Revision of the manuscript “Historical modelling of changes in Lake Erken thermal conditions” by Moras S, Ayala A.I. and Pierson D.C.

Changes after manuscript revision

We revised the manuscript following referees’ suggestions. The major changes in the manuscript are the following:

- After noticing a small error in the model calibration in the first manuscript version (observed water temperature data of April and November were not taken into account during calibration) we updated all the tables and figures with the new results. This update, however does not change the main outcomes of the study.
- The introduction has been rewritten by adding references to studies that analyzed with historical lake water temperature data. Moreover, we better addressed the aim of the study.
- We better clarified the model limitation in simulating ice cover in the method section.
- We added the analysis of observed ice cover duration of Lake Erken between 1941-2017. This analysis was not present in the first version of the manuscript.
- We reported the results of the model performance in different season to evaluate if the model performs differently in different seasons.
- We added a paragraph discussing the model performance in the discussion section
- Three more tables have been added to the manuscript. One table shows how much data we retrieved from the different meteorological station for each meteorological parameter we used to drive the model. Another table shows the model performance of the model on seasonal basis. The third table we added is a table of results of the lake metrics we investigated.
- All the figures have been update with the new results and we followed referees’ suggestion to improve the readability of the figures.
- We added a new section in the supplementary material that describe the mismatch between modelled and observed water temperature in winter

For the detailed changes in the manuscript, please the the marked up version of the manuscript below.

Authors’ response to Referee 1

We would like to thank Referee 1 for the valuable comments he provided for our manuscript, that contribute to improve the quality of our work. See below detailed answers to the comments.

Referee 1 – Page 3, line 10. Daily precipitation was used in driving the model, while the other six datasets were put into the model as hourly resolution. This sounds strange to me. Are different climate variables allowed to put into the model with different temporal resolution?

Authors’ response. The seven climatic parameters used to drive the models are grouped into three input datasets: a meteo_file, which contains wind speed, air pressure, air temperature, cloud cover and relative humidity data; a swr_file in which shortwave radiation data are stored and a precip_file where precipitation data are located. Within the same dataset, the parameters must have the same time-resolution, but it is possible each of the three datasets to differ in time-resolution. GOTM model allows to set a factor that converts the unit of measurement used in the the precip_file input (in our case mm/day) into the unit of measurement used in GOTM for precipitation (m/s). This possibility gave us the chance to use the most suitable time resolution for precipitation in our study, since no weather station around Lake Erken measured precipitation on hourly basis. For our long-term simulations we presented in our paper, we assume a constant water level. Therefore, precipitation had only minor effects on the model output.

Referee 1 – Page 4, line 30. Why the measured water temperatures with 30 minutes resolution were averaged to daily, not the hourly mean values for the model calibration? In this way, the diurnal variation of the water temperature is missing. Could you give an explanation here?
Authors' response. This is a good point and we are aware that using hourly values for model calibration would have taken into account the diurnal variation of water temperature. Our choice to average 30 minutes water temperature to daily values have been made by the fact that a calibration using hourly values was computationally too intensive. We set ACPy to run 10000 simulations to obtain the best parameter set. We calibrated the model using a daily water temperature dataset of 94244 data points. This process takes ~24 hours using daily values. The use of hourly data for model calibration would have been a very time-consuming process. In addition, most of the metrics of change in thermal structure used in our paper were most conveniently calculated using mean daily data. Therefore, we felt that it would be most appropriate to develop model calibration based on mean daily output.

Referee 1 – Page 5, line 3. I am afraid the wind factor of 1.28 is a little bit high, since wind is measured in or quite close to the lake (based on Figure 1). Could you explain why you use such a high wind factor here?
Authors' response. There are two possible explanations here. First, the dominant wind speed (ws) direction is along the longest east-west fetch of Lake Erken that is ~10 km as opposed to the north – south fetch that is only 2-3 km. The 1D model input for wind is only a mean velocity and does not account for the effects of fetch. Given that wind is often blowing along the longest fetch that would have that would have the greatest effect on the measured temperature measurements used for calibration at the Eastern end of the lake, it is reasonable to expect an elevated wind factor. Secondly, it is actually the wind speed cubed that is used in the model equations that effect turbulent mixing. Under variable and gusty conditions cubing the mean hourly wind speed calculated by our data logger measuring at 1 minute intervals

$$\left(\frac{\sum_{60}^{1} ws}{60}\right)^3$$

may underestimate the true effects of wind which would more properly be calculated as the mean of of all cubed wind speed measurements made during the hour.

$$\left(\frac{\sum_{60}^{1} ws^3}{60}\right)$$

This effect would also result in an elevated wind factor.

Referee 1 – Page 6, line 1. how did you define the thermocline depth in the study? As I know, there are two ways in defining the thermocline depth in rLakeAnalyzer (i.e. seasonal=TRUE/FALSE). The results, from the two approaches, are different (see " Read, J. S., Hamilton, D. P. P., Jones, I. D., Muraoka, K., Winslow, L. A., Kroiss, R., Wu, C. H. & Gaiser, E. (2011). Derivation of lake mixing and stratification indices from high resolution lake buoy data. Environmental Model ling and Software 26:1325 1336 ")").
Authors' response. We did not specify which condition I used to define thermocline depth in our R code. However, not specifying any condition as we did gives the same result of the condition "seasonal = TRUE".

Referee 1 – Page 9, line 4. As stronger evidence for such changing trend, could you also use the measured water temperature to do a Mann Kendall test? In the paper, all the statistical test s are based on the simulated temperature, it is better to prove the simulated trend also based on the temperature, it is better to prove the simulated trend also based on the measured values. If it takes you so much time to do this work for all the three cases (i.e whole lake, epilimnion and hyplimnion), I recommended to test the observed trend for the summer epilimnion because the simulated temperatures of the layer significantly increased in the whole period.
Authors' response. Even though Lake Erken has a relatively long measured water temperature record compared to other lakes, there are still significant data gaps within the dataset. There were several years with no (or very few) measured

temperature before the deployment of the automatic floating station in 1988. There are significant data gaps in Erken temperature record after 1988 as well, during the maintenance/failures of the floating station for example. Since our trend analysis is based on seasonal means, performing a trend analysis on measured water temperature with several missing data would have made our results unrealistic. Having such data gaps in our water temperature record is actually the main reason why we developed the approach described in this study in order to get a more consistent and reliable water temperature historical record using a hydrodynamic model.

Referee 1 – Page 12, line 7. I am confused here, you said that the summer epilimentic temperature significantly increased for the whole period, but not significantly increased in two sub intervals? To me, it sounds like a paradox. Please check it.

Authors' response. When Mann-Kendall test is performed on the two sub-intervals (1961-1988 and 1989-2017) of summer epilimnetic temperature, positive trends are detected but they are not significant. This means that the two sub-intervals are too short to detect a significant trend. Indeed, when the trend test is performed on the entire study period (1961-2017) the summer epilimnetic temperature shows a significant increasing trend. From our results, we can infer that summer epilimnetic temperature was subjected to a slower but more stable warming compared to, for example, spring and autumn epilimnetic temperature, which showed a more abrupt increase in water temperature in the most recent sub-interval (1989-2017).

Referee 1. Also, as shown in Blenckner 2002, Lake Erken is always ice covered for the whole winter and the ice melts between March and early May. It is a weak point to use GOTM, without an ice module, to simulate such a lake with a long ice duration. I suggest adding some sentences, in this part, to clarify this limitation. Considering the future model development, it is a valuable work to include ice part into GOTM which could also be added into the Discussion.

Authors' response. GOTM developers are currently working on integrating GOTM with an ice module, but this was not available for this work. The GOTM model used for the simulations documented here did not have a functioning ice model, but instead cut off surface heat exchange when the simulated surface water temperature became negative. This provided a very simple way to make continuous simulations that include freezing conditions that would normally lead to the formation of ice. However, the temperature profiles during winter were not realistic, and could not be used for model calibration. This can be seen in figures 1-2 (below) where a comparison between simulated an observed water temperature at 1m and 15 m depth is reported for year 2009. At 1 m depth, simulated and observed temperature are rather similar throughout the entire year. However, at 15 m depth, the model does not take into account the heat loss from sediment during ice-cover, which cause an increase in bottom water temperature. During winter, there is a clear mismatch between simulated and observed water temperature. For this reason, all data collected between 1 December - 31 March are excluded from the temperature data used for model calibration and only data between 1 April and 30 November are used for model calibration. Yours is a valuable comment and we better clarified this limitation in our Methods and Discussion.

Authors' response to Referee 2

We would like to thank Referee 2 for the comments provided on our manuscript. We added three of the references (Vincent 2009; Skowron 2017; Sadro et al., 2019) proposed by the referee in our revised manuscript. Moreover, we improved the X and Y axis description following the referee's suggestions.

Authors' response to Referee 3

We would like to thank Referee 3 for the valuable comments and criticisms on the manuscript. The detailed comments provided will be certainly useful to improve the overall quality of this work. Before answering to the specific comments, however, we would like to better clarify our vision of the manuscript. We do not agree with the referee on the fact that our work represents only a case study application. On the contrary, we described an effective methodology that is able to reconstruct historical lake water temperature that can be applied and extended to many other lakes, not only Lake Erken specifically. Therefore, we believe that this study advances scientific progress. We also think, however, that the specific comments provided by the referee are extremely helpful to better elucidate the general purpose of this work. Please, see below for our responses to the specific comments.

Referee (1) – Page 2, lines 5-8. I do not understand the claim that these studies do not use observations to validate their models. This is simply not true. Stefan et al., 1998, include a section called model adequacy tests; Taner et al., 2011, state that they use a previously calibrated model (validation results not shown); and Winslow et al., 2017, include a section called technical validation.

Indeed it would be extremely bad practice to parameterize a model without calibration-validation. Please make your intentions and motivations for the study here much clearer.

This would also be improved by expanding the current introduction. At present there are only two paragraphs which cover very little literature. There is a need to root this work within the wider research.

Authors' response (1) – Page 2, lines 5-8. This is bad wording on our part. We did not mean to imply that the mentioned studies did not validate their models. What we were trying to say was that these studies have focused on simulating future changes in lake thermal structure (with model validation to present conditions) while in this paper we are advocating for running simulations farther back in time than is normally done for validation purposes in order to provide evidence that climate change has already affected lake thermal structure.

However, the introduction has been rewritten and the references to these papers have been removed. The new first paragraph of the introduction is the following:

"*Changes in the thermal structure and mixing regimes of lakes are connected to changes in several climatic factors such as air temperature, solar radiation, cloud cover, wind speed and humidity (Woolway and Merchant, 2019). The alteration of lake hydrodynamic properties has consequences on lake chemistry, biology and ultimately on the ecosystem services that lakes provide (Adrian et al., 2009; Vincent, 2009). Since climatic conditions have changed markedly in the last century and they are expected to change considerably in the next decades (IPCC, 2013), the importance of evaluating how freshwater bodies are affected by climate change becomes evident. A direct assessment of how lakes have already been affected by climate change is to analyse historical trends in lake water temperature data. However, the availability of long-term data of lake water temperature is still scarce. For example, there are very few lakes around the world with a long-term record (defined here as >50 years) of water temperature profiles (e.g. Jankowski et al., 2006; Skowron, 2017). Instead, the availability of long-term historical data (>50 years) is often limited to surface water temperature of one or few lakes (e.g. Livingstone and Dokulil, 2001; Kainz et al., 2017) and the time frame of surface temperature data available for the majority of lakes is limited to 2-3 decades. For example, Sharma et al. (2015) compiled a worldwide database with lake surface water temperature between 1985-2009. The same time frame was used by Schneider and Hook (2010) that reported an average warming trend of 0.045 ± 0.011°C/year of lake surface water temperature in 167 large lakes (>500 km2) using satellite-derived measurements; similarly*

O'Reilly et al. (2015) reported an average warming trend of 0.34 °C/decade for lake summer surface water temperature in 235 lake worldwide retrieved from both in-situ and satellite data. Even though these studies have demonstrated a rapid warming trend among lakes, the analysis of only surface water temperature is not sufficient to obtain a complete evaluation of the changes in the thermal structure that encompass, for example, temperature trends in the water column and phenology of thermal stratification. Moreover, the scarcity of water temperature data before 1980s it difficult to assess earlier thermal conditions for the majority of lakes. A longer record of historical data (> 50 years) provides more background information, allows better documentation of the changes that have already taken place, and leads to more accurate predictions of lake thermal conditions in future decades. One of the best arguments to counter climate change sceptics is well documented long-term records of the ongoing effects of climate change."

Referee (2) – Page 4, lines 1-22. How much data is actually missing from the dataset for each parameter? A table, or similar, detailing the quantity and quality of the data would be extremely helpful.

Why are the additional sites only considered when there is missing data? Are the limited stations used truly representative of conditions across the entire lake? Would the coverage and overall consistency of the observed data not be improved if the same data was used at all times? Please clearly justify your decision-making here - deliberately excluding valid data is problematic.

Similarly, please indicate the locations of the additional stations on your map. It would be useful if the reader could understand the locations of these additional stations relative to the three detailed.

Authors' response (2) – Page 4, lines 1-22. A detailed description of the number of missing data is available in the supplementary material (tables 1-4). We put these tables in the supplementary material for a better readability of the paper. However, we added a summary table in the manuscript (table 1) with the number of data retrieved from different station for each parameter.

Our meteorological data are either collected from a small island (fig1a, letter B) in the lake or from a meteorological station only a few hundred meters from the lake shore (fig 1a, letter B). Given the station locations, we considered this ideal data for forcing a lake model and it was our assumption that these data should be used when available. When data were missing, we found that the neural network models made use of as many of the surrounding data sources as possible providing the most accurate replacement values. We do not believe that we were excluding valid data, based on our belief (and we suspect a widely held belief) that locally collected data would be most appropriate for modeling. Data from additional sites was only used as a substitute when the most valid data were not available. We added a map showing the location of the additional stations in the manuscript (fig.1)

Referee (3) – Page 5, lines 13-15; 21-24. The authors appear to use the same data for calibration-validation. Why is this? Please justify - the standard is to employ a split-sampling approach. Further, the aim is to minimize the variance in the GOF statistics across the calibration-validation period. Without defined periods, you cannot determine the consistency of the model performance.

L21 - What is meant by best? Was the algorithm run multiple times? How is the best one determined when three GOF statistics are used? Please clarify.

You introduce figures 2-4 but provide no further commentary on these. There is no clear discussion with regards to how this indicates good performance. Indeed, you do not refer to your GOF statistics through these figures at all. The reporting of the model performance needs to be significantly expanded. Please also consider reporting model performance per month and/or season - this may help to give insights into whether the model performs worse immediately following the ice-cover period.

Please also note that Figures 3 and 4 do not actually add anything to the reporting of model performance - they give no indication of the GOF of the model. Additionally, the use of inconsistent x-y scales is bad practice and misleading. If producing the figures in R then it is possible to fix the axes across plots/facets.

As a more minor comment - it is not necessary to define the three equations, tehy are standard mathematical equations. What is more important is to explain why these are relevant - what insight does using these GOF statistics provide?

Authors' response (3) – Page 5, lines 13-15; 21-24. We agree that for typical applications of models where the goal is to make simulations to future or otherwise different conditions than are covered by the record of measured calibration data it is appropriate to employ a split calibration and validation strategy. However in our case the goal was not to simulate outside of the period of available calibration data, but to use the model to provide a complete and consistent record over a period in which calibration data were available but incomplete (especially in the earlier part of the record). In such a case we believe it is better to make full use of all measured calibration data rather than removing some for a separate validation run. This should ensure that the calibration encompasses the widest possible range of variability and provides parameter values that are most appropriate for the entire period simulated in our study.

L21 - In the ACPy calibration the best set of parameter is calculated by minimizing the log likelihood function. We have now added a specific reference to this in the text of the manuscript.

We reported the model performance in figures 3-4.

The use of inconsistent x-y scales on fig. 3-4 are now. Now all years are shown in the figures.

We removed the standard mathematical equation in the revised version.

Referee (4) – Page 5, lines 26-28. Please consider expanding on the limitation of ice-cover - perhaps in the discussion? For example, it would be helpful to suggest how this might be addressed, being unable to account for almost six months of the year is problematic. Similarly, this should be acknowledged in the section where you describe GOTM. For instance, why do you not simply use a model that does include an ice-cover module given the length of time the lake is ice-covered?

Authors' response (4) – Page 5, lines 26-28. The GOTM model used for the simulations documented here did not have a functioning ice model, but instead cut off surface heat exchange when the simulated surface water temperature became negative. This provided a very simple way to make continuous simulations that include freezing conditions that would normally lead to the formation of ice. However, the temperature profiles during winter were not realistic, and could not be used for model calibration. This can be seen in figures 1-2 (below) where a comparison between simulated and observed water temperature at 1 m and 15 m depth is shown for year 2009. At 1 m depth, simulated and observed temperature are rather similar throughout the entire year. However, at 15 m depth, the model does not take into account the heat loss from sediment during ice-cover, which cause an increase in bottom water temperature. During winter, there is a clear mismatch between simulated and observed water temperature. For this reason, all data collected between 1 December - 31 March are excluded from the temperature data used for model calibration and only data between 1 April - 30 November are used for calibrating the model. From the example year shown below (and all other years not shown) it is evident that the measured water temperature quite closely matches the simulated temperature during the period used for calibration. Furthermore, the onset and loss of stratification always falls within this period (1 Apr – 30 Nov), showing that the lack of a fully functioning ice model will not influence simulated estimates of the timing and duration of thermal stratification. Figures 1-2 described here have now been added to the supplementary material. We expanded the description of the limitation of GOTM to simulate ice-cover in section 2.2. Besides that, we analyzed observed Lake Erken ice-cover data between 1941-2017. The results have now been added in the manuscript. The discussion have been now expanded with a paragraph describing ice-cover dynamics at Lake Erken.

Moreover, the reason we used the GOTM model is that this model was also used within the PROGNOS project (http://prognoswater.org/) to provide real-time predictions of water quality using short-term weather forecast data. In this study, the GOTM model which was already set up and tested for Lake Erken as part of PROGNOS, was used here for a different application, namely simulating long-term changes in the lake thermal structure. This model has the advantage that it can be coupled to biogeochemical models, which is crucial for the aims of PROGNOS.

