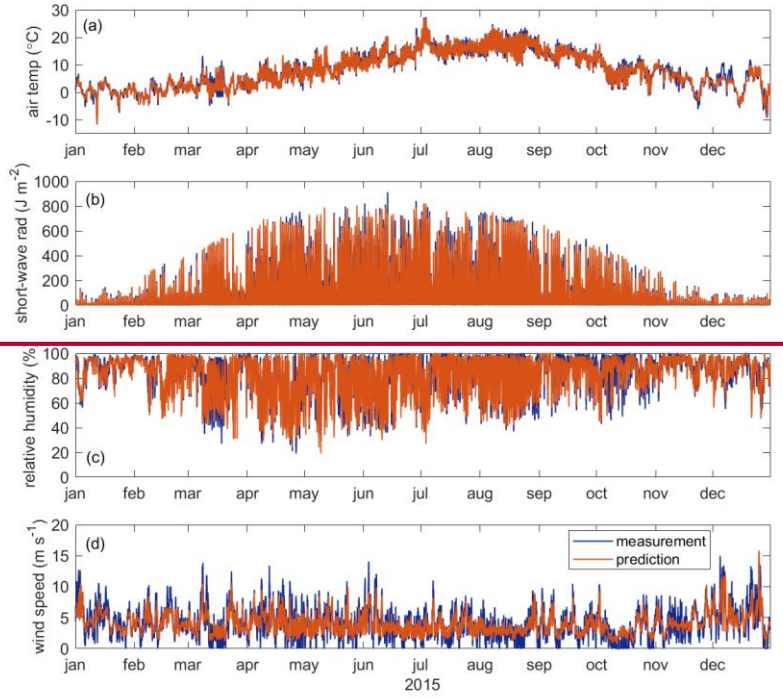
Table 6. Average thermal metrics for reference period (1981-2010), and average projected change in thermal metrics for mid-century and late-century for RCP 6.0.

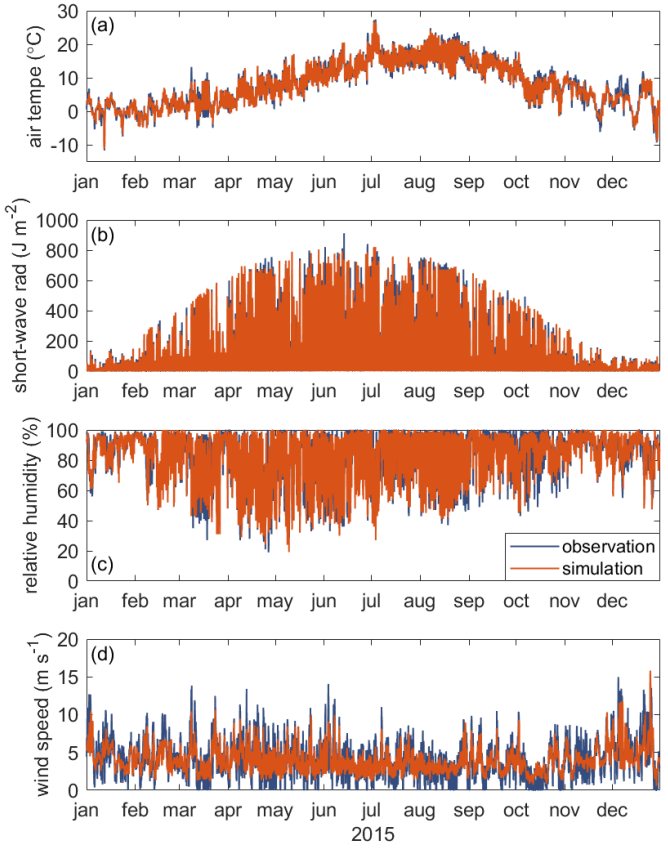
		RCP 6.0					
		24h met			1h met		
		reference period	mid-century	late century	reference period	mid-century	late century
surface temperature (ºC)	GFDL-ESM2M	13.71	1.03	1.67	13.89	0.81	1.28
	HadGEM2-ES	13.56	3.04	4.04	13.84	2.29	2.98
	IPSL-CM5A-LR	13.64	2.60	3.62	13.86	1.92	2.69
	MIROC5	13.41	1.98	2.97	13.69	1.52	2.28
	ensemble	13.58	2.16	3.08	13.82	1.64	2.31
bottom temperature (ºC)	GFDL-ESM2M	9.23	0.94	1.24	9.66	0.68	0.90
	HadGEM2-ES	9.32	0.42	0.91	9.75	0.15	0.42
	IPSL-CM5A-LR	9.29	1.18	1.19	9.61	0.88	0.79
	MIROC5	8.94	0.98	1.18	9.34	0.77	0.90
	ensemble	9.19	0.88	1.13	9.59	0.62	0.75
whole-lake temperature (ºC)	GFDL-ESM2M	12.44	1.06	1.61	12.83	0.79	1.20
	HadGEM2-ES	12.44	1.96	2.81	12.89	1.45	1.98
	IPSL-CM5A-LR	12.36	2.12	2.75	12.72	1.57	1.99
	MIROC5	12.11	1.69	2.41	12.59	1.27	1.84
	ensemble	12.34	1.71	2.39	12.76	1.27	1.75
Schmidt stability (J m^{-2})	GFDL-ESM2M	69.90	4.94	12.26	65.42	4.50	9.79
	HadGEM2-ES	66.43	59.78	77.57	63.18	48.50	61.43
	IPSL-CM5A-LR	67.52	38.67	64.62	65.73	28.06	49.23
	MIROC5	68.96	23.08	42.42	66.39	17.49	31.83
	ensemble	68.20	31.62	49.22	65.18	24.64	38.07
thermocline depth (m)	GFDL-ESM2M	-7.82	-0.39	-0.22	-8.50	-0.17	-0.08
	HadGEM2-ES	-8.23	1.02	1.26	-8.77	0.98	1.23
	IPSL-CM5A-LR	-7.83	0.28	0.59	-8.26	0.24	0.64
	MIROC5	-7.83	0.09	0.34	-8.51	0.28	0.49
	ensemble	-7.93	0.25	0.49	-8.51	0.33	0.57
duration (days)	GFDL-ESM2M	126	-9	-8	129	-10	-9
	HadGEM2-ES	123	-8	-8	126	-9	-9
	IPSL-CM5A-LR	123	-8	-8	126	-9	-8

	MIROC5	124	-8	-8	128	-9	-9
	ensemble	124.08	-11	-11	127	-12	-12
onset (day)	GFDL-ESM2M	131	-10	-10	131	-11	-11
	HadGEM2-ES	133	-10	-10	133	-11	-11
	IPSL-CM5A-LR	133	-10	-10	133	-11	-11
	MIROC5	134	-11	-10	133	-11	-11
	ensemble	133	-65	-58	133	-61	-56
loss (day)	GFDL-ESM2M	257	-7	11	263	-15	-2
	HadGEM2-ES	255	-29	-3	258	-38	-16
	IPSL-CM5A-LR	260	-46	-27	263	-49	-35
	MIROC5	257	-37	-19	260	-41	-27
	ensemble	257	7	8	261	8	8

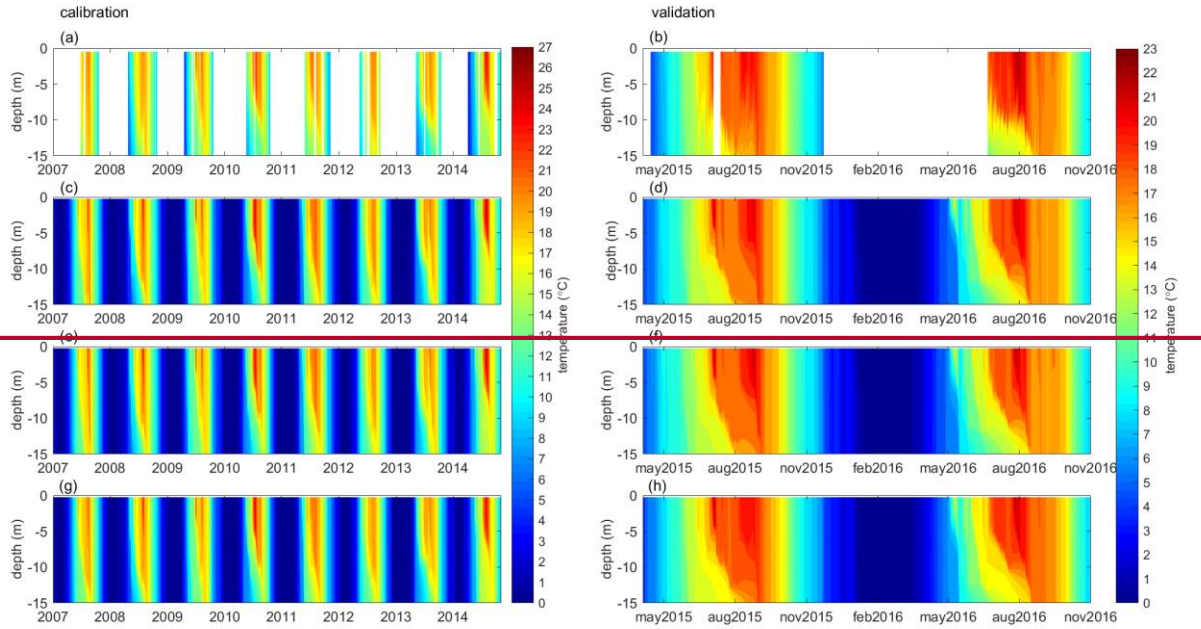
830

Figures


(a)
air temp (°C)
30
20
10
0
-10
jan feb mar apr may jun jul aug sep oct nov dec
(b)
short-wave rad (J m-2)
1000
800
600
400
200
0
jan feb mar apr may jun jul aug sep oct nov dec
(c)
relative humidity (%
100
80
60
40
20
0
jan feb mar apr may jun jul aug sep oct nov dec
(d)
wind speed (m s-1)
20
15
10
5
0
measurement
prediction
jan feb mar apr may jun jul aug sep oct nov dec
2015



835 **Figure 1. ~~Measured~~GRNN temporal disaggregation of meteorological forcing data. Observations vs ~~predicted~~simulations (a) air temperature, (b) short-wave radiation, (c) relative humidity and (d) wind speed for 2015 (validation data set).**


calibration
validation
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
depth (m)
0
-5
-10
-15
2007
2008
2009
2010
2011
2012
2013
2014
may2015
aug2015
nov2015
feb2016
may2016
aug2016
nov2016
temperature (°C)