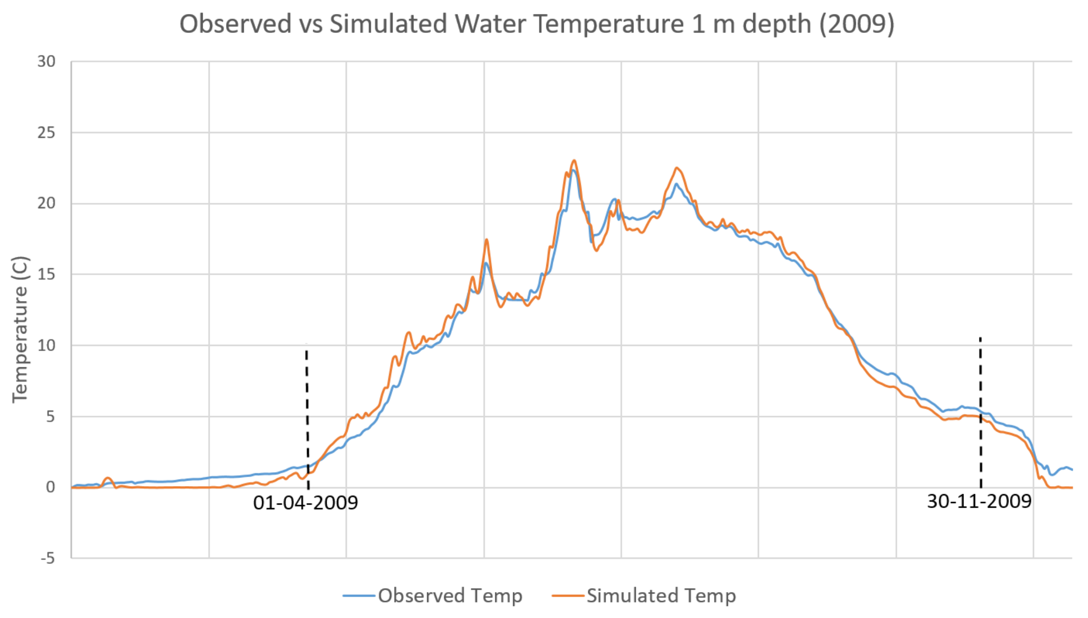


Figure 1

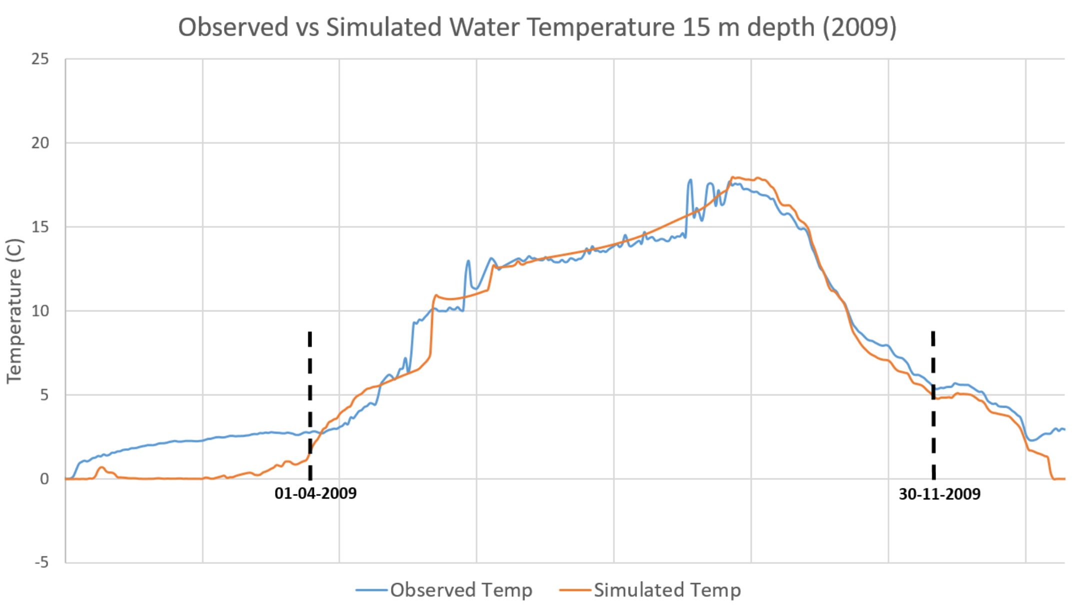


Figure 2

Referee (5) – Page 9, figure 5. Again, please use consistent y-axes and begin at zero. This is bad practice and misleading. Limit the x-axis to the start and end-year.

A continuous line should not be used to represent point data (single seasons per year). This data should be represented as points, or as a single continuous line containing all months.

Finally, please add space between the figures and their titles - at present it looks like the plot titles are related to the dashed line. Including the dashed and solid line in the legend would help. Three duplicated legends are not necessary, replace with a single legend.

Authors' response (5) – Page 9, figure 5. Figure 5 ha been improved following your suggestions

Referee (6) – Page 2, line 9. Define what is meant by a long record, for hydrological modelling of rivers this would mean > 30 years, for hydroecology > 15 years is considered long.

Authors' response (6) – Page 2, line 9. From our point of view the definition of long record is somewhat arbitrary. We think that a long record should encompass the historical changes in climate that have already occurred. For example, a record of 50 years of data can be considered long.

Referee (7) – Page 1, line 6. Need to explain what the abbreviation is - for example, consider: "General Ocean Turbulence model (GOTM), a hydrodynamic model configured in Lake Mode".

Authors' response (7) – Page 1, line 6. Thank you for the suggestion. This has been added in the revised manuscript

Referee (8) – Page 2, lines 10-11. Please provide a citation if making a claim such as this.

Authors' response (8) – Page 2, lines 10-11. This is based on the experience of one author (Don Pierson) who has worked for public water utilities and found that documenting the effects of climate change that have already occurred adds support for policies that mitigate future expected changes. To our knowledge, no reference is available.

Referee (9) – Page 2, line 17. What does significant mean? How much? Can you give a percentage or some other kind of numerical indication?

Authors' response (9) – Page 2, line 17. The word "significant" has been changed to "large". In the revised manuscript, we added the following sentence in section 2.5 (lines 12-13): "*The total number of observed water temperature data in Apr.-Nov. between 1961-2017 was 103454. The number of days with at least one single observed measurements was 6674 days between 1961-2017*."

Referee (10) – Pages 2-3, lines 29-30; 1-2. Extremely limited detail on why the lake was considered. Why should the reader care about the results from this particular work? What is interesting about it?
More information on the case study would also be useful. For example, an overview of the average climate, seasonality, the ecology of the area and anthropogenic influences.

Authors' response (10) – Pages 2-3, lines 29-30; 1-2. Lake Erken has been extensively studied in the last 70 years and it has a considerable amount of water temperature data available, which made it a good study case for testing the methodological approach of this paper. Moreover, we think that this paper describes an important methodology to reconstruct complete records of past water temperature of lakes using readily available meteorological data. Thus, the relevance of our work is not only related to the ecological importance of Lake Erken in itself.
We clarified this point in the introduction of the revised paper. Moreover, we expanded the lake description following your suggestion.

Referee (11) – Page 3, line 2. What months represent winter? The reader cannot tell how many months the lake is actually ice-covered. Also, please clarify if it is the entirety of the lake which is ice-covered.

Authors' response (11) – Page 3, line 2. In this paper, we considered the period of Dec-Mar as winter, and it is during this period when the lake is ice-covered. However, in some years, the onset of ice-cover starts in January or later and it also occasionally ends in April. Yes, the lake is ice-covered in its entirety. This information is added in the revised manuscript

Referee (12) – Page 3, lines 9-12. Repetitive - could simplify to say: "The model utilises six of these climatic parameters (excluding DP) at an hourly timestep; DP is input on a daily timestep."

Authors' response (12) – Page 3, lines 9-12. Thank you for the suggestion. This has been modified following your suggestion.

Referee (13) – Page 3, lines 4-13. More information is required for GOTM. Why choose this model specifically? Why is it well-suited for this application? Please also describe the structure of the model, what key processes does it capture? Define and describe the parameters of the model (Table 1). What are the limitations of the model?
It is also worth stating that GOTM, and all the other software/codes used, are Open Source.

Authors' response (13) – Page 3, lines 4-13. GOTM is mainly used as a stand-alone model for hydrodynamic applications in natural water, and simulates processes such as surface heat fluxes, surface mixed-layer dynamics and stratification processes. The adjusted model parameters in this study are non-dimensional scaling factors affecting the heat-flux, shortwave radiation and wind which are adjusted to minimize the difference between observed and modelled temperature. The minimum turbulent kinetic energy (k_min) and the e-folding depth for visible fraction (g2) are parameters that strongly influence the vertical distribution of light and temperature in the water column. Low values of g2 represent a higher extinction coefficient promoting higher surface temperature.
A known limitation of the model is the lack of an ice-module and a complete energy balance of the ice including ice growth and ice decay is not calculated by GOTM at this time.
We expanded the description of GOTM and its model parameters. We added that GOTM is open source.

Referee (14) – Page 3, figure 1. I am aware that the images used for review are not the final high-resolution images. However, this map looks equivalent to a screenshot. A north arrow and, critically, a scale bar, are missing. Additionally, labelling of features such as the roads and the island are unnecessary. Please consider producing a map using GIS Or similar software (mapping options are available in R). A map of Sweden indicating the location of the lake, which would highlight the relative scale, are also necessary.

Authors' response (14) – Page 3, figure 1. The figure has been substituted with a better quality map following your suggestion.

Referee (15) – Pages 3-4, lines 15-17; 1-3. Inconsistent use of meteorological station and weather station - please be consistent. For conciseness, the authors could simply state: "Driving climatic parameters were retrieved from meteorological stations at…".

Authors' response (15) – Pages 3-4, lines 15-17; 1-3. The naming is now consistent in the revised manuscript

Referee (16) – Pages 3-4, lines 15-17; 1-3. Clearer signposting is required, please refer to the letters that each station represents in the main body text.

Authors' response (16) – Pages 3-4, lines 15-17; 1-3. A reference to the point in the map is now in the revised manuscript.

Referee (17) – Page 3, line 15. Primarily retrieved from? What does primarily mean specifically?

Authors' response (17) – Page 3, line 15. We modified the sentence as follows: "*Driving meteorological parameters were retrieved whenever possible from the Erken laboratory meteorological station…*"

Referee (18) – Page 1, line 9. Real is not very clear - consider replacing with "observed" (or similar).

Authors' response (18) – Page 1, line 9. Thanks for the suggestion. This has been corrected.

Referee (19) – Page 3, line 16. Is the Malma weather station the Erken laboratory meteorological station? This inconsistency is reflected in the caption as well.

Authors' response (19) – Page 3, line 16. Yes, it is. This is now clarified in the revised manuscript.

Referee (20) – Page 4, lines 10; 21. What is meant by best? Please clarify how this is judged.

Authors' response (20) – Page 4, lines 10; 21. We removed lines 8-14 (page 4) and rewrite them as follows: "*To make maximum use of data from surrounding stations we used Artificial Neural Network function fitting analysis (ANN nftool) to predict missing meteorological data at Erken. The analysis was carried out using MATLAB version R2017b (MathWorks Inc. Natick, Massachussets)".* We modified lines 20-21 (page 4) into: "*Offsite and local dataset overlap for at leat 8-10 years to get a reasonable number of data to perform ANN function fitting analysis that describes the input-target relationship.*"

Referee (21) – Page 4, lines 24-25. Is the lake always ice-free April-November? Additionally, please replace was with "is" - I presume that the ice-free period has not recently changed, therefore this should be in the present-tense.

Authors' response (21) – Page 4, lines 24-25. This period is usually longer than the total period of ice cover which can be variable from year to year. There are occasions when ice continues into April, but the April to November period is definitely representative of ice-free conditions and using data from this period to calibrate the model will definitely avoid errors associated with GOTM's simplistic simulation of ice cover. During the revision of the manuscript we noticed that water temperature data of April and November were not used when calibrating the model. We re-ran the model calibration including April and November measured water temperature and we updated the calibration parameters. The new calibration provided very similar results to the calibration showed in the discussion paper. These are the values of the updated calibrated parameters and model statistics

- Heat-flux factor: 0.863009
- Short- wave radiation factor: 0.970753
- Wind factor: 1.28701
- Minimum turbulent kinetic energy: 1.64873e-06
- e-folding depth for visible fraction: 2.63732

- ln Likelihood: -60469.715
- Bias (°C): -0.04707
- MAE (°C): 0.7529
- RMSE (°C): 1.089
- Correlation: 0.9717

The tables are now updated with these new values in the revised paper.

The sentence is now corrected with the present- tense.

Referee (22) – Page 4, lines 24-32. Why is this text part of model calibration? This is still text relating to the input data. Perhaps consider combining 2.3 Data sources of driving parameters-2.3 Model calibration (paragraph 1) into a single Data section.

Additionally, there is an issue with the section numbering: 2.3-2.4-2.3

Authors' response (22) – Page 4, lines 24-32. We renamed the section "2.3 Data sources of driving parameters"(page 3, line 14) into "2.3 Data sources of meteorological parameters" and renamed the section "2.3 Model calibration" (page 4, line 23) into "2.5 Observed water temperature data and model calibration".

We corrected the section numbering in the revised manuscript.

Referee (23) – Page 5, lines 1-2. As with the hydrodynamic model, the reader needs to know why this approach is used. What is the rationale?

Authors' response (23) – Page 5, lines 1-2. We added the following sentence in the text: **"***ACPy is a utility that eliminates the need for time consuming manual calibration of hydrodynamic and water quality models. This allows for more extensive testing and evaluation of model calibrations, ultimately providing more accurate and repeatable results".*

Referee (24) – Page 5, lines 5-7. These lines are unclear, please consider rewording.

Authors' response (24) – Page 5, lines 5-7. We modified lines 5-7 as follows: "*Simulations were run between 1961 and 2017 but in order to obtain stable initial conditions the model was run over an additional one year spin up using a copy of the 1961 data. In this way, 1961 data were both used as a spin-up year and then reused in the proceeding calibration*."

Referee (25) – Page 5, lines 8-13. The authors state that an algorithm is used in the parameterization. What is the stopping criteria? How does the algorithm select a parameter set?

Authors' response (25) – Page 5, lines 8-13. ACPy was set to run 10000 simulation during calibration to get a stable solution to obtain the optimal parameter set that minimizes the log likelihood function. Manual testing found that additional simulations added very little if any improvement to the model calibration.

Referee (26) – Pages 5-6, lines 28; 1-3. Please explain to the reader why they should care about these metrics - why are they important? What do they indicate?

Authors' response (26) – Pages 5-6, lines 28; 1-3. We rephrased lines 26-27 (page5) as follows: "*We summarized the model temperature output by calculating a number of statistics that can qualify the ecological consequence of changes in thermal stratification using the Lake Analyzer R Package. The ecological implications of the changes of these metrics due to climate change are discussed in detail in the Discussion section".*

Referee (27) – Page 6, line 12. Did you test for autocorrelation? Was it all autocorrelated? Please be clearer.

Authors' response (27) – Page 6, line 12. Yes, we tested for autocorrelation using acf and pacf function in R. This is now mentioned in the methods (section 2.6). In table 4 are now mentioned which datasets were autocorrelated.

Referee (28) – Page 6, lines 13-14. Please correct Figure 3 accordingly - the time-series should not extend beyond the point for which it is useful!

Authors' response (28) – Page 6, lines 13-14. This has been modified in the revised manuscript following your suggestion.

Referee (29) – Page 1, lines 10-12. Suggest the author's state why the results are split into these sub-intervals; until very late on the paper I presume the split was because pre-1988 records were patchy.

Authors' response (29) – Page 1, lines 10-12. Thanks for the suggestion. We added a sentence in the abstract that specifieed that the splitting was selected because of an abrupt change in air temperature.

Referee (30) – Page 6, line 21. Please consider moving this line to the start of the section. Please also include the package version for Lake Analyzer (the citation seems relatively old).

Authors' response (30) – Page 6, line 21. The line is now at the beginning of the section. The package version used here is 1.11.4 and the citation has changed with the following: "Winslow, L., Read, J., Woolway, R., Brentrup, J., Leach, T.,Zwart, J., Albers, S., and Collinge, D: rLakeAnalyzer: Lake Physics Tools. R package version 1.11.4. https://CRAN.R-project.org/package=rLakeAnalyzer, 2018."

Referee (31) – Page 9, line 2. What data did this use? The pre-1988 data which included data from mixed stations and the post-1988 data which was much more consistent? Can this finding be trusted?

Authors' response (31) – Page 9, line 2. Yes, pre-1988 data mostly included data from mixed stations. However, these data have been adjusted to better represent Lake Erken local conditions (at Malma met station) using Neural Network function fitting analysis. This analysis report a very good agreement between Erken air temperature (target data) and output data (see suplementary material). Given that, we think that we think that air temperature dataset used in this study is reliable.

Referee (32) – Page 9, line 2. In the discussion, please explain to the reader why this matters, what it indicates etc. - It is not made clear.

Authors' response (32) – Page 9, line 2. From our point of view, an abrupt change in air temperature support the fact that a more rapid change in water temperature is occurring in the last decades and that the effect of climate change on thermal conditions is accelerating. We added the following sentence in the revised paper: *"The Pettitt test showed that a significant abrupt change in annual mean air temperature occurred in 1988 (p < 0.001). Therefore, in addition to checking for trends in lake thermal structure over the entire simulation period we also evaluated the possibility of trends occurring over the subinterval 1961-1988 and 1989-2017, and we tested whether a more rapid change in water temperature is occurring after 1988, following a step-change in annual mean temperature."*

Referee (33) – Page 9, line 5. Please define your terms, e.g. epilmnetic.

Authors' response (33) – Page 9, line 5. The terms epilimnetic and hypolimnetic are now defined in the revised paper.

Referee (34) – Page 9, lines 9-14. Please be consistent in the number of significant figures for temperature.

Authors' response (34) – Page 9, lines 9-14. Thanks for the suggestion. This will be corrected in the revised paper.

Referee (35) – Page 9, line 14. Please start a new paragraph before discussing thermal stratification.

Authors' response (35) – Page 9, line 14. Thanks for the suggestion. This is now corrected in the revised paper.

Referee (36) – Pages 9-10. As a decadal mean, it would be useful to see the reporting of confidence intervals for these values. Perhaps consider a table of results.

Authors' response (36) – Pages 9-10. Thanks for your suggestion. A table of results is now added to the revised manuscript.

Referee (37) – Page 11, lines 1-8. You cannot claim that there was a good match. No valid assessment of model performance was provided. This needs to be significantly addressed before such a claim can be asserted.

Authors' response (37) – Page 11, lines 1-8. We now referred to the model performance in the revised paper.

Referee (38) – Page 11, line 12. I do not agree that it indicates the reliabilty, the wording is too strong. It could be described as a positive indication.

Authors' response (38) – Page 11, line 12. We changed the word "reliability" into "consistency".

Referee (39) – Page 12, lines 1-11. Much of this text appears to be results.

Authors' response (39) – Page 12, lines 1-11. From our point of view, summarizing the major results in the discussion section while at the same time providing a possible explanation for them improves understanding for the reader and also improves the readability of the manuscript.

Referee (40) – Page 1, lines 10-11; 15-16. State the months associated with your seasons.

Authors' response (40) – Page 1, lines 10-11; 15-16. The months are now added.

Referee (41) – Page 12, lines 14-17. The provision of a confidence interval would help to expand upon this further (it could also improve or worsen the difference in results).

Authors' response (41) – Page 12, lines 14-17. A confidence interval is not reported by O'Reilly et al. (2015), but only the Sen's slope of lake summer water temperature trends.

Referee (42) – Page 12, lines 18-22. Does O'Reilly account for the influence of ice-cover? If yes, could this not also account for some of the discrepancy? Please weight the pros and cons of this study versus theirs accordingly.

Authors' response (42) – Page 12, lines 18-22. Trends reported by O'Reilly et al. do not account for ice-cover since the work reports on summer lake surface temperature trends. However, the paper does state that lakes that are always completely ice-covered during winter (this is the case of Lake Erken) are experiencing a faster warming trend compared to lakes that do not freeze during winter.