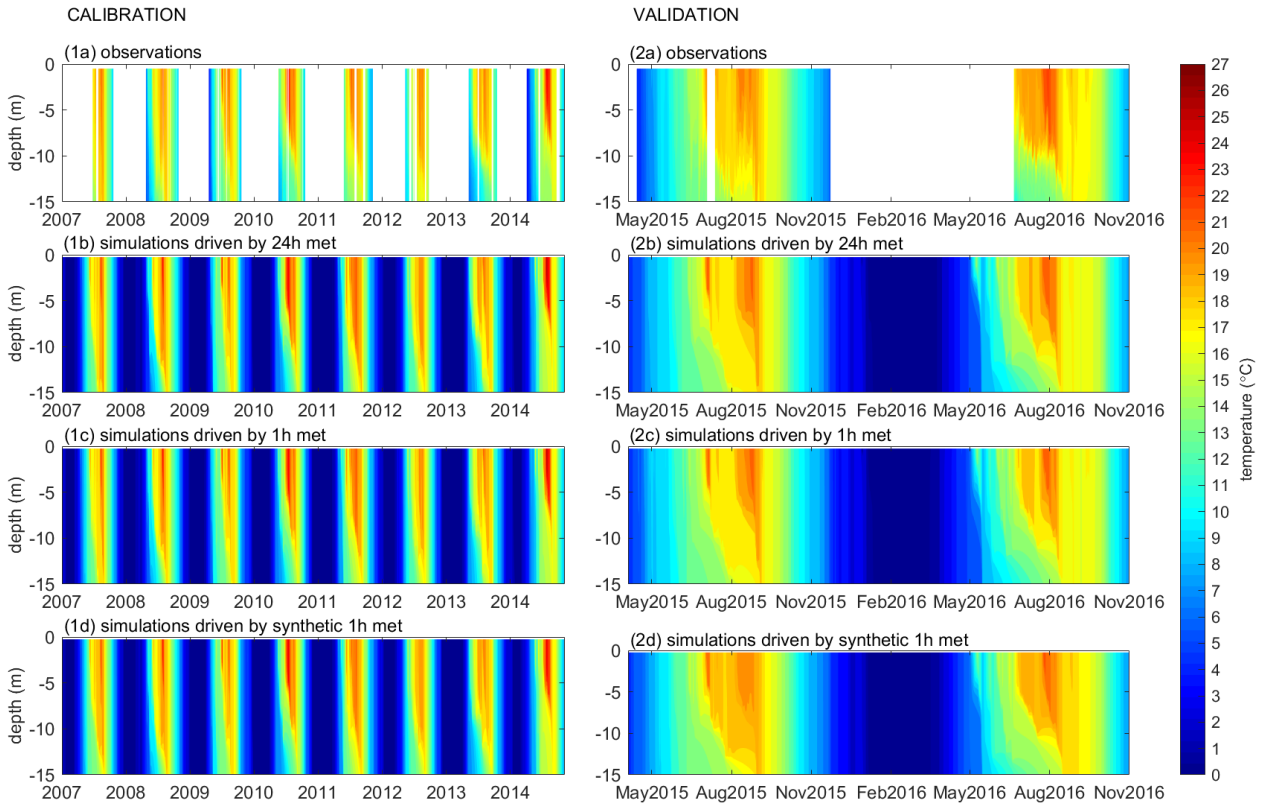
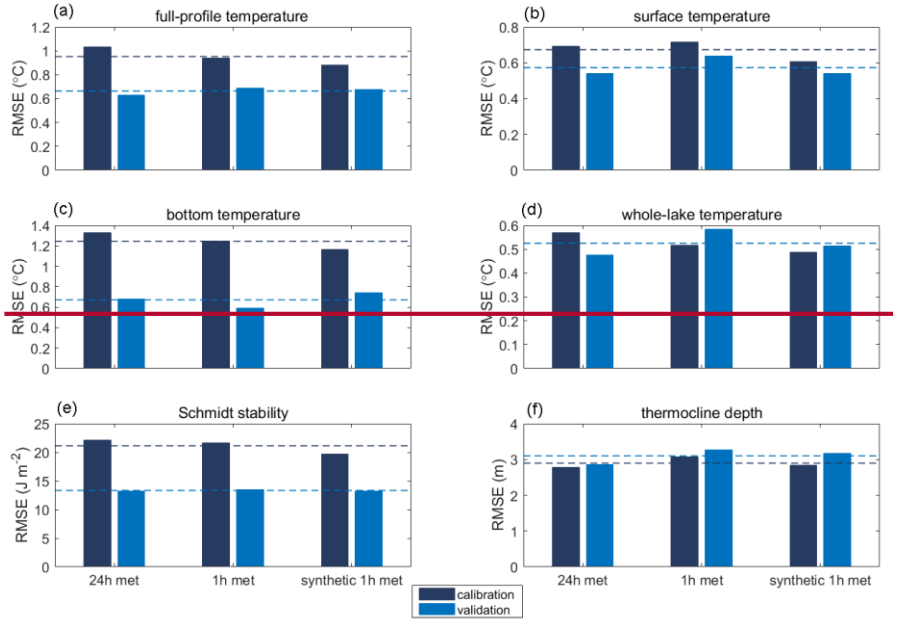
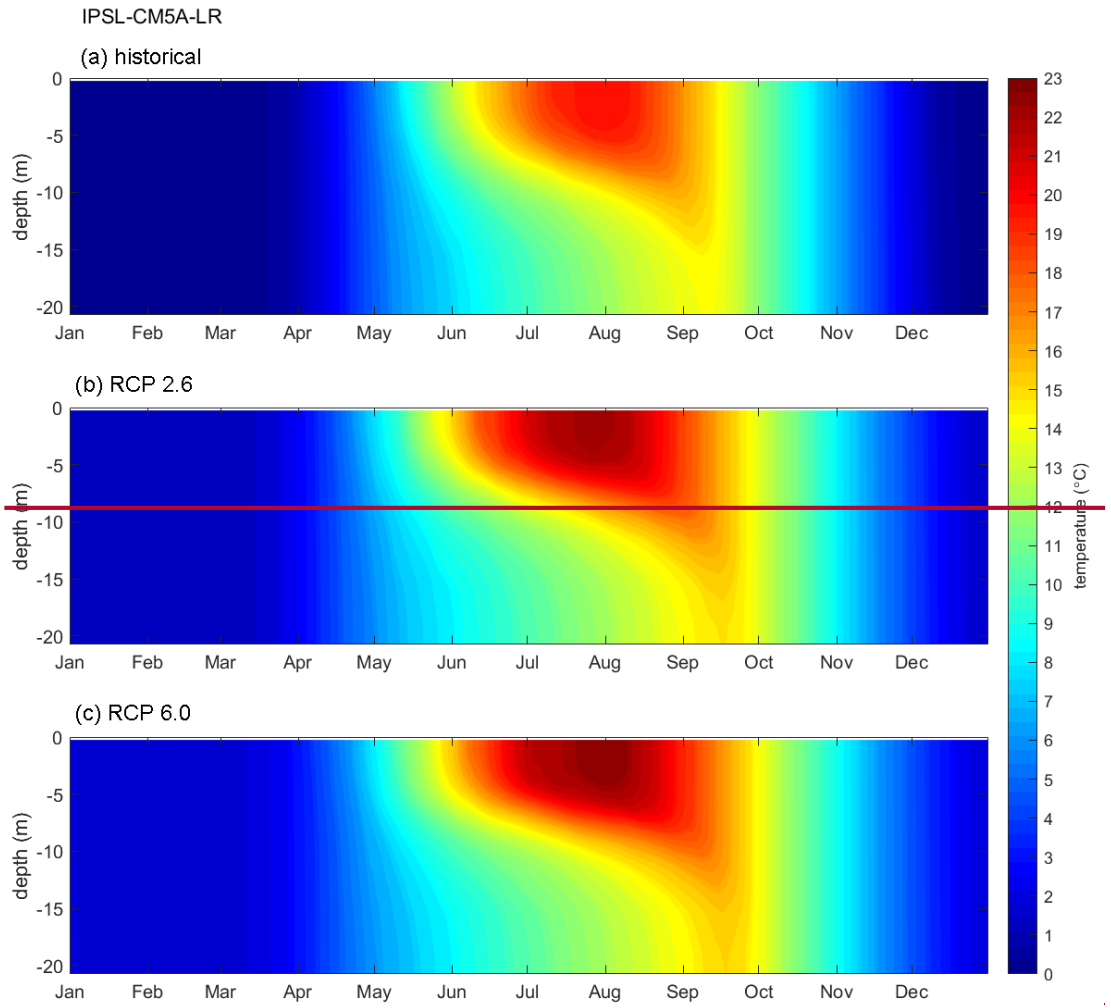