Referee (43) – Page 13, lines 25-34. Suggest that the authors consider leading the discussion with this text. At present, it is not clear to the reader why this work or the results are relevant - the implications are not made clear.

Authors' response (43) – Page 13, lines 25-34. We consider that the first part of the discussion is useful to demonstrate the validity of the first aim of this work, which is to provide a valuable and reliable method to extend historical water temperature records back in time. In the second part (lines 25-34 page 13 and lines 1-7 page 14) we provided the most important ecological implications that a warmer climate might have on Lake Erken specifically. This order follows the same order of how the aims are described in the introduction.

Referee (44) – Page 14, lines 8-12. The assertion of "accurately" cannot be made whilst there is no robust consideration of model performance.

Authors' response (44) – Page 14, lines 8-12. We rephrased lines 8-9 (page 14) in the following way: "*The present study has shown that the use of the GOTM model to reconstruct the past 57-years of thermal condition of Lake Erken provided a valuable source of information that could be used to detect changes in its thermal structure. This methodology can be extended to other lakes that have incomplete records of water temperature data. The use of local meteorological data to drive model simulations such as those demonstrated here can be used to extend water temperature records further back in time or fill data gaps where they exist*."

Referee (45) – Page 14, lines 13-19. Suggest that a dedicated conclusion would help to wrap up the paper and reassert the aims/objectives and relevance of the work.

Authors' response (45) – Page 14, lines 13-19. A dedicated conclusion is now present in the revised paper.

Referee (46) – Page 1, lines 23-25. Abstract does not necessarily make clear why this matters - what is the need for the work?

Authors' response (46) – Page 1, lines 23-25. We removed line 5 (page 1) and write as follows: "*Historical lake water temperature records are a valuable source of information to assess the influence of climate change on lake thermal structure. However, in most cases such records span a short period of time and/or are incomplete, providing a less credible assessment of change.*"

Referee (47) – Page 1, lines 27-29. This first sentence is repetitive; also not convinced that Samal et al., 2012 is the best citation for this critical statement. There are other more relevant seminal works that the authors may cite.

Authors' response (47) – Page 1, lines 27-29. We rephrased lines 27-29 as follows: "*Changes in the thermal structure and mixing regimes of lakes are a consequence of changes in several climatic factors such as air temperature, solar radiation, cloud cover, wind speed and humidity (Woolway and Merchant, 2019).*"

Referee (48) – Page 2, lines 1-2. Again, repetition - it is self-evident that a rise in lake water temperature increases water temperature - please be more concise.

Authors' response (48) – Page 2, lines 1-2. Thanks for your suggestion. The citation of Arhonditsis et al. has been removed

Referee (49) – Page 2, lines 5-7. It would be helpful to explain what some of the conclusions of these studies are/were - it makes it clearer to the reader why there is a need for this.

Authors' response (49) – Page 2, lines 5-7. We now removed the reference to this paper and change the first paragraph of the introduction (see Author,s response 1).

Referee (50) – Page 6, lines 4-7. I would like to highlight that the level of description here is excellent and represents the level that should be achieved throughout the manuscript.

Authors' response (50) – Page 6, lines 4-7. Thanks for your comment!

Historical modelling of changes in Lake Erken thermal conditions

Simone Moras, Ana I. Ayala, Don C. Pierson

Department of Ecology and Genetics, Uppsala U~~u~~niversity, Uppsala, 75236, Sweden

Correspondence to: Simone Moras (simone.moras@ebc.uu.se)

Abstract. ~~The thermal structure of lakes is strictly related to climate and to the variability of thermal and mixing dynamics. In this study, a physical hydrodynamic model (GOTM)~~ Historical lake water temperature records are a valuable source of information to assess the influence of climate change on lake thermal structure. However, in most cases such records span a short period of time and/or are incomplete, providing a less credible assessment of change. In this study, the hydrodynamic model GOTM (General Ocean Turbulence Model, a hydrodynamic model configured in Lake Mode) was used to reconstruct daily time-step water temperature of Lake Erken (Sweden) over the period 1961-2017, using seven climatic parameters as forcing data: wind speed (WS), air temperature (Air T), atmospheric pressure (Air P), relative humidity (RH), cloud cover (CC), precipitation (DP) and shortwave radiation (SWR). The model was calibrated against observed~~real~~ water temperature data collected during the study interval, and the calibrated model revealed a good match between modelled and observed temperature (RMSE=1.~~112~~ 089 °C). From the long-term simulations of water temperature, this study focused on detecting possible trends in water temperature over the entire study interval 1961-2017 and in the sub-intervals 1961-~~1987~~ 1988 and ~~1988~~1989-2017, since an abrupt change in air temperature was detected in 1988. The analysis of the simulated temperature showed that epilimnetic temperature has increased on average by 0.444°C/decade and 0.792 °C/decade in spring and autumn in the sub-interval 1989-2017. Summer epilimnetic temperature has increased by 0.351 °C/decade over the entire interval 1961-2017. Hypolimnetic temperature has increased significantly in spring over the entire interval 1961-2017 by 0.148 °C/decade and by 0.816 °C/decade in autumn in the sub-interval 1989-2016. Whole-lake temperature showed a significant increasing trend in the sub-interval 1989-2017 during spring (0.404 °C/decade), and autumn (0.789 °C/decade, interval 1989-2016), while a significant trend was detected in summer over the entire study interval 1961-2017 (0.239 °C/decade). Moreover, this study showed that changes in the phenology of thermal stratification have occurred over the 57-years period of study. Since 1961, the stability of stratification (Schmidt Stability) has increased by 5.365 Jm^{-2}/decade. The duration of thermal stratification has increased by 7.297 days/decade, correspondent with an earlier onset of stratification of ~ 16 days and to a delay of stratification termination of ~ 26 days. The average thermocline depth during stratification became shallower by ~1.345 m, and surface-bottom temperature difference increased over time by 0.249 °C/decade. ~~The analysis of the simulated temperature showed that epilimnetic temperature has increased on average by +0.43°C/decade and +0.809 °C/decade in spring and autumn in the sub-interval 1988-2017. Summer epilimnetic temperature has increased by +0.348 °C/decade over the entire interval 1961-2017. Hypolimnetic temperature has increased significantly in the sub-interval 1988-2016 by +0.827 °C/decade in autumn. Whole-lake temperature showed a significant increasing trend in the sub-interval 1988-2017 during spring (+0.378 °C/decade) and in autumn (+0.809 °C/decade). Moreover, this study showed that changes in the phenology of thermal~~

stratification, have occurred over the 57-years period of study. Since 1961 the stability of stratification (Schmidt Stability) has increased by 5.535 Jm^{-2}/decade. The duration of thermal stratification has increased by 7.083 days/decade, correspondent with an earlier onset of stratification of ~ 16 days and to a delay of stratification termination of ~ 26 days. The average thermocline depth during stratification became shallower by ~1.242 m, and surface-bottom temperature difference increased over time by +0.249 °C/decade. The creation of daily-time step water temperature dataset not only provided evidence of changes in Erken thermal structure over the last decades, but it is also a valuable resource of information that can help in future research on the ecology of Lake Erken. The use of readily available meteorological data to reconstruct Lake Erken's past water temperature is shown to be a useful method to evaluate long-term changes in lake thermal structure, and it is a method that can be extended to other lakes.

1. Introduction

The variability of thermal stratification and the dynamics of mixing in lakes and reservoirs is closely coupled to several atmospheric factors such as solar radiation, air temperature and wind speed, highlighting the fact that changes in water thermal conditions are strongly related to changes in climatic conditions (Samal et al. 2012). As a consequence of a warmer climate on a global scale, several studies have shown that lake water temperature is strongly affected (Adrian et al. 2009), causing shorter ice cover periods (Blenckner et al. 2002, Butcher et al. 2015, Kainz et al. 2017), increased water temperature (Arhonditsis et al. 2004) and stronger summer stratification (Jankowski et al. 2006). Since climatic conditions have changed markedly in the last century and they are expected to change considerably in the next decades (IPCC, 2013), the importance of predicting how freshwater bodies will be affected by such changes becomes evident. Several studies have used modelling tools to assess how lake water temperature and more generally the entire lake ecosystem will respond under different future scenarios (Stefan et al. 1998, Taner et al. 2011, Winslow et al. 2017). However, future predictions on how lakes will respond to climate change are less credible without model validation based on comparison with historical water temperature data. A long record of historical data provides more background information and allows better documentation of the changes that have already taken place and leads to more accurate predictions of lake thermal condition in future decades. Changes in the thermal structure and mixing regimes of lakes are connected to changes in several climatic factors such as air temperature, solar radiation, cloud cover, wind speed and humidity (Woolway and Merchant, 2019). The alteration of lake hydrodynamic properties has consequences on lake chemistry, biology and ultimately on the ecosystem services that lakes provide (Adrian et al., 2009; Vincent, 2009). Since climatic conditions have changed markedly in the last century and they are expected to change considerably in the next decades (IPCC, 2013), the importance of evaluating how freshwater bodies are affected by climate change becomes evident. A direct assessment of how lakes have already been affected by climate change is to analyse historical trends in lake water temperature data. However, the availability of long-term data of lake water temperature is still scarce. For example, there are very few lakes around the world with a long-term record (defined here as >50 years) of water temperature profiles (e.g. Jankowski et al., 2006; Skowron, 2017). Instead, the availability of long-term historical data is often limited to surface water temperature of one or few lakes (e.g. Livingstone and Dokulil, 2001; Kainz et al., 2017) and the time frame of surface temperature data available for the majority of lakes is limited to 2-3 decades. For example, Sharma et al. (2015) compiled a worldwide database with lake surface water temperature between 1985-2009. The same time frame was used by Schneider and Hook (2010) that reported an average warming trend of 0.045 ± 0.011°C/year of lake surface water temperature in 167 large lakes (>500 km2) using satellite-derived measurements; similarly O'Reilly et al. (2015) reported an average warming trend of 0.34 °C/decade for lake summer surface water temperature in 235 lake worldwide retrieved from both in-situ and satellite data. Even though these studies have demonstrated a rapid warming trend among lakes, the analysis of only surface water temperature is not sufficient to obtain a complete evaluation of the changes in the thermal structure that encompass, for example, temperature trends in the water column and phenology of thermal stratification. Moreover, the scarcity of water temperature data before 1980s makes difficult to assess earlier thermal conditions for the majority of lakes. A longer record of historical data provides more background information, allows better documentation of the changes that have already taken place, and leads to more accurate predictions of lake thermal conditions in future decades. One of the best arguments to counter climate change sceptics is well documented long-term records of the ongoing effects of climate change.

For these reasons, we aimed to create a daily time-step dataset of historical water temperature data for Lake Erken (Sweden) using a hydrodynamic model. Lake Erken has been studied extensively over the past 70 years (Pettersson 2012) and automated hourly measurements of water temperature and meteorological data have been collected from the lake since October 1988. For this reason, the aim of this study was to use a hydrodynamic model to extend records of lake water temperature back in time, in order to provide a longer and more consistent picture of the changes in thermal structure of lakes. We tested this approach on Lake Erken, which has been studied extensively over the past 70 years (Pettersson, 2012). Automated hourly measurements of water temperature and meteorological data have been collected from the lake since October 1988. Before 1988, water temperature measurements were taken manually only during periodic sampling campaigns during periodic sampling campaigns, or recorded using strip chart recorders from a limited number of depths during several years (1961-1963). As a consequence, information about the thermal state of the lake before 1988 was missing or patchy infrequent for most of the time, and even after 1988 there are significant gaps in the measured temperature data. Thus, the information available for Lake Erken made it a good study case for testing the methodological approach of this study. The aim of this study is to use a hydrodynamic model to extend records of lake water temperature further back in time until 1961, in order to provide a longer and more consistent picture of the changes that have occurred in Lake Erken over the last five decades. The model is driven by meteorological data collected from Uppsala University's field station at Lake Erken (http://www.ieg.uu.se/erken-laboratory/) and from nearby stations to create daily time-step water temperature profiles for the entire period 1961-2017. This work aims to (1) demonstrate the validity of using modelled temperature to reconstruct past water temperature of Lake Erken for the period 1961-2017, providing a valuable method that can be extended to other lakes and (2), to evaluate how water temperature and other metrics of lake stratification have changed over the study period. Here, we created a complete daily time step water temperature record for Lake Erken, extending the information on its thermal structure further back in time until 1961, in order to provide a longer and more consistent picture of the changes that have occurred in the lake over the last five decades. The GOTM hydrodynamic model used here is driven by meteorological data collected from Uppsala University's field station at Lake Erken (http://www.ieg.uu.se/erken-laboratory/) and from nearby stations to create daily time-step water temperature profiles for the entire period 1961-2017. In this work we evaluated (1) the validity of using modelled temperature to reconstruct past water temperature of Lake Erken for the period 1961-2017, providing a valuable method that can be extended to other lakes, and (2) how water temperature and other metrics of lake stratification have changed over the study period. -Finally, the creation of a reliable consistent and complete 57-year dataset of daily water temperature profiles will be a valuable source of information for future research on Lake Erken that will help to better our understanding of many ecological processes that can be affected by changes in thermal conditions.

2. Methods

2.1 The lake

The lake investigated in this study was Lake Erken (59.4166 N, 18.2500 E) a mesotrophic lake located ~60 km North-East from Stockholm (Sweden) at an altitude of 10 m above the sea level. The lake covers an area of about 24 km2 and its catchment area is relatively small (141 km2), mainly covered by forest and with no major anthropic activities (Malmaeus et al., 2005). Lake Erken has a mean depth of 9 m and a maximum depth of 21 m while its water retention time is around 7 years (Blenckner et al., 2002). Little is the contribution of inflows on lake hydrodynamics (Pierson et al., 1992). The lake is always ice-covered in its entirety during winter between December-February to March-May (Blenckner et al., 2002; Persson and Jones, 2008) and is always stratified during summer months between May-June to August-September (Persson and Jones, 2008).~~and it covers an area of about 24 km^2. Lake Erken's mean depth is 9 m and its maximum depth is 21 m. The retention time of the lake is around 7 years. The lake is ice-covered during winter until March-early May (Blenckner et al. 2002).~~

2.2 The model

The model used in this work is GOTM (General Ocean Turbulence Model). GOTM is an open source 1-dimensional physical model for hydrodynamic applications in natural waters, it simulates processes such as surface heat fluxes, surface mixed-layer dynamics and stratification processes.~~The model used in this work is GOTM (General Ocean Turbulence Model). It is a 1-dimensional physical model built to simulate water temperature and other physical parameters using meteorological parameters as driving factors.~~ Detailed information about GOTM can be found in Burchard (2002) and on the website www.gotm.net. In this study, seven climatic parameters were used as input to the GOTM model over the study period (January 1st, 1961 - October 31st, 2017): wind speed (WS; m/s), air temperature (Air T: °C), relative humidity (RH; %), atmospheric pressure (Air P; hPa), cloud cover (CC; dimensionless value between 0-1), shortwave radiation (SWR; W/m^2) and precipitation (DP; mm/day). The model utilises six of these climatic parameters (excluding DP) at an hourly time step; DP is input on a daily time step. ~~To run the model, hourly datasets were created for six climatic parameters over the considered study period (WS, Air T, Air P, RH,~~

CC, SWR), while DP dataset was based on daily values. For the purpose of this studystudy, the lake was considered to have a fixed water level equal to the long-term mean. This assumption was justified given the lakes long retention time and that the mean annual variation in lake level is only 48 cm. The GOTM model used for the simulations documented here did not have a fully functioning ice model, but instead cut off surface heat exchange when the simulated surface water temperature become negative. This provides a very simple way to make continuous simulations that include freezing conditions that would normally lead to the formation of ice. Also, the model does not take into account the heat loss from lake sediments during ice-cover, which causes an increase in bottom water temperature. For this reason, the temperature profiles during winter (Dec.-Mar.) were not realistic and were excluded from model calibration. However, the onset and loss of stratification always falls between 1 Apr – 30 Nov (the period used for calibration), showing that the lack of a fully functioning ice model did not influence simulated estimates of the timing and duration of thermal stratification. A visual example that describe the mismatch between modelled and observed temperature in winter is available in the supplementary material (fig. S1-S2).

2.3 Data sources of meteorological parametersData sources of driving parameters

Driving meteorological parameter were retrieved whenever possible from the Erken laboratory meteorological station Driving climatic parameters were primarily retrieved from the Erken laboratory meteorological station (Malma islet; 59.8391 N, 18.6296 E, fig. 1a letter A and Svanberga meteorological station (59.8321 N, 18.6348 E, fig. 1a, letter B), about 800 m from the Erken laboratory meteorological station. 1) and Svanberga station (59.8321 N, 18.6348 E, fig. 1), about 800 m from the Malma weather station. The Svanberga weather station is managed by SMHI (Swedish Meteorological and Hydrological Institute) and climatic data were downloaded from SMHI website (http://opendata-download-metobs.smhi.se/explore/). /). Meteorological data from neighboring meteorological stations were used when data from Erken or Svanberga were not available (fig. 1b). When data could not be retrieved from neighboring stations, missing data were replaced by linear interpolation. An overview of the number of meteorological data retrieved from different sources is given in table 1. A detailed description on the methodology used to retrieve these data is available in supplementary material.

Climatic data from neighboring stations were used when data from Erken or Svanberga were not available. A detailed description on the methodology used to retrieve these climatic data is available in supplementary material.

2.4 Missing data replacement and missing data estimation with Artificial Neural Network analysis

To simulate Lake Erken water temperature at daily time step using GOTM, a continuous hourly record of meteorological forcing data was created by merging the data sources described above. In the case of DP missing data were replaced by taking data from the closest stations to Lake Erken (see supplementary material). For cloud cover that was only available from one station (Svenska Hogarna: 59.4445 N, 19.5059 E), missing data were replaced by linear interpolation. Similar methods were first used to estimate WS, Air T, Air P, RH and SWR. To make maximum use of data from surrounding stations we used an Artificial Neural Network fitting analysis (ANN nftool) to predict missing meteorological data at Erken (DP and CC excluded).

The analysis was carried out using MATLAB version R2017b (MathWorks Inc. Natick, Massachusetts).However, for these remaining meteorological variables that showed significant inter-station variability, we found that Artificial Neural Network (ANN) algorithms provided the best estimate of local Erken meteorological data as judged by comparison of modeled and observed water temperature. ANN function fitting analysis was used to predict missing data using the ANN fitting tool (nftool) in MATLAB version R2017b (MathWorks Inc. Natick, Massachusetts). Compared to other methods, ANN nftool also made maximum use of data from surrounding stations to predict missing meteorological data at Erken.

ANN algorithms were used to estimate each the driving parameters during occasions when no local measurements were recorded at Malma Erken meteorological station (WS, Air T and SWR) and Svanberga (Air P and RH). Input data were those collected from the nearest (less than 60 km away) meteorological stations to Lake Erken and Svanberga. Input data were retrieved from SMHI database, except for SWR data that were retrieved from measurements made at the Swedish Agricultural University (SLU) near Uppsala. The choice of the climatic datasets to use as input data was based on two characteristics:

1. Offsite datasets that have recorded data when data from Lake Erken or Svanberga were not available.
2. Offsite and local dataset overlap for at least 8 years to get a reasonable number of data to perform ANN function fitting analysis that describes input-target relationship.Offsite and local datasets overlap for at least 8-10 years. We found that the ANN fitting tool requires this length of data overlap in order to obtain the best fit function that describes input-target relationship.

A detailed description of the Neural Network analysis is available in supplementary material.

2.5 Observed water temperature data and model calibration2.3 Model Calibration

The model was calibrated using measured profiles of averaged daily water temperature collected between April 1st - November 30th when the lake is usually ice-free. The model was calibrated using measured profiles of averaged daily water temperature collected between April and November when the lake was ice-free. Observed data in the period 1961-1988 were collected manually during occasional sampling campaigns, and from strip chart data recordings made at the island Erken meteorological station. Most of the observed temperature data were measured at 0.5, 5, 10 and 20 m depth. A much greater number of observed data were available for the period 1989-2017 when an automated floating station (59.84297 N, 18.635433 E, fig. 1, letter C) was deployed to collect water temperature data during ice-free period. The floating station measured water temperature data every 0.5 meters, from 0.5 m to 15 m depth. Profiles were stored every 30 minutes and these were averaged to provide a daily mean profile. Also between 1989-2017 water temperature was digitally recorded in a manner similar to the old strip chart recordings. These measurements were made year-round from the 1, 3 and 15 m depths at the Erken meteorological stationMalma island station, and used for calibration at times the floating system was not deployed. The total number of observed water temperature data in Apr.-Nov. between 1961-2017 was 103454. The number of days with at least one single observed measurements was 6674 days between 1961-2017.

The program ACPy (Auto Calibration Python) was used to calibrate the model (webpage: www.bolding-bruggeman.com/portfolio/acpy/). ACPy is a utility that eliminates the need for time-consuming manual calibration of hydrodynamic and water quality models. This allows for more extensive testing and evaluation of model calibrations, ultimately providing more accurate and repeatable results. A set of model parameters was calibrated and adjusted within their feasible range (see table 2) in order to minimize the difference between the simulated and measured water temperature. Simulations were run between 1961 and 2017 but in order to obtain stable initial conditions the model was run for an additional one year spin up using a copy of the 1961 data. In this way, 1961 data were both used as spin-up year prior to the calibration and then reused in the proceeding calibration.A set of model parameters was calibrated and adjusted within their feasible range (see table 1) in order to minimize the difference between the simulated and measured water temperature. A 1-year simulation spin-up was used to minimize errors in the calibration by allowing the initial state of the model to better reflect typical lake conditions. To make full use of the input data when spinning up the model a copy of the 1961 data was appended to the beginning of the input data and this year was used as the spin up. In this way, 1961 data were both used as spin-up year and discarded from calibration, and then reused in the proceeding calibration.

In ACPy, a Differential Evolution algorithm (Storn and Price, 1997) is used to calculate a log likelihood function which compares the modelled water temperature to the observed temperature.

The likelihood Λ is defined in Eq. (1):

$$\Lambda = - \sum_i \frac{(x_{obs,i} - x_{mod,i})^2}{var_x} \quad (1)$$

where $x_{obs,i}$ is the observed temperature, $x_{mod,i}$ is the modeled temperature and var_x is the variance between the modelled and observed temperature. ACPy was set to run 10000 simulation during calibration in order to obtain a stable solution and get the optimal parameter set that minimizes the log likelihood function. Following the calibration, model fit was judged based on estimates of bias, mean absolute error (MAE), and the root mean squared error (RMSE).defined in Eq. (2), Eq. (3) and Eq. (4) respectively:

$$bias = \sum_{i=1}^{N} \frac{(x_{mod,i} - x_{obs,i})}{N} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^{N} |x_{obs,i} - x_{mod,i}|}{N} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x_{obs,i} - x_{mod,i})^2}{N}} \quad (4)$$

For typical applications of models where the goal is to make simulations to future or otherwise different conditions that are not covered by the record of measured calibration data, it is appropriate to employ a split calibration and validation strategy. However, in our case the goal was not to simulate outside of the period of available calibration data, but to use the model to provide a complete and consistent record over a period in which calibration data were available but incomplete. Therefore, we used here the entire record of measured temperature over the study period (1961-2017) to judge the validity of the calibration. This ensured that the calibration encompasses the widest possible range of variability and that our simulations within the calibration period would have the greatest degree of accuracy. The adjusted model parameters in this study are non-dimensional scaling factors that adjust the heat-flux, shortwave radiation and wind as well as the minimum turbulent kinetic energy and the e-folding depth for visible fraction of incoming radiation which are parameters that strongly influence the vertical distribution of light and temperature in the water column. Using ACPy all of these parameters were adjusted to minimize the difference between observed and modelled temperature. For the purpose of generating the long-term time series of temperature data we used the full calibration period (all seasons but winter) because it encompassed the full variation of conditions that occur over the entire simulation period. However, in addition to that, we also performed seasonal calibrations in order to validate that the model performed well in all seasons (winter excluded), especially when simulating the onset of stratification after the ice break-up.

The best set of parameters calculated from the ACPy full calibration period (table 2) were then used for the final simulation which produced the data analyzed in the remainder of this paper. The model fit results for the full period and seasonal calibration are shown in table 3. The comparison between observed and simulated water temperature and the error distribution after full calibration period are shown in figure 2. We also calculated the model performance at different depths (0.5, 5, 10 and 15 meters) after full calibration period. The results are shown in figure 3-4.Since the purpose of this study was to reconstruct the lake water temperature over a pre-defined period (1961-2017), we used the entire record of measured temperature over this period to judge the validity of the calibration. This ensured that the greatest range of environmental conditions would be represented, and that our simulations within the calibration period would have the greatest degree of accuracy. The best set of parameters calculated from the ACPy calibration (table 1) were then used for the final simulation which produced the data analyzed in the remainder of this paper. The error distribution between the modelled and the observed water temperature after calibration is shown in figure 2, while the comparison between modelled and observed temperature at different depths is shown in figures 3-4.

2.6 2.4 Statistical analysis

The entire statistical analysis was carried out using R Studio version 3.4.1 (R Studio Team 2016). We summarized the model temperature output by calculating a number of statistics that can qualify the ecological consequence of changes in thermal stratification using Lake Analyzer R Package (Winslow et al., 2018).The Lake Analyzer R package (Read et al. 2011) was used to retrieve several metrics describing lake thermal structure at a daily time step for the entire study period from the modelled water temperature profiles. Lake Analyzer calculates volumetrically averaged epilimnetic (upper layer during stratification) , hypolimnetic (lower layer during stratification) and whole-lake temperature, the thermocline depth (°C), Schmidt stability (J m-2, Schmidt 1928, Idso 1973), and difference between surface and bottom temperature (ΔT, °C). Lake Analyzer calculates volumetrically averaged epilimnetic, hypolimnetic and whole lake temperature (°C), Schmidt stability (J m^{-2}, Schmidt 1928, Idso 1973), thermocline depth (m) and difference between surface and bottom temperature (ΔT, °C). In addition, we calculated the length of the growing season for each year, which is defined as the number of days in which epilimnetic temperature exceeds 9 °C (Håkanson and Boulion, 2001).

In this study, the lake was considered stratified when the difference between surface temperature and bottom temperature (ΔT) was greater than 1 °C (Woolway et al. 2014). The onset of stratification was considered to be the first day of the first period of 4 or more consecutive days in which ΔT > 1 °C (Yang et al. 2016). and, in general, stratification events shorter than 4 days were also not considered when estimating duration and loss of stratification.Thus, stratification events shorter than 4 days were not considered in the analysis of stratification duration.

To assess if water temperature trends vary with seasons, the seasonal averages of the simulated water temperature were analyzed using the non-parametric Mann-Kendall test (Mann 1945, Kendall 1975). This test assesses whether a statistically significant monotonic increase or decrease over time is occurring. The values of such trends were estimated using the non-parametric Sen's slope (Sen 1968), which is the median of all pairwise slopes of the considered data. For autocorrelated data, the modified version of Mann-Kendall test proposed by Hamed and Rao (1998) was used instead of the traditional Mann-Kendall test, which does not account for autocorrelation. Since the simulation stops in October 2017, autumn water temperature of that year were not taken into consideration in the data analysis. The Mann-Kendall and Sen's slope were also used to evaluate trends in average Schmidt stability during thermal stratification, thermocline depth, stratification duration, onset and termination of stratification and growing season length. Data autocorrelation was tested using acf and pacf function in RStudio. For auto-correlated data, the modified version of Mann-Kendall test proposed by Hamed and Rao (1998) was used instead of the traditional Mann-Kendall test, which does not account for autocorrelation. We used the Pettitt test (Pettitt 1979) to assess whether an abrupt change in annual mean air temperature occurred during the study period. Since Lake Erken is always ice-covered during winter and the GOTM model does not contain an ice-cover module, the simulated winter lake temperature were underestimated, especially in the bottom layers, where GOTM did not simulate the effect of heat flux from the sediment into the water. For this reason, trends in winter lake temperature were not analyzed in this study. However, the availability of manual observations of the timing of ice cover since 1941 (for 10 out of 68 years ice cover data are not available) for Lake Erken made it possible for us to test for trends in ice cover length during the interval 1941-2017, and to make comparison with the other simulated lake metrics. Mann-Kendall test, Sen's slope and Pettitt test were therefore used to analyse the record of ice cover length obtained from observational data. A synthesis of the statistical analysis results is reported in table 4.In addition, since Lake Erken is always ice-covered during winter and the GOTM model does not contain an ice-cover module, the simulated winter lake temperature might have been underestimated, especially in the bottom layers, where GOTM might have not simulated the effect of heat flux from the sediment into the water. For this reason, trends in winter lake temperature were not analyzed in this study. The Mann-Kendall, modified Mann-Kendall test and Sen's slope were also used to evaluate trends in average Schmidt stability during thermal stratification, thermocline depth, stratification duration, onset and termination of stratification and growing season length. We used the Pettitt's test (Pettitt 1979) to assess whether an abrupt change in annual mean air temperature occurred during the study period. The entire statistical analysis was carried out using R Studio version 3.4.1 (R Studio Team 2016).

3. Results

3.1 Model performance

The GOTM model was able to accurately reconstruct water temperature of Lake Erken in the time interval 1961-2017. Overall, the calibrated model showed a RMSE of 1.089 ºC, a MAE of 0.753 ºC (table 3, fig. 2). The modelled temperature showed a slightly cold temperature bias (-0.047 ºC). The comparison between observed and modelled temperature at specific depths (0.5, 5, 10 and 15 m) showed a good model performance throughout the entire water column. At 0.5 m depth (fig. 3a), the calculated RMSE was 0.827 ºC and the MAE was 0.614 ºC. The modelled temperature at 0.5 m were slightly warmer than the observed water temperature, since the measured bias was 0.086 ºC. At 5 m depth (fig. 3b), RMSE and MSE were very similar to the ones calculated at 0.5 m with a value of 0.840 ºC and 0.618 ºC respectively. A slightly colder temperature bias was found (-0.004 ºC). The comparison of modelled and observed temperature at 10 m depth showed a RMSE of 1.187 ºC, a MAE of 0.811 ºC and a temperature bias of 0.003 ºC (fig. 4a). At 15 m depth, the RMSE was 1.155 ºC, the MAE was 0.803 ºC and the temperature bias was -0.137 ºC (fig. 4b).

3.2 Reconstructed thermal structure

The Pettitt test showed that a significant abrupt change in annual mean air temperature occurred in 1988 ($p < 0.001$). Therefore, in addition to checking for trends in lake thermal structure over the entire simulation period we also evaluated the possibility of trends occurring over the subinterval 1961-1988 and 1989-2017, and we tested whether a more rapid change in water temperature is occurring after 1988, following a step-change in annual mean temperature. The results suggest that Lake Erken did indeed change more rapidly since 1989. The Mann-Kendall test showed that during summer (Jun.-Aug.) a significant increase in whole-lake and epilimnetic temperature of 0.239 °C/decade (p-value < 0.001, fig. 5a) and of 0.351 °C/decade (p-value < 0.001, fig. 5b) respectively occurred over the entire study period (1961-2017) but not when the trend analysis was performed in the sub-intervals. Similarly, a slightly increasing trend was also detected for hypolimnetic spring temperature (0.148 °C/decade, p-value < 0.05, fig. 5c) over the entire study period but not in the sub-intervals. No other significant trends were detected over the entire simulation period or over the sub-interval 1961-1988. In contrast, the results suggest that Lake Erken did change more rapidly since 1989. Significant positive trends were detected from 1989 onwards during both the spring and autumn. Since 1989, spring (Apr. – May) whole-lake temperature showed an average increasing trend of +0.404 °C/decade (p-value < 0.05, fig. 5a) and epilimnetic temperature an average increasing trend of +0.444 °C/decade (p-value < 0.05, fig. 5b). The same pattern is shown during autumn months (September-November) with no trends detected in the sub-interval 1961-1987, while significant increasing trends were detected in the sub-interval 1988-2016 for whole-lake (+0.789 °C/decade, p-value < 0.001, fig. 5a), epilimnetic (+0.792 °C/decade, p-value <0.001, fig. 5b) and hypolimnetic temperature (0.816 °C/decade, p-value < 0.01, fig. 5c).

Other metrics of thermal statification showed long-term trends that were significant over the entire simulation period. The length of the growing season showed a positive significant increase, which, on average, was of 3.793 days/decade (p-value < 0.001, fig. 6) in the interval 1961-2017. With regards of thermal stability, the trend analysis of Schmidt stability revealed that more energy is required to mix the lake during stratified conditions in recent years if compared to the first years of the study period (5.365 Jm-2/decade, p-value < 0.01, fig. 7). This greater stability also corresponded with a longer duration of stratification. From 1961, the duration of lake stratification increased, on average, by 7.297 days/decade (p-value < 0.001, fig. 8a). The longer period of stratification is the result of both an earlier onset of thermal stratification, which now occurs on average ~16 days earlier since 1961 (-2.903 days/decade, p-value < 0.01, fig. 8b) and a later loss of thermal stratification that now is on average delayed by ~26 days (4.583 days/decade, p-value < 0.001, fig. 8c). The difference between surface and bottom temperature is often used as a simple indicator of thermal stratification. Its mean annual value during the stratified period increased significantly over time, increasing, on average, by 0.253 °C/ decade (p-value < 0.05, fig. 9). Mean annual thermocline depth during lake stratification period shows a significant decrease over the entire study period, with an average decrease of ~1.345 m since 1961 (-0.236 °C/decade, p-value < 0.01, fig. 10).

Regarding ice cover duration, the Pettitt test showed that an abrupt change in ice cover duration occurred in 1988 in the interval 1941-2017. Therefore, similarly to water temperature analysis, trend tests were performed in two sub-interval, 1941-1988 and 1989-2017. Trend in ice cover length did not significantly change within the sub-intervals. However, a significant decrease in ice cover length was detected when trend analysis were performed on the entire interval (-7.343 days/decade, p < 0.001, fig. 11).The Pettitt's test showed that a significant abrupt change in annual mean air temperature occurred in 1988 (p < 0.001). Therefore, in addition to checking for trend in lake thermal structure over the entire simulation period we also evaluated the possibility of trends occur over the period starting in 1988. The results suggest that Lake Erken did indeed change more rapidly since 1988. The Mann-Kendall test showed that during summer (June-August) a significant increase in epilimnetic temperature of +0.348 °C/decade (p-value < 0.001, fig. 5b) occurred over the entire study period (1961-2017). However, no other significant trends were detected over the entire simulation period or over the sub-interval 1961-1987. In contrast, significant positive trends were detected from 1988 onwards during both the spring and autumn. Since 1988, spring (April – May) whole-lake temperature showed an average increasing trend of +0.378 °C/decade (p-value < 0.05, fig. 5a) and epilimnetic temperature an average increasing trend of +0.43 °C/decade (p-value < 0.05, fig. 5b). The same pattern is showed during autumn months (September-November) with no trends detected in the sub-interval 1961-1987, while significant increasing trends were detected in the sub-interval 1988-2016 for whole-lake (+0.809 °C/decade, p-value < 0.001, fig. 5a) epilimnetic (+0.809 °C/decade, p-value <0.001, fig. 5b). Also during the autumn hypolimnetic temperature showed an increasing trend of +0.827 °C/decade (p-value < 0.001, fig. 5c). Other metrics of thermal statification showed long-term trends that were significant over the entire simulation period. The length of the growing season showed a positive significant increase, which, on average was of +3.684 days/decade (p-value < 0.001, fig. 6) in the interval 1961-2017. With regards of thermal stability, the trend analysis of Schmidt stability revealed that more energy is required to mix the lake during stratified conditions in recent years if compared to the first years of the study period (+5.535 Jm^{-2}/decade, p-value < 0.01, fig. 7). This greater stability also corresponded with a longer duration of stratification. From 1961, the duration of lake stratification increased, on average, by 7.083 days/decade (p-value < 0.001, fig. 8a).

The longer period of stratification is the result of both an earlier onset of thermal stratification, which now occurs on average ~16 days earlier since 1961 (-2.841 days/decade, p-value < 0.01, fig. 8b) and a later loss of thermal stratification that now is on average delayed by ~26 days (+4.545 days/decade, p-value < 0.001, fig. 8c). The difference between surface and bottom temperature is often used as a simple indicator of thermal stratification. Its mean annual value during the stratified period increased significantly over time, increasing, on average, by +0.249 °C/ decade (p-value < 0.05, fig. 9). Mean annual thermocline depth during lake stratification period shows a significant decrease over the entire study period, with an average decrease of ~1.242m since 1961 (-0.218 °C/decade, p-value < 0.01, fig. 10).

4. Discussion

The model used in this study revealed a good match between observed and simulated water temperature during the entire study period 1961-2017 (RMSE = 1.089 ºC, MAE = 0.753 ºC, bias = -0.047). In particular, the GOTM model was able to reproduce past water temperature with a high level of accuracy not only when meteorological driving data were available from the Erken field meteorological station (1988-2017), but also during the period 1961-1988, when most of the meteorological data were estimated using Artificial Neural Network Analysis. Indeed, the model was able to well describe summer water temperature during the three-year period 1961-1963, when it was possible to compare frequent water temperature measurements recorded from several depths by strip chart recorders (fig. 3-4)..

The model performance was very similar at 0.5 m and 5 m depth, with a RMSE of 0.827 ºC and 0.840 ºC respectively. Slightly greater errors were found at 10 m and 15 m depth, with a RMSE of 1.187 ºC and 1.155 ºC respectively. Since Erken is subjected to internal seiche movements, it is likely that higher errors found at deeper points (especially when a thermocline is present) between modelled and observed temperature could be at least partially explained by the limitation of a 1D model like GOTM to describe seiche movements. On seasonal basis, the model performed well in all the season (winter is excluded from the calibration). In spring (Apr.-May), the model showed a RMSE of 0.952 ºC, a MAE of 0.721 ºC and a very low temperature bias (-0.008 ºC). These values revealed that, despite the lack of a fully functioning ice-module in the GOTM version used here, the model performed very well during a period in which ice-cover could still occur in some years at Lake Erken. The good model performance during spring also adds confidence to the onset of stratification calculated from the modelled water temperature profiles, which often starts in April or May. In summer (Jun.-Aug.), the model showed slightly higher errors (RMSE = 1.240 ºC, MAE = 0.903 ºC and bias = -0.027 ºC). Lake Erken is always stratified during summer and, similarly to higher errors found at deeper points, higher errors of the model in this season could be related to seiche movements around the thermocline that are hard to predict with a 1D model. This is corroborated by the fact that in autumn (Sep.-Nov.), when the lake is fully mixed, the model performed better (RMSE = 0.530 ºC, MAE = 0.361 ºC, bias = -0.005 ºC).