Figure 2. GOTM water temperature simulations. Daily averaged water temperature in Lake Erken for the ~~validation (a) (c) (e) (g) and~~ calibration (~~b) (d) (f) (h~~1a)-(1b)-(1c)-(1d) and validation (2a)-(2b)-(2c)-(2d) periods: observations (~~a) (b~~1a)-(2a), simulations driven by daily meteorological data (~~c) (d~~1b)-(2b), hourly meteorological data (~~e) (f~~1c)-(2c) and synthetic hourly meteorological data (~~g) (h~~1d)-(2d).



845 **~~Figure 3. GOTM model performance metrics for prediction of (a) full-profile temperature which compared simulated and measured data at all possible depths, (b) surface temperature, (c) bottom temperature, (d) whole-lake temperature, (e) Schmidt stability and (f) thermocline depth. The mean (horizontal line) is also shown~~**


IPSL-CM5A-LR
(a) historical
(b) RCP 2.6
(c) RCP 6.0
depth (m)
0
-5
-10
-15
-20
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
temperature (°C)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

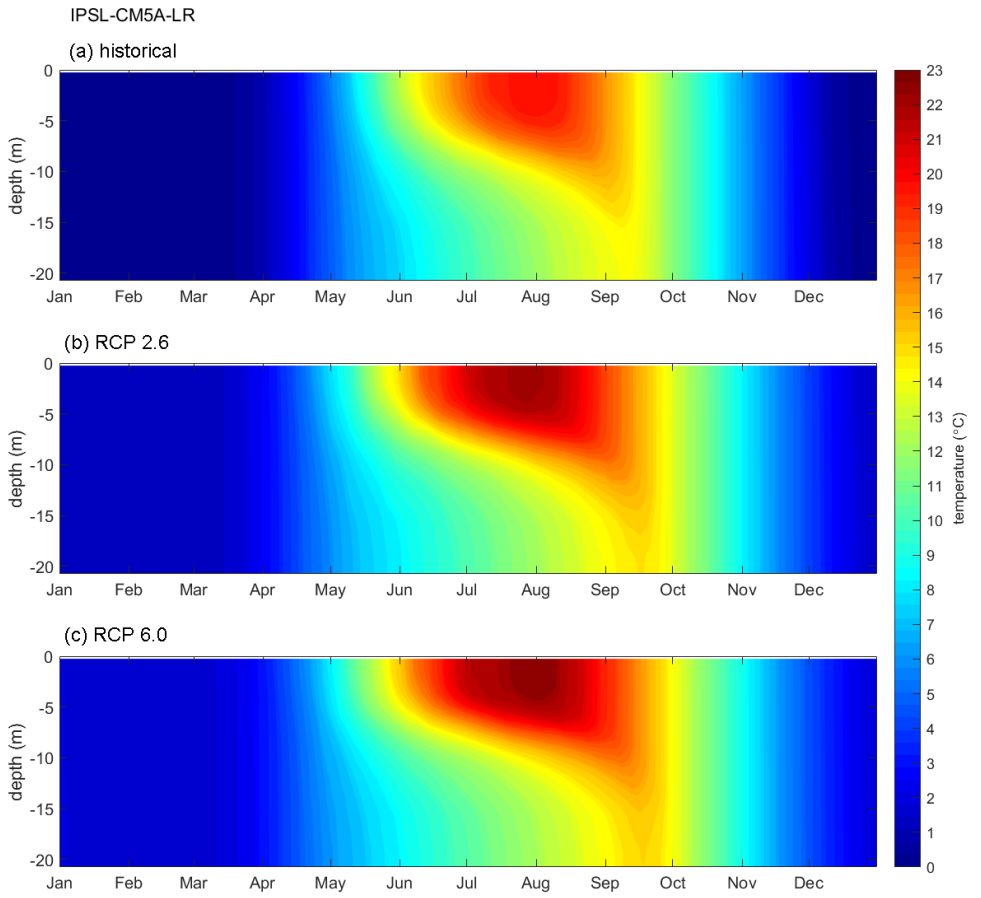
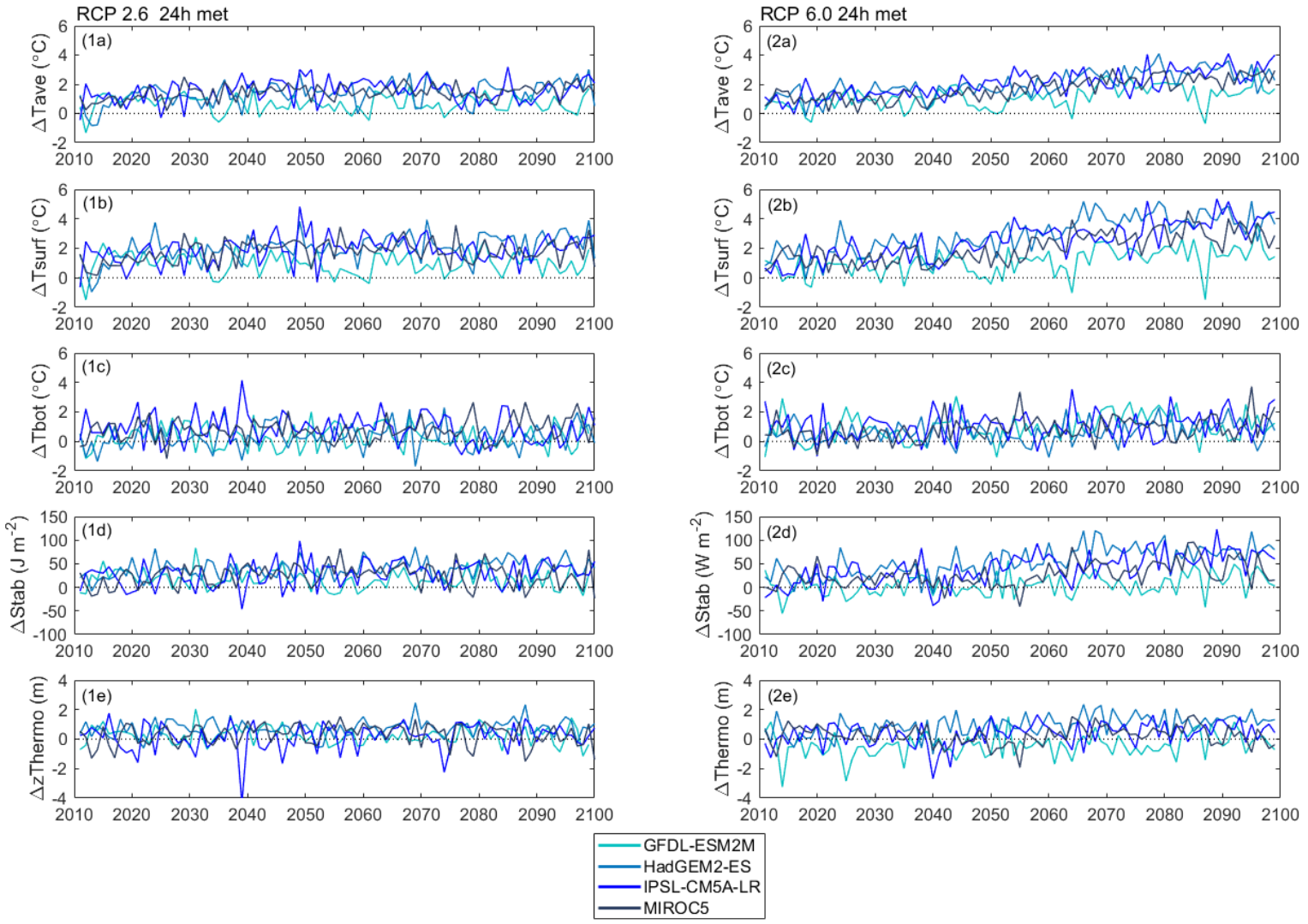


Figure 3.~~Figure 4.~~ Temperature isopleth diagrams for the (a) historical, (b) RCP 2.6 and (c) RCP 6.0 scenarios showing results from the lake model forced with daily IPSL-CM5A-LR projections. The temperature matrix used to make these plots was created by averaging the simulated daily temperature profiles for every year in each scenario.


RCP 2.6 24h met
RCP 6.0 24h met
(1a)
(2a)
(1b)
(2b)
(1c)
(2c)
(1d)
(2d)
(1e)
(2e)
ΔTave (°C)
ΔTsurf (°C)
ΔTbot (°C)
ΔStab (J m-2)
ΔStab (W m-2)
ΔzThermo (m)
ΔThermo (m)
GFDL-ESM2M
HadGEM2-ES
IPSL-CM5A-LR
MIROC5

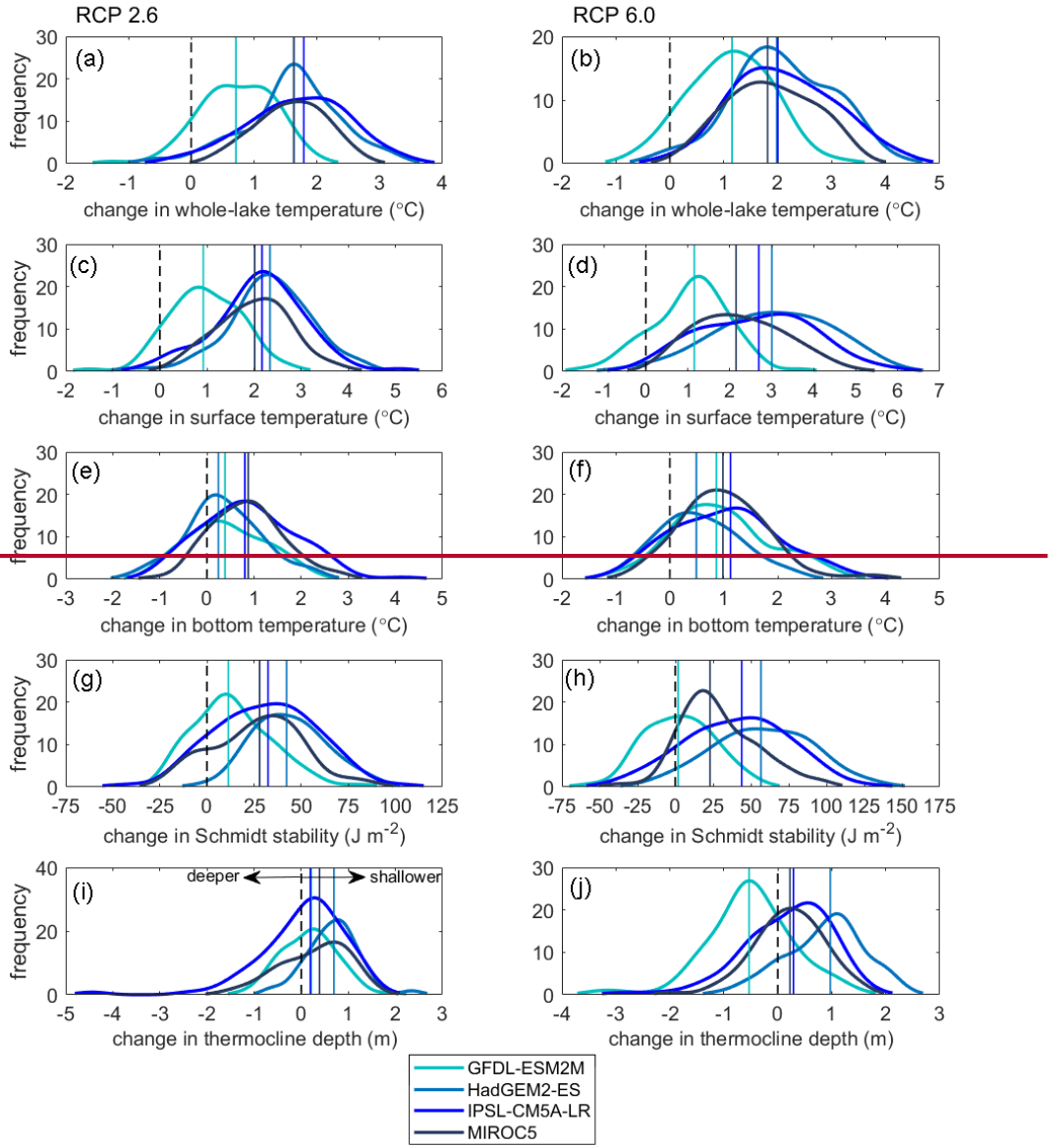
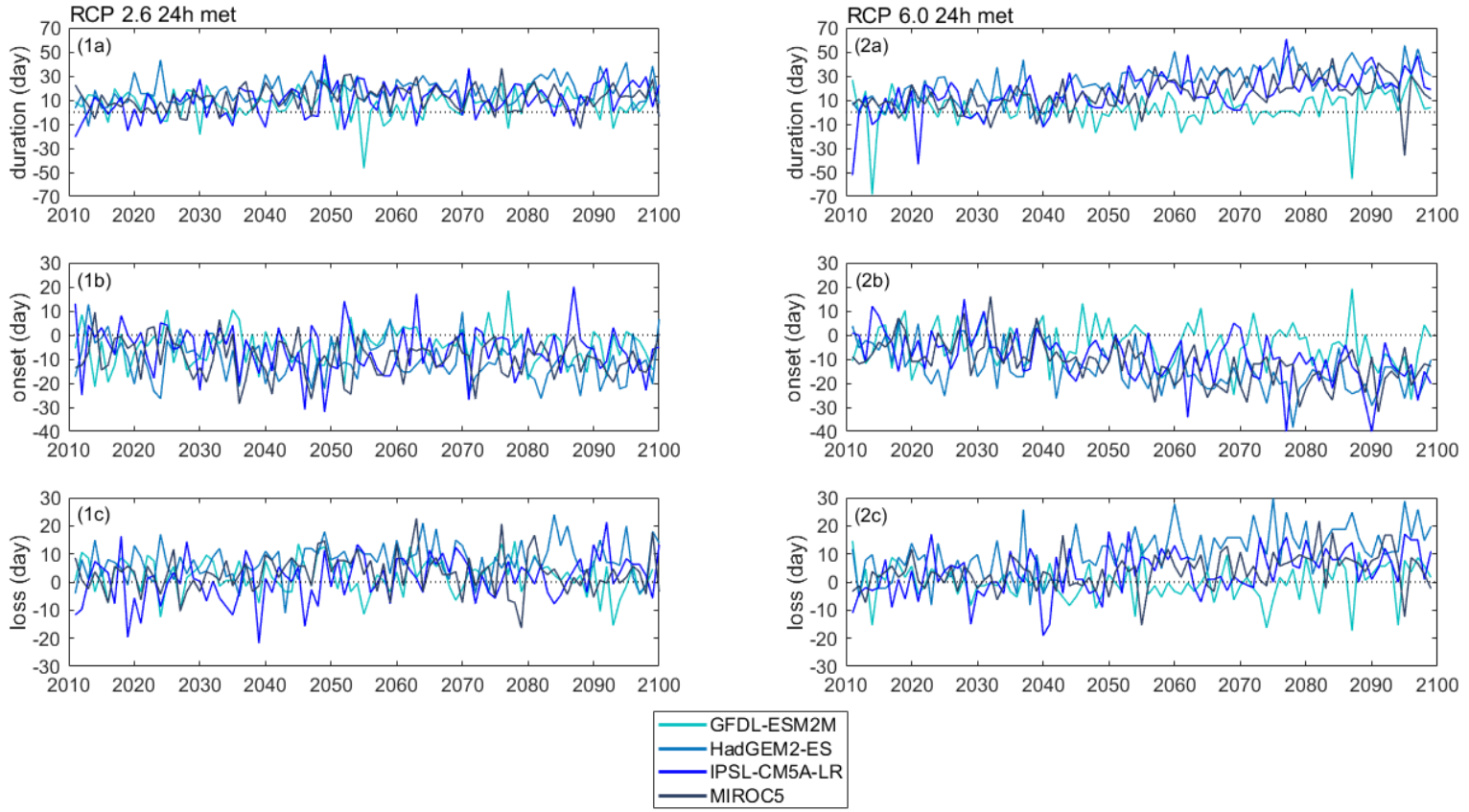

RCP 2.6
RCP 6.0
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
frequency
change in whole-lake temperature (°C)
change in surface temperature (°C)
change in bottom temperature (°C)
change in Schmidt stability (J m-2)
change in thermocline depth (m)
deeper
shallower
GFDL-ESM2M
HadGEM2-ES
IPSL-CM5A-LR
MIROC5

Figure 4.

Evolution of annual average projected ~~Figure 5. Changes in~~ anomalies ~~calculated from annually averaged~~ (from April to September) ~~(a) (b~~ for (1a)-(2a) whole-lake temperature, (~~c) (d~~1b)-(2b) surface temperature, (~~e) (f~~1c)-(2c) bottom temperature, (~~g) (h~~1d)-(2d) Schmidt stability~~, (i) (j~~ and (1e)-(2e) thermocline depth ~~under (a) (c) (e) (g) (i) RCP 2.6 and (b) (d) (f) (h) (j) RCP 6.0,~~ showing results ~~from~~when the lake model was forced with daily GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR ~~-~~and MIROC5 projections~~. All changes are for 2006-2099~~ from 2011 to 2100 under RCP 2.6 and 6.0. Anomalies are relative to ~~1975-2005. The median (vertical line) is also shown.~~reference period (1981-2010).