T~~Moreover, t~~he Pettitt's test showed that an abrupt change in air temperature occurred in 1988, consistent with the results of Temnerud and Weyhenmeyer (2008), who detected that most of the abrupt changes in air temperature (interval 1961-2005) ~~across different sites in Sweden~~ occurred in 1988 and 1989 across different sites in Sweden. The similarity of these findings with the present work demonstrates the ~~reliability~~ consistency of air temperature data used to drive GOTM model, to that previously evaluated by Temnerud and Weyhenmeyer (2008), and also supports our finding that trends in water temperature were strongest during the period 1989-2017, and that this interval plays the most important role in defining the trends water temperature warming. ~~and also supports our finding that trends in water temperature were strongest during this period, which found that the 1988-2017 interval plays an important role in defining the trends water temperature warming.~~

Since the majority of the increasing trends were detected only in this sub-interval, it is apparent that most of the increase in Erken water temperature has occurred during the last three decades rather than during the entire study period. Overall, autumn is the season that showed the highest increase in water temperature between 1989-2017 (whole-lake: 0.789 °C/decade, epilimnion: 0.792 °C/decade, hypolimnion: 0.816 °C/decade). A lesser trend was detected during spring between 1989-2017 for whole-lake and epilimnetic temperature (0.404 °C/decade, and 0.444 °C/decade respectively). Summer whole-lake and epilimentic temperature showed a constantly increasing trend throughout the entire study period, with a significant increase since 1961 (0.239 °C/decade and 0.351 °C/decade respectively), but no significant increase in the sub-intervals, suggesting that the first and most persistent effects of global warming have occurred during summer. Otherwise, the more recent and more significant trends are most apparent in spring and autumn. These results also showed that while epilimnetic temperature increased in each season, hypolimnetic temperature showed a significant increase in autumn between 1989-2017 (0.816 °C/decade) in autumn and in spring, even though the trend detected in this season is pretty low (0.148 °C/decade over the interval 1961-2017), while no significant trends were detected in summer. The marked trend detected in autumn could results from the entrainment of warmer epilimnetic waters into the hypolimnion as the seasonal thermocline deepens. In general, however, the lower and fewer increasing trends detected in the hypolimnion compared to epilimnion and whole-lake temperature could be related to the fact that hypolimnetic temperature is less affected by meteorological variability than epilimnetic temperature (Adrian et al., 2009).

~~The majority of the increasing trends were detected only in this sub-interval, suggesting that most of the increase in Erken water temperature has occurred during the last three decades rather than during the entire study period. Overall, autumn is the season that showed the highest increase in water temperature between 1988-2017 (epilimnion: +0.809 °C/decade, hypolimnion: +0.827 °C/decade, whole-lake: +0.809 °C/decade). A lesser trend was detected during spring between 1988-2017 for epilimnetic and whole-lake temperature (+0.43 °C/decade, and +0.378 °C/decade respectively).Only summer epilimentic temperature showed a constantly increasing trend throughout the entire study period, with a significant increase since 1961, but no significant increase in the sub-intervals, suggesting that the first and most persistent effects of global~~

~~warming have occurred during summer, while the more recent and more significant trends are most apparent in spring and autumn. These results also show that while epilimnetic temperature increased in each season, hypolimnetic temperature showed a significant increase only in autumn between 1988-2017. This probably results from the fact that hypolimnetic temperature is less affected by meteorological variability than epilimnetic temperature (Adrian et al. 2009).~~

A large-scale study carried out by O'Reilly et al. (2015), compared trends in summer surface water temperature from 235 lakes located in different climatic regions between 1985-2009. A global mean trend of increasing summer surface water temperature (+0.34 °C/decade) was detected, with trends from individual lakes ranging from -0.7 - +1.3 °C /decade. The study reports a surface water temperature trend for Lake Erken derived from measured data of ~~+~~0.61 °C/decade (see O'Reilly et al. 2015, supporting information), while in the present work the trend of the surface summer modelled temperature ~~(0.5 m depth)~~ calculated over the same time period (1985-2009) is somewhat greater~~-~~: 1.145 °C/decade.~~+1.115 °C/decade~~.

The trend detected by O'Reilly et al. (2015) is calculated using a dataset with temporal gaps in summer water temperature record (14 years with recorded data between 1985-2009), while in the present work the trend was calculated using a complete long-term dataset of simulated summer surface water temperatures. Furthermore, the same article suggests that water temperature trends detected for lakes with data gaps might have been underestimated, suggesting that the trend detected in the present study for Lake Erken could be more accurate. This illustrates the value of using more complete and consistent modelled data to calculate trends, and also indicates that the lake is warming at a rate near the global maximum. The rapid rate~~d~~ of warming estimated from our work is also consistent with the conclusions of O'Reilly et al. (2015) that lakes located in Northern Europe are warming more rapidly than the global average, and also of Kraemer et al. (2017) that lakes at greater latitudes are warming faster than tropical lakes. Temperature trends obtained in the present study are consistent with these findings.

A prolonged duration of high surface water temperature and an increase of epilimnetic temperature can have impacts on lake mixing dynamics leading to a higher thermal stability (Jankowski et al. 2006;~~,~~ Butcher et al. 2015). Such increases in water temperature can explain why Schmidt stability has also increased over the period 1961-2017 (5.365 Jm^{-2}/decade~~5.353 Jm^{-2}/decade~~) and why the duration of stratification has also increased by about 40 days since 1961, shifting both the onset and the end of the stratification. Compared to 1961, the present onset of thermal stratification occurs on average 16 days earlier. However, the higher thermal stability has even a greater effect on the loss of stratification. From the simulated temperature, the end of the stratification now occurs on average 26 days later if compared with the 1960s. Very similar results were reported by Arvola et al. (2009) who used less frequently measured temperature data to estimate that the loss of stratification in Lake Erken was delayed by almost one month since the 1960, further verifying the reliability of the model based approach used in this study for detecting such variation. However, Arvola et al. (2009) did not detect the trends in the onset and duration of stratification or lake warming that were detected here. That these trends are now detected shows the value of using model simulations to provide long-term consistent temperature records that are more amenable to trend analysis.

Higher surface temperature has also increased the difference between surface and bottom temperature over the period 1961-2017 (0.253 ºC/decade) thereby increasing the gradient of the thermocline, while reducing the mean thermocline depth (-0.236 ºC/decade) ~~(figure 9) inducing a greater steepness of the thermocline, and that mean thermocline depth (figure 10) have~~

decreased over time. This could be due to a lower wind speed in recent times. Trend analysis of mean annual wind speed have revealed that there is a significant decreasing trend in wind speed over the study period. Since 1961, the wind speed has decreased on average by 0.775 m/s (Sen's slope = -0.136 ms^{-1}/decade, $p < 0.001$). Another possible explanation to a shallower thermocline could be related to a reduction of heat fluxes from water to air, which might have weakened the convective mixing of the upper layers (Monismith and MacIntyre 2009). However, heat fluxes have not been analyzed in this study, and further research is needed to better understand the causes behind reductions in thermocline depth.

A known limitation of the GOTM model version used in this work is the lack of an ice-module to simulate the onset, the loss and the duration of the lake ice-cover. Simulated and observed temperature close to the surface, are similar throughout the entire year. However, during winter months (Dec.-Mar.) the model does not take into account the heat loss from sediment during ice-cover, which cause an increase in bottom water temperature. Despite the lack of simulated ice cover, it was possible to analyse the duration of ice cover thanks to a yearly observations of the onset and loss of ice cover made at Lake Erken since 1941. Ice cover duration is decreased since 1941 of 7.343 days/decade. A step-change in ice cover duration was detected in 1988, consistent with the step change in air temperature detected in the present study and by Temnerud and Weyhenmeyer (2008). When trend analysis was performed on the two sub-intervals 1941-1988 and 1989-2017 no significant trends in ice cover duration were detected. However, the ice cover duration showed a greater interannual variability in the sub-interval 1989-2017 compared to the sub-interval 1941-1988. Within the sub-interval 1941-2017, the variability in ice cover duration ranges from a minimum of 68 days in 1961 to a maximum of 168 days in 1958, while in the sub-interval 1989-2017 the ice cover duration ranges from 12 days in 2008 to 142 days in 2011. Such increase in variability of ice cover duration could be related to warming of climatic conditions (Magnuson et al., 2000). Sadro et al. (2019) found that decline of snowpack in mountain lakes in Sierra Nevada (California) causes a warming response in lake temperature. Since Lake Erken and most lake in Scandinavia are always ice covered during winter and snowfall occurs every year, understanding the dynamics of snowfall and ice cover phenology could be of extreme importance to better understand thermal response of lakes to climate change.

Changes in lake water temperature and stratification patterns can have a broad influence on many aspects of lake ecosystems, both biotic and abiotic. For example, a longer duration of thermal stratification could lead to a depletion in hypolimnetic oxygen (Jankovski et al., 2006;, Butcher et al., 2015), potentially reducing the natural range of lacustrine fish (Jones et al. 2008) or otherwise influence the vertical distribution of living organisms (Woolway et al., 2014). Moreover, an earlier onset of thermal stratification and warmer lake temperature could change the seasonal dynamics of phytoplankton species (Thackeray et al., 2008). A previous model simulation conducted by Blenckner et al. (2002) on the ecology of Lake Erken concluded that warmer water temperature and changes in mixing dynamics due to climate change are likely to boost nutrient concentration and phytoplankton production, with consequences for the entire lake ecosystem in the coming decades. Given the relative long retention time of lake Erken (7 years), the importance of internal phosphorus loading due to changes in thermal stratification could make the lake more susceptible to climate change than other Swedish lakes with shorter retention times and higher levels of external nutrient loading (Malmaeus et al., 2005). Moreover, the relative long retention time of lake Erken (7 years) could enhance the importance of internal phosphorus loading due to warmer temperature, making the lake more susceptible to climate change than other Swedish lakes with shorter retention time (Malmeus et al. 2005). Another related issuepotential impact could be an increase of carbon emission from the lake, since a recent study has revealed that an increase in nutrient concentration coupled with a rise in water temperature can have a positive and synergistic effect on methane ebullition (Davidson et al. 2018). In a warmer world not only methane, but also CO_2 emissions from boreal lakes are likely to increase (Weyhenmeyer et al. 2015). Finally, a general indicator of the effects of warmer conditions on the biological dynamics of

5. Conclusion

The present study has shown that the use of the GOTM model to reconstruct the past 57-years of thermal condition of Lake Erken provided a valuable source of information that could be used to detect changes in its thermal structure. This methodology can be extended to other lakes that have incomplete records of water temperature data. The use of local meteorological data to drive model simulations such as those demonstrated here can be used to extend water temperature records further back in time or fill data gaps where they exist. This work also shows that water temperature has been rising faster in the last three decades compared to the previous decades both in the epilimnion and hypolimnion and that other metrics describing thermal stratification have changed over the entire 57-year study period.

It is likely that increasing water temperature will cause many secondary effects with serious and to some extent unpredictable repercussions on lake ecosystems. This work can be seen as a baseline for future research on Lake Erken that involve climate-related investigations. A further step towards a better understanding of how the lake ecosystems will respond to climate change is to couple a biogeochemical model with the physical model GOTM using FABM - Framework for Aquatic Biogeochemical Models (Bruggemann and Bolding, 2014). Then parameters such as chlorophyll, nutrient and dissolved organic carbon concentrations can be simulated and analyzed. The coupling of physical and biogeochemical models could, therefore, be a valuable tool to facilitate the mitigation of detrimental effects of a warmer world on lake ecosystems.Lake Erken is the growing season indicator of Håkanson and Boulion (2001). Our simulations showed a significant increase in the number of days in which epilimnetic temperature was greater than the suggested 9 ° C threshold during the 1961-2017 study period (+3.684 days/decade).

The present study has shown that the GOTM model accurately reconstructed the past 57-years of thermal conditions in Lake Erken, and the use of modelled data was a valuable tool for detecting changes in its thermal structure. This work also pointed shows that water temperature has been rising faster in the last three decades compared to the previous decades both in the epilimnion and hypolimnion and that other metrics describing thermal stratification have changed over the entire 57-year study period.

In conclusion, it is likely that increasing water temperature will cause many secondary effects with serious and to some extent unpredictable repercussions on lake ecosystems. This work can be seen as a baseline for future research on Lake Erken that involve climate-related investigations. A further step towards a better understanding of how the lake ecosystems will respond to climate change is to couple a biogeochemical model with the physical model GOTM using FABM - Framework for Aquatic Biogeochemical Models (Bruggemann and Bolding 2014). Then parameters such as chlorophyll, nutrient and dissolved organic carbon concentrations can be simulated and analyzed. The coupling of physical and biogeochemical models could, therefore, be a valuable tool to facilitate the mitigation of detrimental effects of a warmer world on lake ecosystems.

Data and code availability

The model configuration, the input data used to run the GOTM model, the output data and the water temperature used to calibrate the model are available on Hydroshare (doi:10.4211/hs.54375615d258461086125d5fc85a4c32doi: 10.4211/hs.7e5ec8c0e2b245199ab13cc9ae08b841). Matlab codes, R codes and all the datasets produced during this study are available upon request from the corresponding author.

Competing interests

The authors declare no competing interests.

Acknowledgements

The authors would like to thank the EU and FORMAS for funding, 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).

References

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingstone, D. M., Sommaruga, R., Straile, D., Van Donk, E., Weheyenmeyer, G. A., and Winder, M.: Lake as sentinel of climate change, Limnol. Oceanogr., 54, 2283–2297, doi: 10.4319/lo.2009.54.6_part_2.2283, 2009.

Arhonditsis, G. B., Brett, M. T., DeGasperi, C. L., and Schindler, D. E.: Effects of climatic variability on the thermal properties of Lake Washington, Limnol. Oceanogr., 49, 256–270, doi:10.4319/lo.2004.49.1.0256, 2004

Arvola, L., George, G., Livingstone, D. M., Järvinen, M., Blenckner, T., Dokulil, M. T., Jennings, E., Aongusa, C. N., Nõges, P., Nõges, T., and Weheyenmeyer, G. A.: The Impact of the Changing Climate on the Thermal Characteristics of Lakes, in: The Impact of Climate Change on European Lakes, edited by: George, G., Aquat. Ecol. Series, 4, 85-101, doi: 10.1007/978-90-481-2945-4_6, 2009.

Blenckner, T., Omstedt, A., and Rummukainen, M.: A Swedish case study of contemporary and possible future consequences of climate change on lake function, Aquat. Sci., 64, 171–184, doi:10.1007/s00027-002-8065-x, 2002

Bruggeman, J. and Bolding, K.: A general framework for aquatic biogeochemical models, Environ. Modell. Software, 61, 249-265, doi: 10.1016/j.envsoft.2014.04.002, 2014.

Burchard, H.: Applied turbulence modeling in marine waters, Lect. Notes Earth Sci., doi: 10.1007/3-540-45419-5, 2002.

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, doi:10.1007/s10584-015-1326-1, 2015.

Davidson, T. A., Audet, J., Jeppesen, E., Landkildehus, F., Lauridsen, T. L., Søndergaard, M., and Syväranta, J.: Synergy between nutrients and warming enhances methane ebullition from experimental lakes, Nat. Clim. Change, 8, 156-160, doi:10.1038/s41558-017-0063-z, 2018.

Håkanson, L. and Boulion, V. V.: A practical approach to predict the duration of the growing season for European lakes, Ecol. Modell., 140, 235–245, doi:10.1016/s0304-3800(01)00319-2, 2001.

Hamed, K. H. and Rao, A. R.: A Modified Mann-Kendall Trend Test for Autocorrelated Data, J. Hydrol., 204, 182-196, doi: 10.1016/S0022-1694(97)00125-X, 1998.

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

~~Håkanson, L. and Boulion, V. V.: A practical approach to predict the duration of the growing season for European lakes, Ecol. Modell., 140, 235–245, doi:10.1016/s0304-3800(01)00319-2, 2001.~~

IPCC 2013. Summary for Policymakers, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and V., Midgley, P. M., 3-29, Cambridge University Press, New York, NY, USA, 2013.

Jankowski, T., Livingstone, D. M., Bührer, H., Forster, R., and Niederhauser, P.: Consequences of the 2003 European heat wave for lake temperature profiles, thermal stability, and hypolimnetic oxygen depletion: Implications for a warmer world, Limnol. Oceanogr., 51, 815–819, doi:10.4319/lo.2006.51.2.0815, 2006.

Jones, I., D., Winfield, I., J., and Carse, F.: Assessment of long-term changes in habitat availability for Arctic charr (Salvelinus alpinus) in a temperate lake using oxygen profiles and hydroacoustic surveys, Freshwater Biol., 53, 393–402, doi:10.1111/j.1365-2427.2007.01902.x, 2008.

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, doi:10.1080/20442041.2017.1294332 2017.

Kendall, M. G.: Rank Correlation Methods, Charles Griffin, London, 1975.

Kraemer, B. M., Chandra, S., Dell, A. I, Dix, M., Kuusisto, E., Livingstone, D. M., Schladow, S. G., Silow, E., Sitoki, L. M., Tamatamah, R., and McIntyre, P. B.: Global patterns in lake ecosystem responses to warming based on the temperature dependence of metabolism, Global Change Biol., 23, 1881–1890, doi:10.1111/gcb.13459, 2017.

Livingstone, D. M., and Dokulil, M. T.: Eighty years of spatially coherent Austrian lake surface temperatures and their relationship to regional air temperature and the North Atlantic Oscillation, Limnol. Oceanogr., 46, 1220-1227, doi: 10.4319/lo.2001.46.5.1220, 2001.

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., Assel, R. A., Barry, R. G., Card, V., Kuusisto, E., Granin, N. G., Prowse, T. D., Stewart, K. M., and Vuglinski, V. S.: Historical trends in lake and river ice cover in the Northern Hemisphere, Science, 289, 1743-1746, doi: 10.1126/science.289.5485.1743, 2000.

Malmaeus, J. M., Blenckner, T., Markensten, H., and Persson, I.: Lake phosphorus dynamics and climate warming: A mechanistic model approach. Ecol. Modell., 190, 1–14, doi:10.1016/j.ecolmodel.2005.03.017, 2005.

Mann, H. B.: Non-parametric tests against trend, Econometrica, 13, 245–259, 1945.

MATLAB. 1992-2017. Neural Network Toolbox. The MathWorks Inc., Natick, Massachusetts, United States.

Monismith, S. G., MacIntyre, S.: The Surface Mixed Layer in Lakes and Reservoirs, in: Encyclopedia of inland waters, edited by: Likens G. E., Elsevier, Oxford, 2009.

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, 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, doi:10.1002/2015GL066235, 2015.

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

Pettersson K.: Limnological Studies in Lake Erken Sweden, in: Encyclopedia of Lakes and Reservoirs, Encyclopedia of Earth Sciences Series, edited by: Bengtsson L., Herschy, R. W, and Fairbridge, R. W., Springer, Dordrech, doi:10.1007/978-1-4020-4410-6_120, 2012.

Pettitt A. N.: A non-parametric approach to the change-point problem, Appl. Stat., 28, 126-135, doi:10.2307/2346729, 1979.

Pierson, D. C., Pettersson, K., and Istvanovics, V.: Temporal changes in biomass specific photosynthesis during the summer: regulation by environmental factors and the importance of phytoplankton succession. Hydrobiologia 243-244, 119–135, 1992.

R Studio Team 2016, RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, URL http://www.rstudio.com/.

~~Read, J. S., Hamilton, D. P., Jones, J. 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. Modell. Software, 26, 1325–1336, doi: 10.1016/j.envsoft.2011.05.006, 2011.~~

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, doi:10.2166/wqrjc.2012.020, 2012.

Sadro, S., Melack, J. M., Sickman, J. O., and Skeen, K.: Climate warming response of mountain lakes affected by variations in snow, Limnol. Oceanogr. Lett., 4, 9-17, doi:10.1002/lol2.10099, 2019.

Schmidt, W.: Über Temperatur und Stabilitätsverhaltnisse von Seen, Geographiska Annaler 10, 145-177, 1928.

Schneider, P., and Hook, S., J.: Space observations of inland water bodies show rapid surface warming since 1985, Geophys. Res. Lett., 37, L22405, doi:10.1029/2010GL045059, 2010.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's Tau, J. Am. Stat. Assoc., 63, 1379–1389, doi:10.2307/2285891, 1968.

Sharma, S., et al.: A global database of lake surface temperatures collected by in situ and satellite methods from 1985–2009, Sci. Data, 2, doi:10.1038/sdata.2015.8, 2015.

Skowron, R.: Water temperature in investigations of Polish lakes, Limnol. Rev., 17, 31-46, doi:10.1515/limre-2017-0004, 2017.

Stefan, H. G., Fang, X., and Hondzo M.: Simulated climate change effect on year-round water temperatures in temperate zone lakes, Clim. Change, 40, 547–576, doi:10.1023/A:1005371600527, 1998.

Storn, R. and Price K.: Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, J. Global Optim., 11, 341–359, doi:10.1023/A:100820282, 1997.

Taner, M. Ü., Carleton, J. N., and Wellman, M.: Integrated model projections of climate change impacts on a North American lake, Ecol. Modell., 222, 3380 – 3393, doi: 10.1016/j.ecolmodel.2011.07.015, 2011.

Temnerud, J. and Weyhenmeyer G. A.: Abrupt changes in air temperature and precipitation: Do they matter for water chemistry?, Global Biogeochem. Cycles, 22, GB2008, doi:10.1029/2007GB003023, 2008.

Thackeray, S. J., Jones I. D., and Maberly S. C.: Long-term change in the phenology of spring phytoplankton: Species-specific responses to nutrient enrichment and climatic change, J. Ecol., 96, 523–535, doi:10.1111/j.1365-2745.2008.01355.x, 2008.

Vincent, W. F.: Effects of climate change on lakes, in: Encyclopedia of Inland Waters, 55-60, edited by: Likens G.E., Academic Press, Oxford, doi:10.1016/B978-012370626-3.00233-7, 2009.

Weyhenmeywer, G. A., Kosten, S., Wallin, M. B., Tranvik, L. J., Jeppesen, E., and Roland F.: Significant fraction of CO2 emissions from boreal lakes derived from hydrologic inorganic carbon inputs, Nat. Geosci., 8, 933-938, doi: 10.1038/ngeo2582, 2015.

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, doi: 10.1038/sdata.2017.53, 2017.Winslow, L., Read, J., Woolway, R., Brentrup, J., Leach, T.,Zwart, J., Albers, S., and Collinge, D.: rLakeAnalyzer: Lake Physics Tools. R package version 1.11.4. https://CRAN.R-project.org/package=rLakeAnalyzer, 2018.

Woolway, R. I., Maberly, S. C., Jones, I. D., and Feuchtmayr, H.: A novel method for estimating the onset of thermal stratification in lakes from surface water measurements, Water Resour. Res., 50, 5131–5140, doi:10.1002/2013wr014975, 2014.

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

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

Tables

Table 1: Number of data points retrieved from difference sources of the driving meteorological parameters for GOTM simulation

Meteorological parameter	**No. of data retrieved from Erken or Svanberga meteorological station**	**No. of data retrieved from neighbouring meteorological stations**	**No. of interpolated data**	**Total data**
WS	282389	188567	28700	499656
Air T	235250	234982	29424	499656
RH	191678	294092	13886	499656
Air P	194881	259563	45212	499656
SWR	398129	101520	7	499656
CC	-	157948	341708	499656
DP	10016	10803	-	20819

~~**Table 1: Best parameter set and model statistics from ACPy calibration**~~ **Best parameter set from ACPy over the entire calibration period (Apr.-Nov.) between 1961-2017**

Model parameter	**Calibrated factor**	**Feasible range**
Heat-flux factor	0.863	0.5-1.5
Short- wave radiation factor	0.971	0.8-1.2
Wind factor	1.287	0.5-2.0
Minimum turbulent kinetic energy	$1.649e^{-6}$	$1.4e^{-7}$-$1.0e^{-5}$
e-folding depth for visible fraction	2.637	0.5-3.5

Table 3: Model performance over the entire calibration period (Apr.-Nov.) and different seasons (spring, summer, autumn)

Calibration interval	Model statistics	Value
Apr.-Nov. (full period)	ln Likelihood	-60469.700
	Bias (°C)	-0.047
	MAE (°C)	0.753
	RMSE (°C)	1.089
Apr.-May (spring)	ln Likelihood	-7782.650
	Bias (°C)	-0.008
	MAE (°C)	0.721
	RMSE (°C)	0.952
Jun.-Aug. (summer)	ln Likelihood	-39818.975
	Bias (°C)	-0.027
	MAE (°C)	0.903
	RMSE (°C)	1.240
Sep.-Nov. (autumn)	ln Likelihood	4122.193
	Bias (°C)	-0.005
	MAE (°C)	0.361
	RMSE (°C)	0.530

Table 4: Trend analysis results of the investigated lake metrics

Lake metrics	Time interval	Mann-Kendall τ	Sen's slope	Sen's Slope 95 % CI	P-value
Whole-lake spring	1961-1988	-	-	-	> 0.05
	1989-2017	0.305	0.404 ºC/decade	[0.076-0.827]	< 0.05
Whole-lake summer	1961-2017	0.308	0.239 ºC/decade	[0.119-0.381]	< 0.001
	1961-1988	-	-	-	> 0.05
	1989-2017	-	-	-	> 0.05
Whole-lake autumn	1961-1988	-	-	-	> 0.05
	1989-2017	0.444	0.789 ºC/decade	[0.265-1.273]	< 0.001
Epilimnion spring	1961-1988	-	-	-	> 0.05
	1989-2017	0.296	0.444 ºC/decade	[0.062-0.932]	< 0.05
Epilimnion summer	1961-2017	0.326	0.351 ºC/decade	[0.164-0.540]	< 0.001
	1961-1988	-	-	-	> 0.05
	1989-2017	-	-	-	> 0.05
Epilimnion autumn	1961-1988	-	-	-	> 0.05
	1989-2017	0.455	0.792 ºC/decade	[0.248-1.262]	< 0.001
Hypolimnion spring	1961-2017	0.187	0.148 ºC/decade	[0.007-0.294]	< 0.05
	1961-1988	-	-	-	> 0.05
	1989-2017	-	-	-	> 0.05
Hypolimnion summer	1961-2017	-	-	-	> 0.05
Hypolimnion autumn	1961-1988	-	-	-	> 0.05
	1989-2017	0.392	0.816 ºC/decade	[0.262-1.323]	< 0.01
Growing season (epi T > 9 ºC)*	1961-2017	0.380	3.793 days/decade	[2.222; 5.319]	< 0.001
Schmidt stability	1961-2017	0.256	5.365 Jm^{-2}/decade	[1.900; 9.023]	< 0.01
Stratification duration*	1961-2017	0.420	7.297 days/decade	[4.667-10.500]	< 0.001
Onset of stratification	1961-2017	-0.266	-2.903 days/decade	[-4.314; -0.889]	< 0.01
End of stratification*	1961-2017	0.397	4.583 days/decade	[2.593; 6.250]	< 0.001
$T_{surface}$-T_{bottom}	1961-2017	0.212	0.253 ºC/decade	[0.048; 0.464]	< 0.05
Thermocline depth	1961-2017	-0.268	-0.236 m/decade	[-0.380; -0.074]	< 0.01
Ice cover length*	1941-2017	-0.307	-7.343 days/decade	[-11.364; -3.438]	< 0.001

*Modified Mann-Kendall test (Hamed and Rao, 1998)

Figures

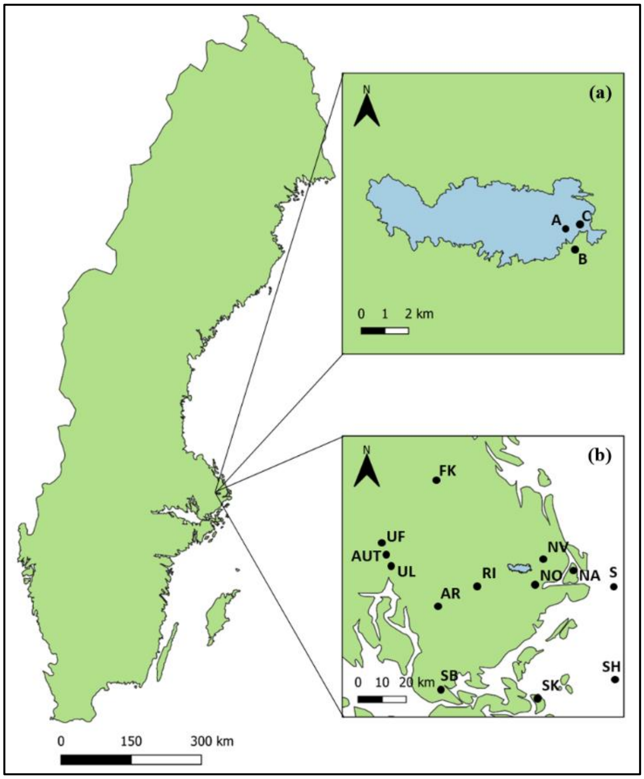


Figure 1: (a) Location of meteorological and floating stations within Lake Erken basin and catchment area. Letter A shows the position of Erken laboratory meteorological station (Malma islet, 59.83909 N, 18.62956 E), letter B identifies Svanberga SMHI weather station (59.8321 N, 18.6348 E), letter C represents the position of the floating station that records water temperature data. (b) Location of SMHI (Swedish Meteorological and Hydrological Institute) and SLU (Swedish Agricultural University) meteorological stations from which input data have been retrieved to run the model. SMHI stations: AR = Arlanda (59.6557 N, 17.9462 E), AUT = Uppsala AUT (59.8586 N, 17.6253 E), FK = Films Kirkby (60.2363 N, 17.9078 E), NA = Norrveda (59.8298 N, 18.9524 E), NO = Norrtälje (59.7506 N, 18.7091 E), NV = Norrtälje-Vasby (59.8524 N, 18.7296 E), RI = Rimbo (59.7487 N, 18.3535 E), S = Söderarm (59.7538 N, 19.4089 E), SB = Stockholm-Bromma (59.3537 N, 17.9513 E), SH = Svenska Hogarna (59.4445 N, 19.5059 E), SK = Skarpö A (59.3455 N, 18.7406 E), UF = Uppsala Flygplats (59.8953 N, 17.5935 E). SLU station: UL = Ultuna (59.8175 N, 17.6536 E).

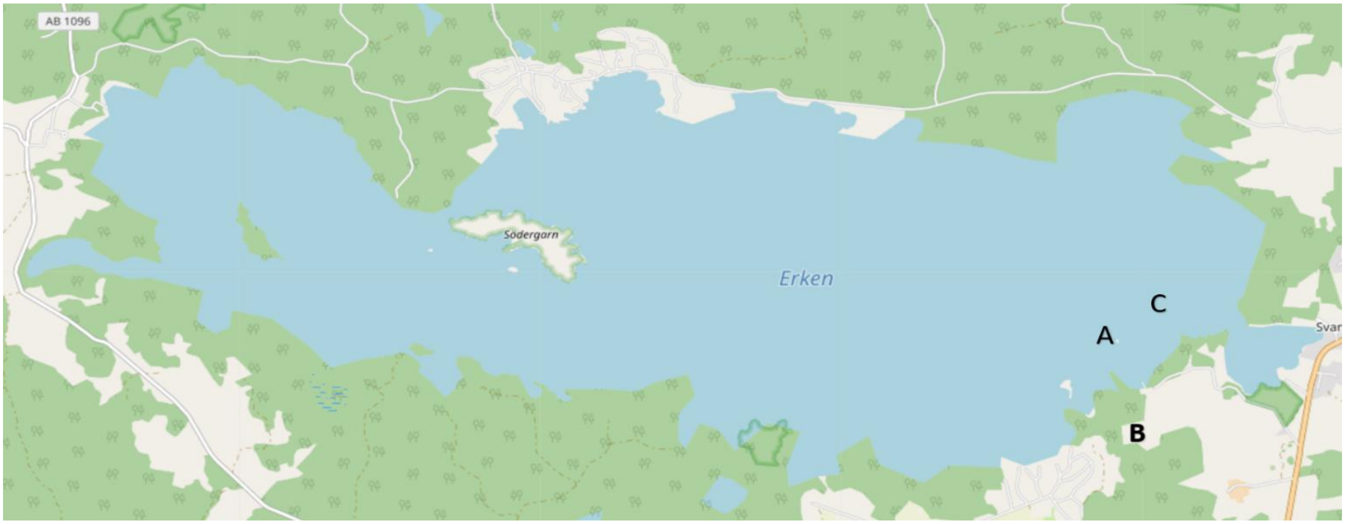


Figure 1: Position of meteorological and floating stations within Lake Erken basin and catchment area. Letter A shows the position of Malma meteorological station (59.83909 N, 18.62956 E), letter B identifies Svanberga SMHI weather station (59.8321 N, 18.6348 E), letter C represents the position of the floating station that records water temperature data (map retrieved from OpenStreetMap: https://www.openstreetmap.org.

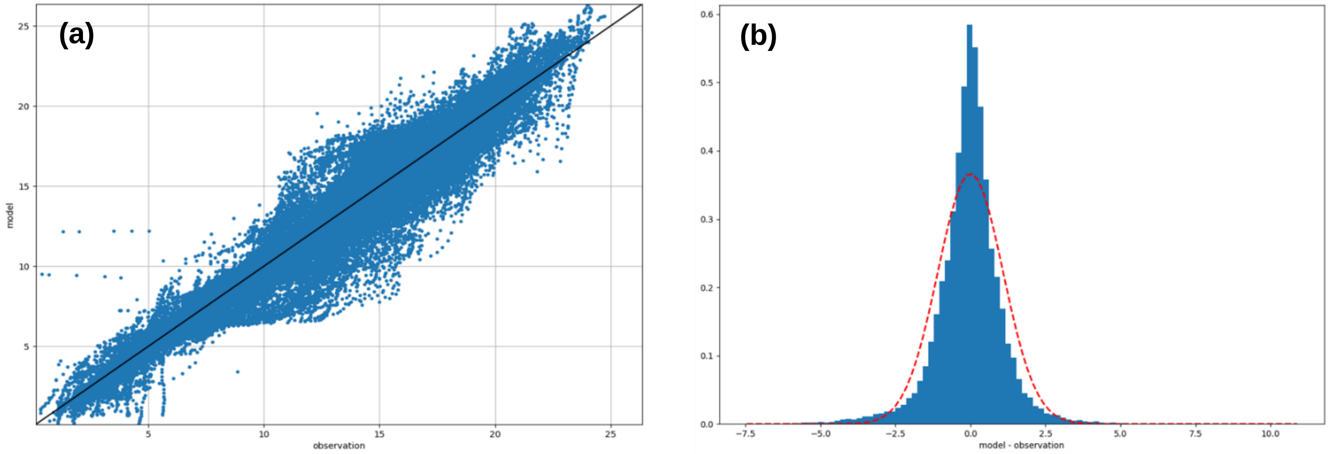
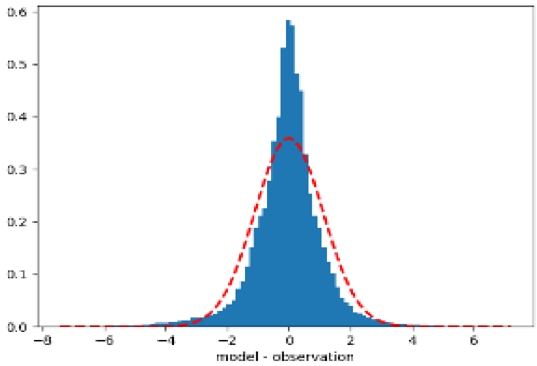


Figure 2: (a) Comparison between observed water temperature and simulated water temperature of Lake Erken in the interval 1961-2017 (correlation = 0.972) and (b) error distribution between modelled water temperature and observed water temperature (Model-Observation) retrieved from ACPy calibration (panel b). RMSE = 1.089, MAE = 0.7529, bias = -0.04707.