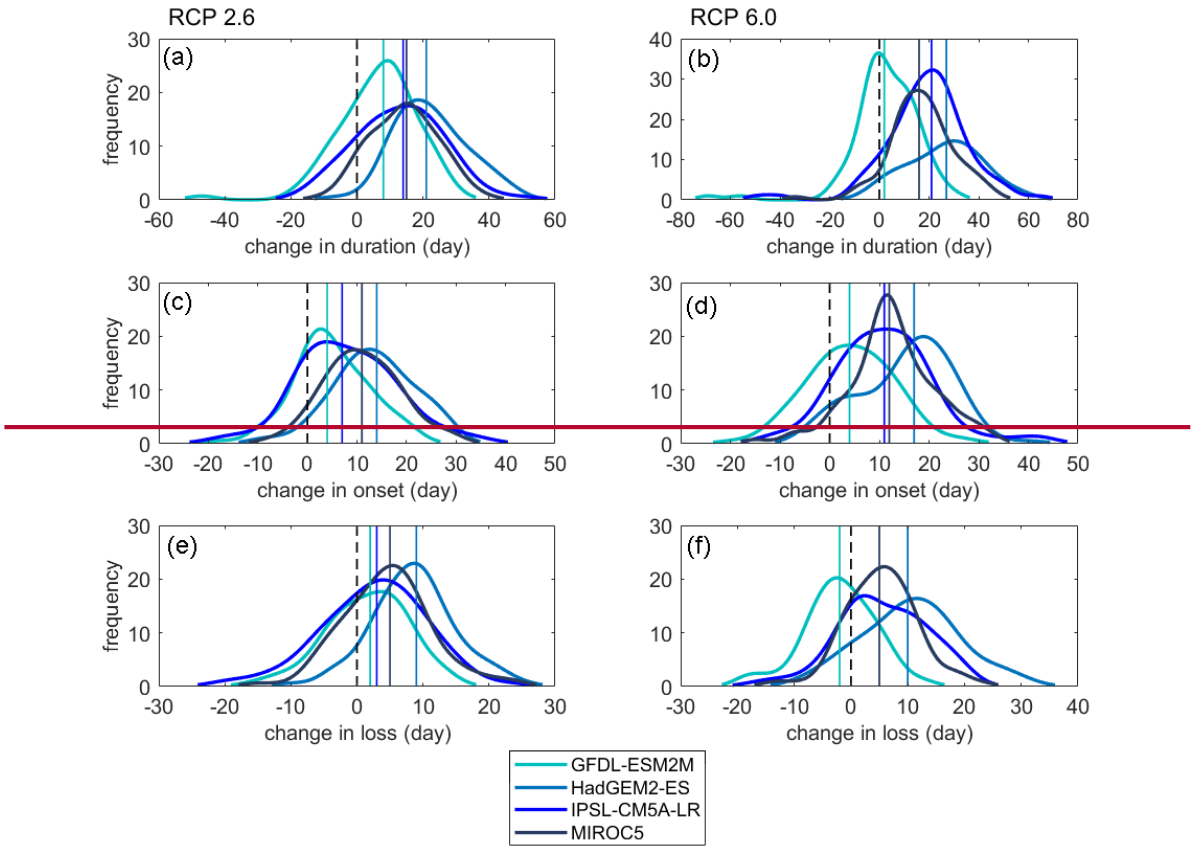


Figure 5.

Evolution of~~Figure 6. Changes in the calculated~~ annual average projected anomalies ~~of the (a) (b~~(from April to September) for (1a)-(2a) duration, (~~c) (d~~1b)-(2b) onset and (~~e) (f~~1c)-(2c) loss of stratification ~~under (a) (c) (e) RCP 2.6 and (b) (d) (f) RCP 6.0,~~ showing results ~~from~~when the lake model was forced with daily GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 projections from 2011 to 2100 under RCP 2.6 and 6.0. Anomalies are relative to reference period (1981-2010).