~~**Figure 2: Error distribution between modelled water temperature and observed water temperature (Model-Observation) retrieved from ACPy calibration.**~~

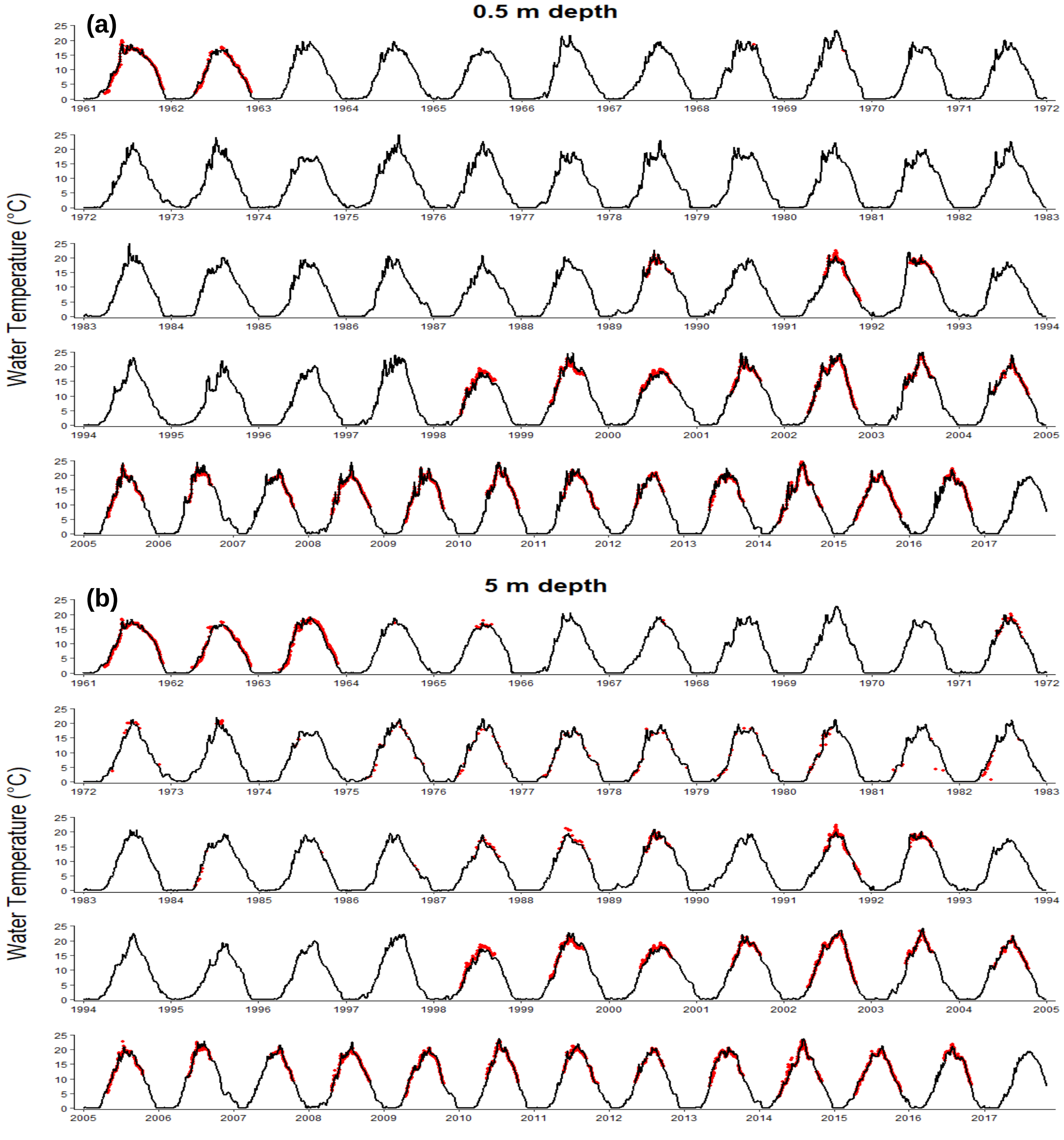
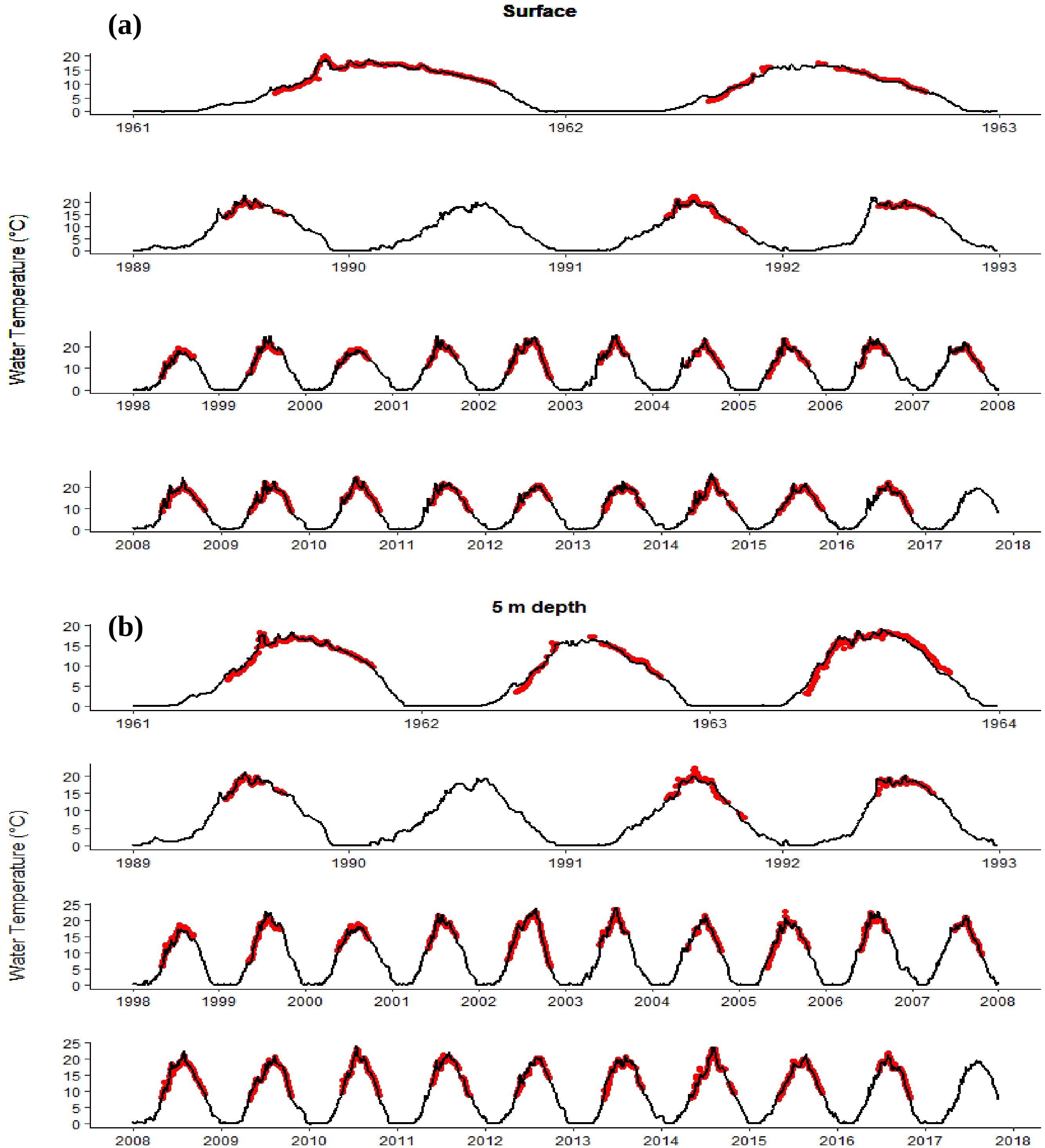


Figure 3: Comparison between Erken modeled (black line) and observed daily temperature (red dots) at 0.5m (a), and 5m (b) depth (0.5m depth: RMSE = 0.827 ºC, MAE = 0.614 ºC, bias = 0.086 ºC; 5m depth: RMSE = 0.840 ºC, MAE = 0.618 ºC, bias = -0.004 ºC).



~~Figure 3: Comparison between Erken modeled (black) and observed daily temperature (red) at surface (a), and 5m (b) depth. For better visualization of the results, years in which observed data were scarce are omitted.~~

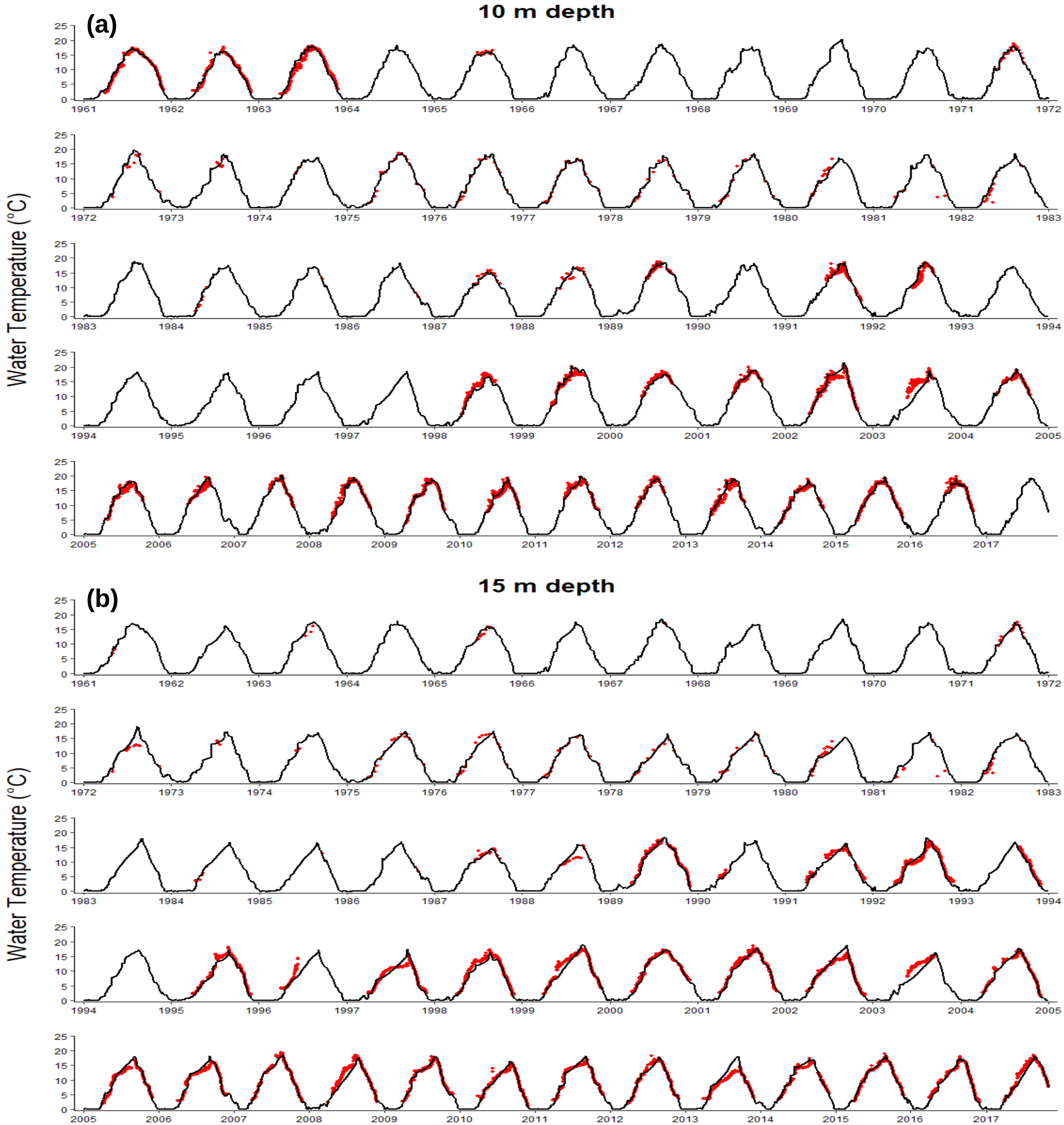
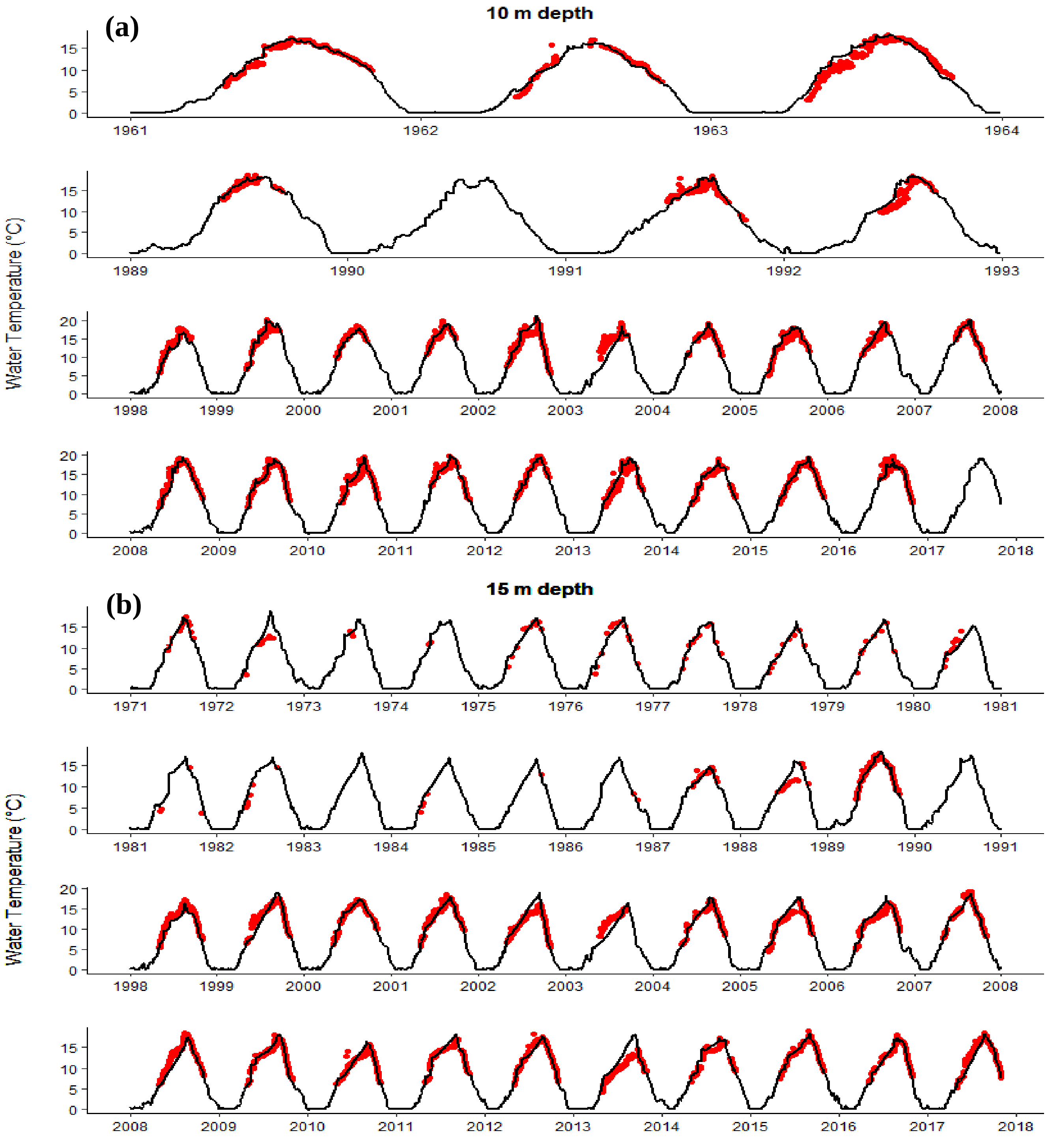


Figure 4: Comparison between Erken modeled (black line) and observed daily temperature (red dots) at 0.5m (a), and 5m (b) depth (10m depth: RMSE = 1.187 ºC, MAE = 0.811 ºC, bias = 0.003 ºC; 15m depth: RMSE = 1.155 ºC, MAE = 0.803 ºC, bias = -0.137 ºC).



~~**Figure 4: Comparison between Erken modeled (black) and observed daily temperature (red) 10m (a) and 15m (b) depth. For better visualization of the results, years in which observed data were scarce are omitted.**~~

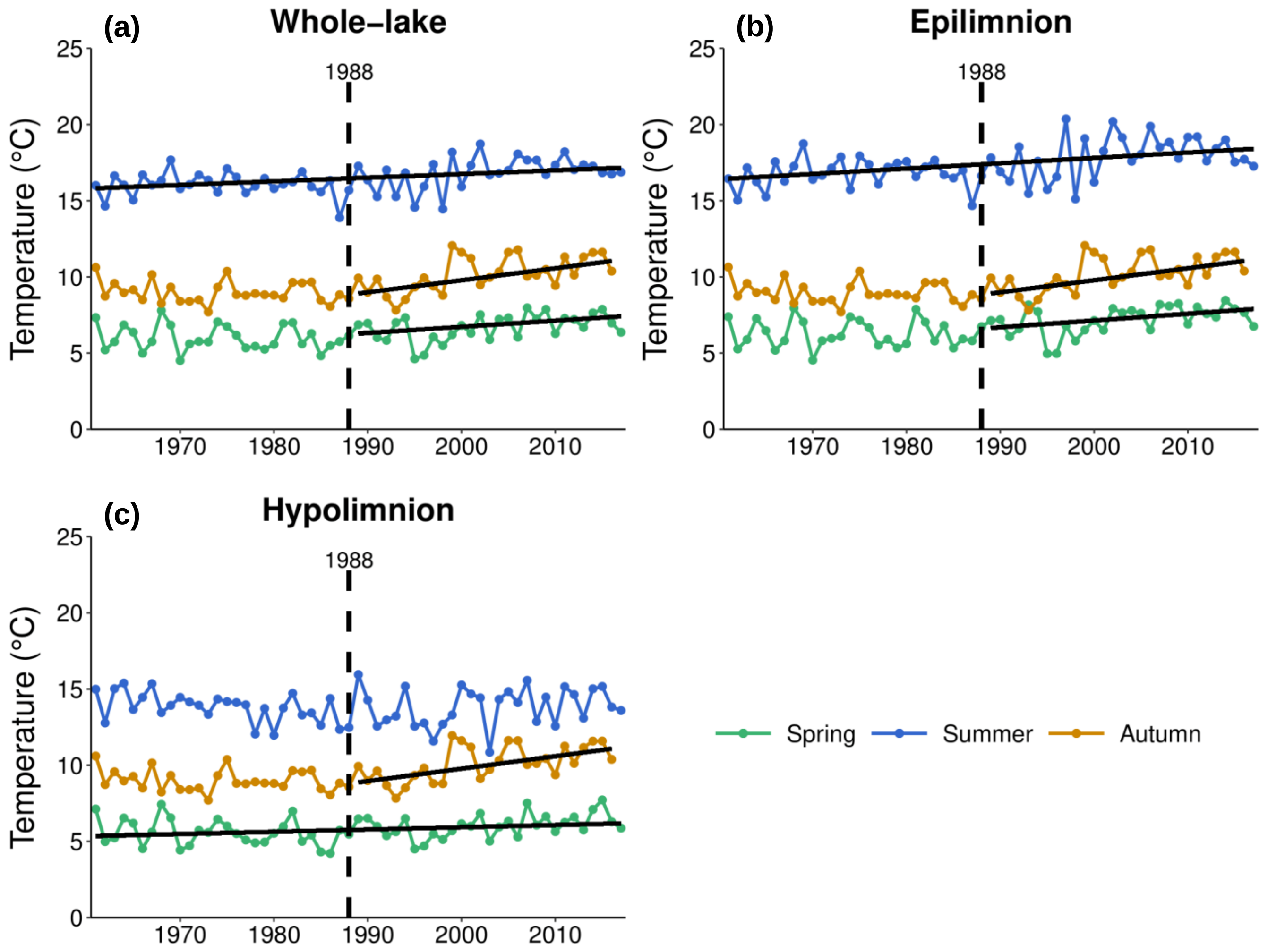
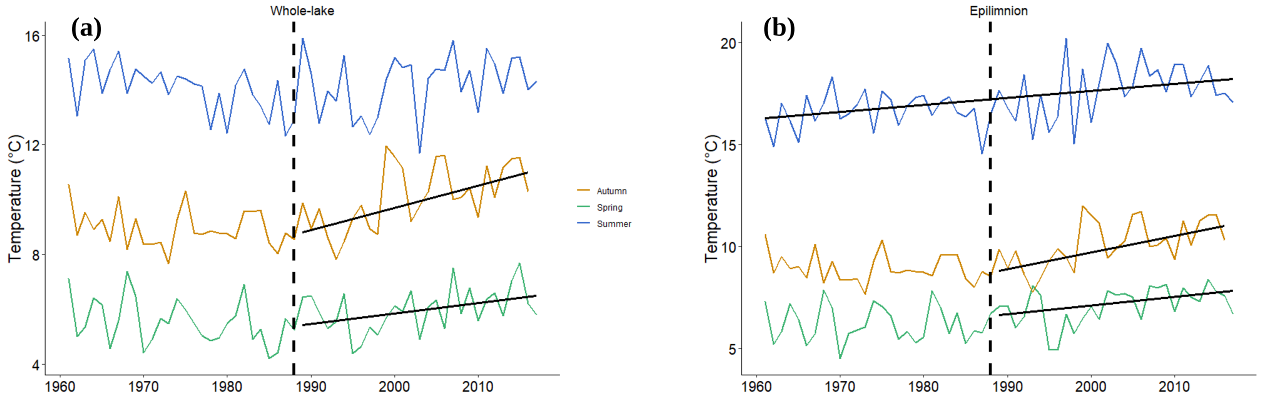
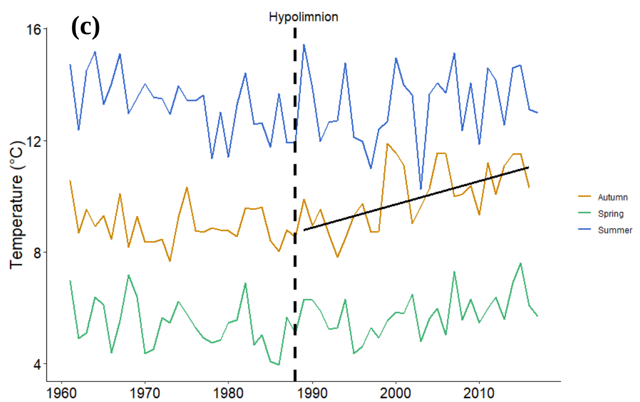


Figure 5: Mean seasonal water temperature and relative trends for whole lake (a), epilimnion (b) and hypolimnion (c). between 1961-2017. Besides performing trend analysis over the entire study period, the Mann-Kendall analysis was performed in two sub-intervals (1961-1988 and 1989-2017) that are divided by the dashed line. Only significant trend lines are displayed. Whole-lake temperature significant trends: spring = 0.404 °C/decade, summer = 0.239 °C/decade, autumn = 0.789 °C/decade. Epilimnetic temperature significant trends: spring = 0.444 °C/decade, summer = 0.351 °C/decade, autumn = 0.792 °C/decade. Hypolimnetic temperature significant trends: spring = 0.148 °C/decade, autumn = 0.816 °C/decade.