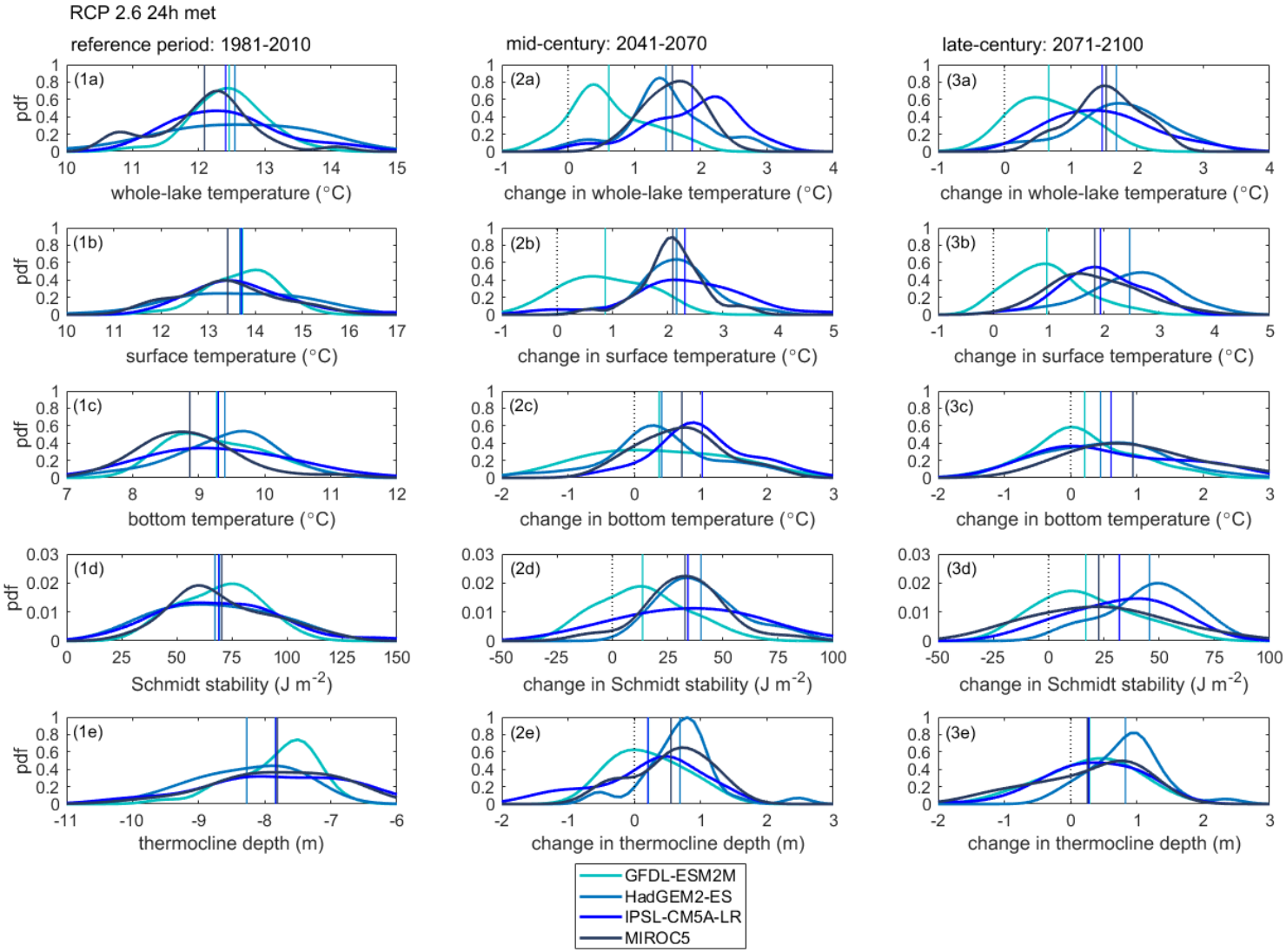


Figure 6. - Changes in annually averaged thermal metrics (from April to September) (2a)-(3a) whole-lake temperature, (2b)-(3b) surface temperature, (2c)-(3c) bottom temperature, (2d)-(3d) Schmidt stability and (2e)-(3e) thermocline depth under RCP 6.0, showing results when the lake model was forced with daily GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 projections. All changes are for mid-century (2041-2070) and late-century (2071-2100) are relative to reference period (1981-2011). The mean (vertical line) is also shown. Changes in thermal metrics greater than 0 show an increase and lower than 0 show a decrease.

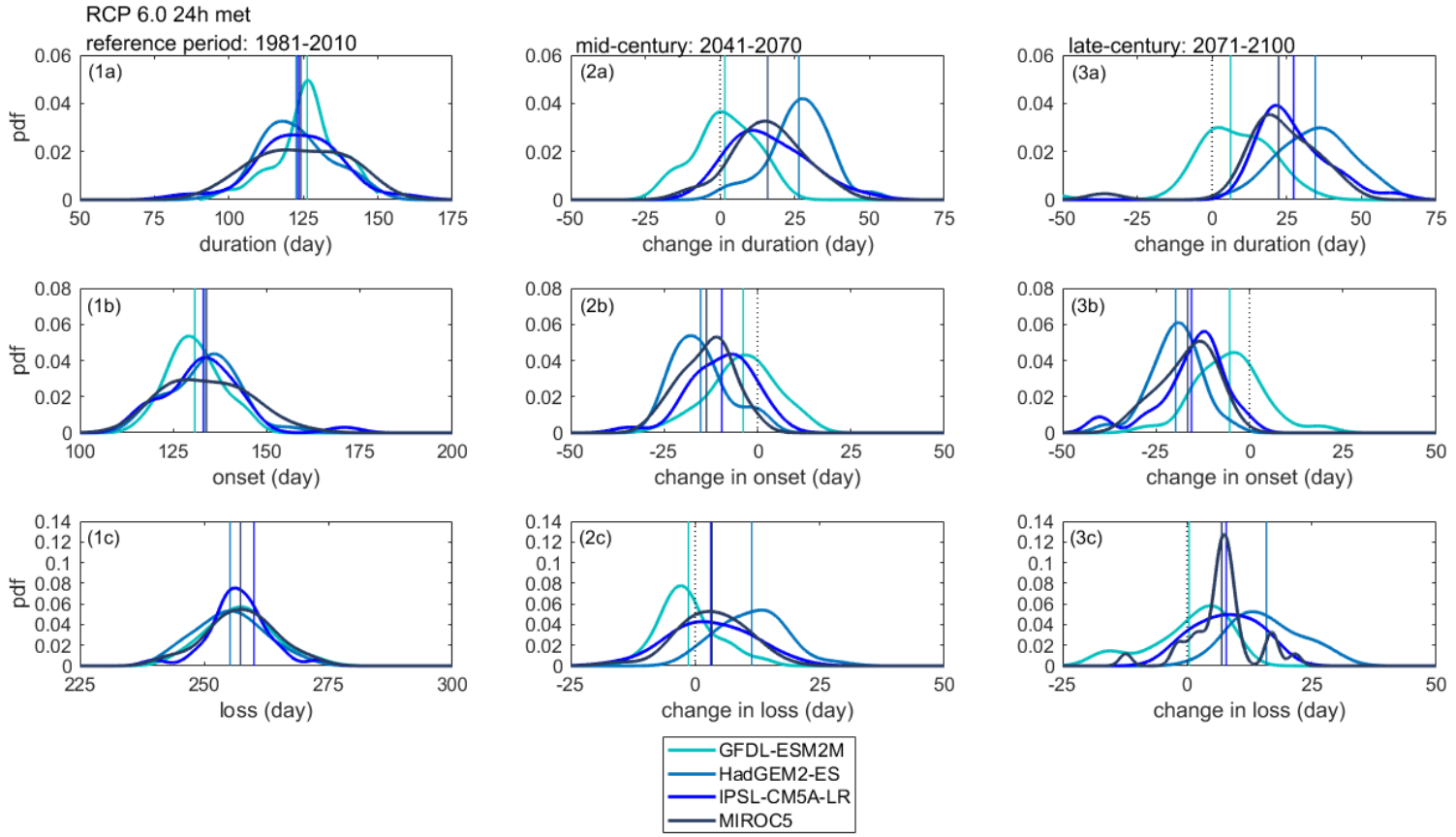


Figure 7. Changes in annually averaged thermal metrics (from April to September) (2a)-(3a) duration, (2b)-(3b) onset and (2c)-(3c) loss of stratification under RCP 6.0, showing results when the lake model was forced with daily GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR and MIROC5 projections. All changes are for ~~2006-2099 relative to 1975-2005.~~mid-century (2041-2070) and late-century (2071-2100) are relative to reference period (1981-2011). The ~~median~~mean (vertical line) is also shown. Changes in thermal metrics greater than 0 show and increase and lower than 0 show a decrease.