~~**Figure 5: Mean seasonal water temperature for whole lake (a), epilimnion (b) and hypolimnion (c). Only the significant trend lines are displayed. The dashed lines divide the two subintervals in which the Mann-Kendall analysis was performed (1961-1987 and 1988-2017). Significant spring trends: whole-lake = +0.378 °C/decade, epilimnion= +0.43 °C/decade. Significant summer trends: epilimnion = +0.348 °C/decade. Significant autumn trends: whole-lake = +0.809 °C/decade, epilimnion = +0.809 °C/decade, hypolimnion = +0.827 °C/decade.**~~

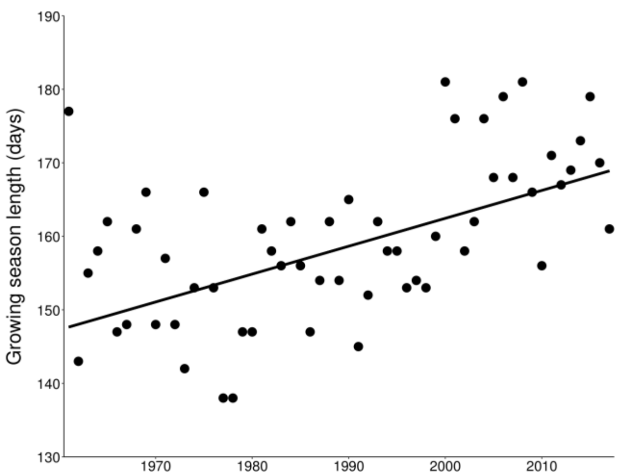


Figure 6: Growing season length calculated from simulated water temperature between 1961-2017 (Sen's slope: +3.793 days/decade, $p < 0.001$).

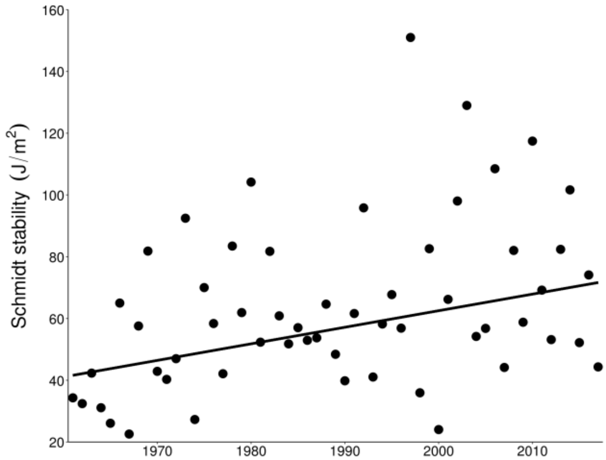
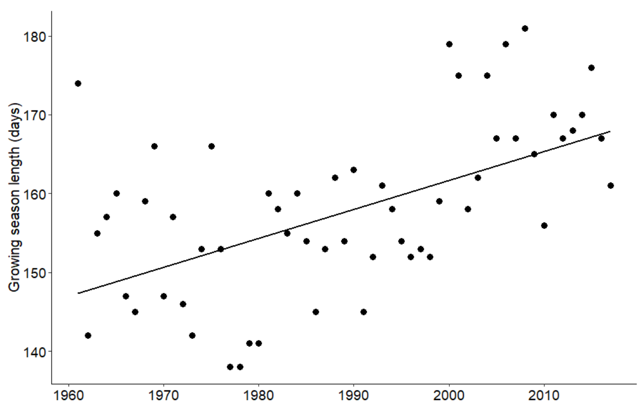
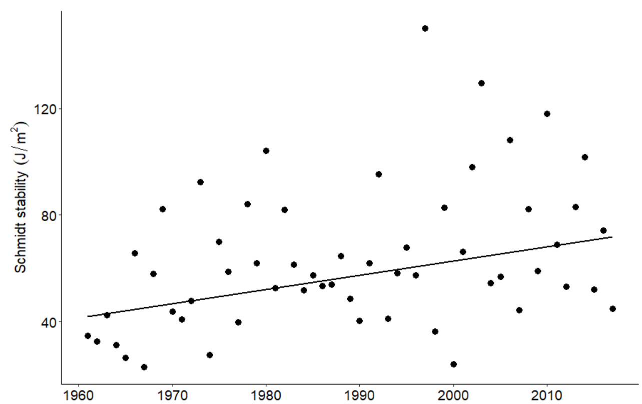


Figure 7: Schmidt stability calculated from simulated water temperature between 1961-2017 (Sen's slope: 5.365 Jm^{-2}/decade, $p < 0.01$).



~~**Figure 6: Growing season length calculated from simulated water temperature between 1961-2017 (Sen's slope: +3.684 days/decade, p < 0.001).**~~



~~**Figure 7: Schmidt stability calculated from simulated water temperature between 1961-2017 (Sen's slope: +5.535 Jm-2/decade, p < 0.01).**~~

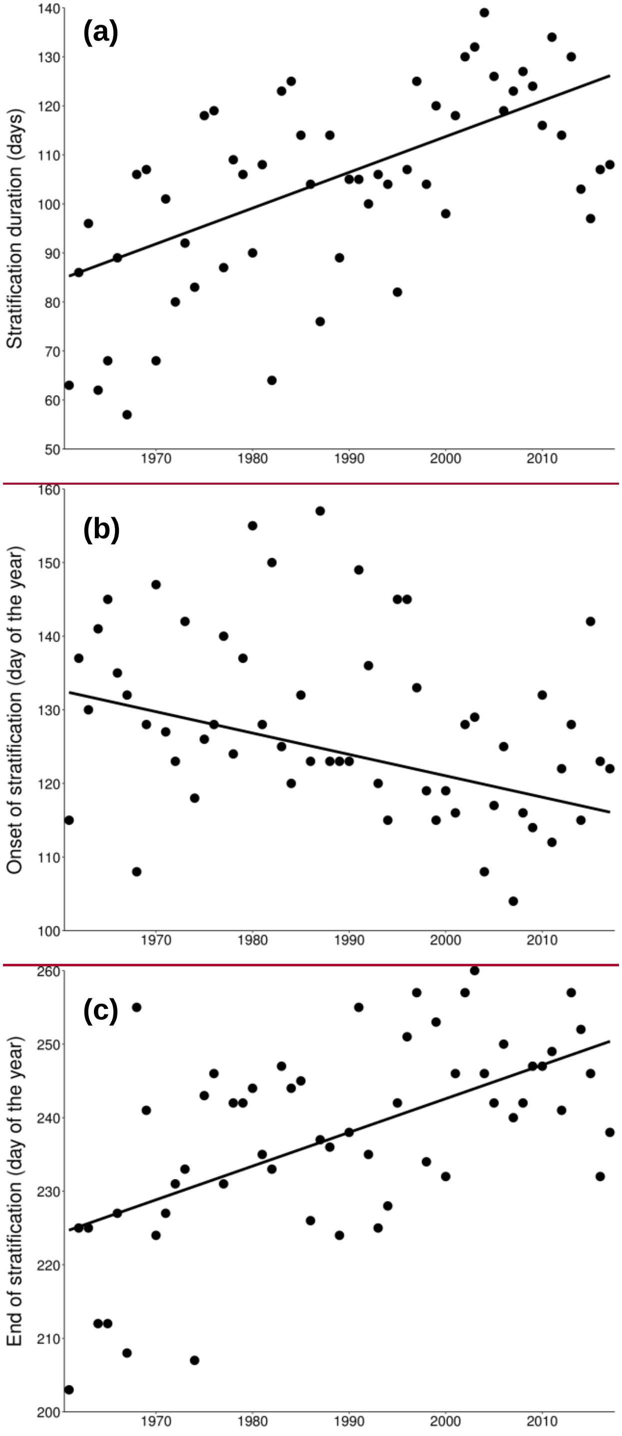
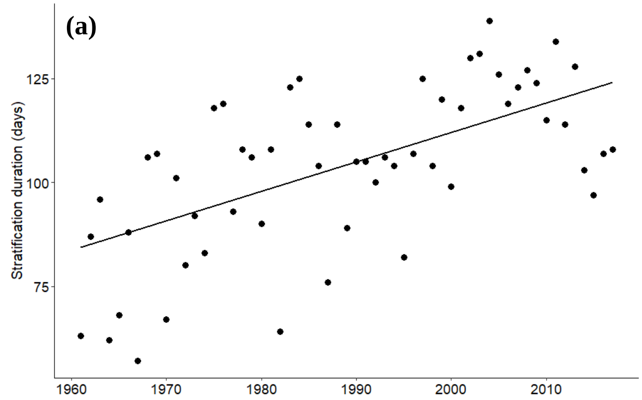
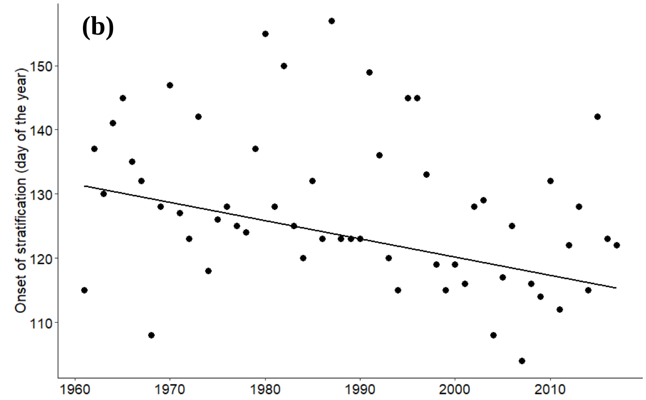
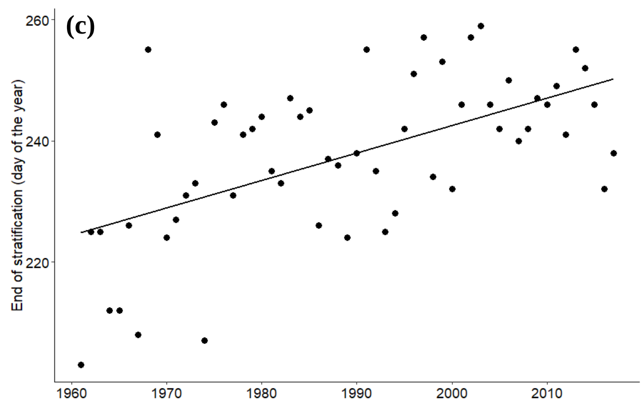


Figure 8: Reconstructed stratification duration (a), onset (b) and termination (c) of Lake Erken between 1961-2017. Stratification duration trend: 7.297 days/decade ($p < 0.001$). Onset of stratification trend: -2.903 days/decade ($p < 0.01$). End of stratification trend: 4.583 days/decade ($p < 0.001$).







~~**Figure 8: Reconstructed stratification duration (a), onset (b) and termination (c) of Lake Erken between 1961-2017. Stratification duration trend: +7.083 days/decade ($p < 0.001$). Onset of stratification trend: -2.841 days/decade ($p < 0.01$). End of stratification trend: +4.545 days/decade ($p < 0.001$).**~~

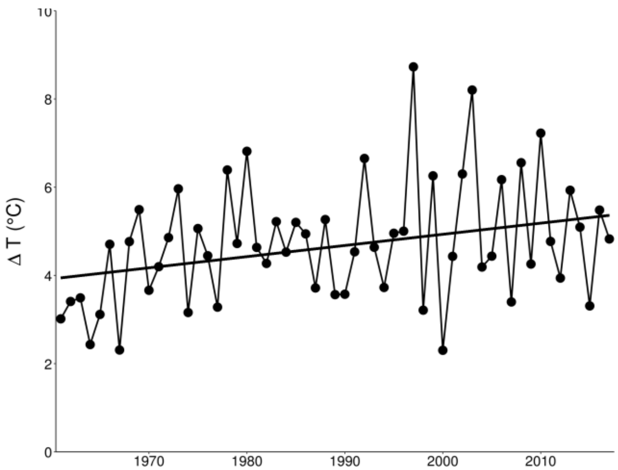


Figure 9: Surface-Bottom modeled temperature difference (ΔT) during stratification between 1961-2017 (Sen's slope: 0.253 ºC/decade, $p < 0.05$).

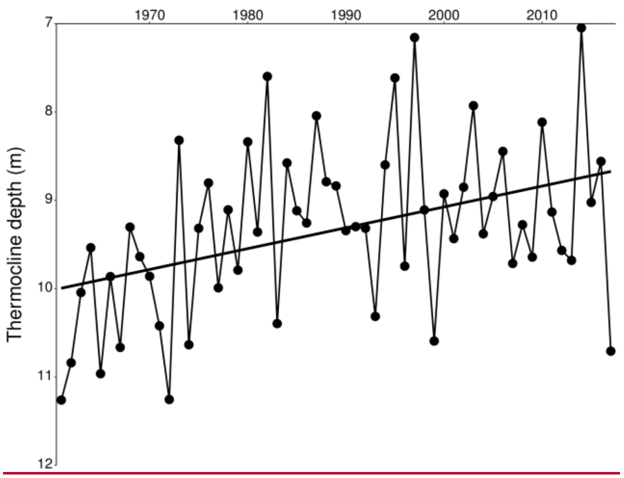
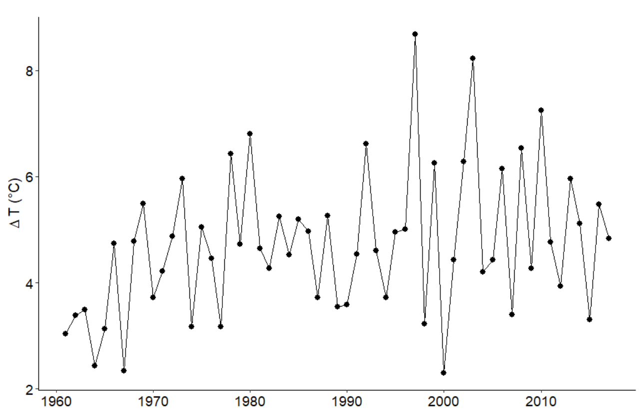
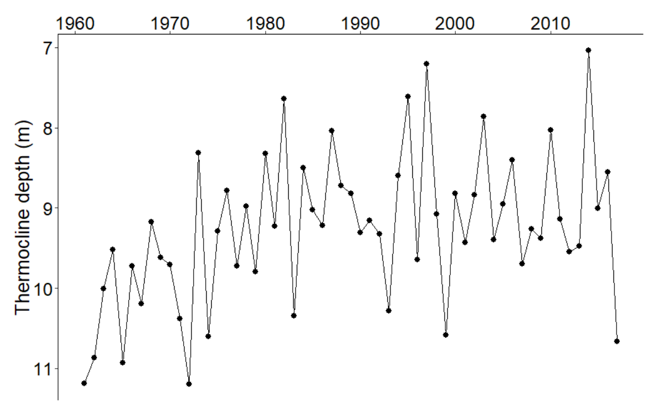


Figure 10: Mean annual thermocline depth between 1961-2017 (Sens's slope: -0.236 m/decade, p < 0.01).



~~**Figure 9: Surface-Bottom modeled temperature difference (ΔT) during stratification between 1961-2017. ΔT increasing rate +0.249 °C/ decade (p < 0.05).**~~



~~**Figure 10: Mean annual thermocline depth between 1961-2017. Average thermocline depth decrease since 1961: 1.6 m (p < 0.001).**~~

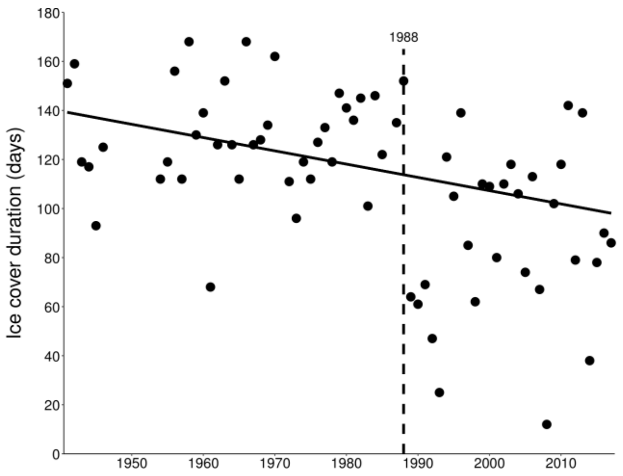


Figure 11: Observed ice cover duration of Lake Erken between 1941-2017 (Sen's slope: -7.343 days/decade, $p < 0.001$). The dashed line shows the year (1988) of abrupt change in ice cover duration (Pettitt test, $p < 0.001$).