

#	Page	Line	Comment
Major1	2	5-8	<p>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.</p> <p>Indeed it would be extremely bad practice to parameterise a model without calibration-validation. Please make your intentions and motivations for the study here much clearer.</p> <p>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 detailing the quantity and quality of the data would be extremely helpful.</p>
Major2	4	1-22	<p>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.</p> <p>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.</p>
Major3	5	13-15; 21-24	<p>standard is to employ a split-sampling approach. Further, the aim is to minimise the variance in the GOF statistics across the calibration-validation period. Without defined periods, you cannot determine the consistency of the model performance.</p> <p>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.</p> <p>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.</p> <p>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</p>
Major4	5	26-28	<p>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 x-axis to the start and end-year.</p>
Major5	9	Figure 5	<p>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.</p> <p>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</p>
Minor 9	2	9	Define what is meant by a long record, for hydrological modelling of rivers this would mean > 30
Minor1	1	6	Need to explain what the abbreviation is - for example, consider: "General Ocean Turbulence model (GOTM), a hydrodynamic model configured in lake mode".
Minor10	2	10-11	Please provide a citation if making a claim such as this.
Minor11	2	17	What does significant mean? How much? Can you give a percentage or some other kind of
Minor12	2-3	29-30; 1;2	<p>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?</p> <p>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.</p>

Minor13	3	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.
Minor14	3	9-12	Repetitive - could simplify to say: "The model utilises six of these climatic parameters (excluding DP)
Minor15	3	4-13	More information is required for GOTM. Why choose this model specifically? i.e. 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?
Minor16	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
Minor17	3-4	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
Minor18	3-4	15-17; 1-3	Clearer signposting is required, please refer to the letters that each station represents in the main
Minor19	3	15	Primarily retrieved from? What does primarily mean specifically?
Minor2	1	9	Real is not very clear - consider replacing with "observed" (or similar).
Minor20	3	16	Is the Malma weather station the Erken laboratory meteorological station? This inconsistency is
Minor21	4	10; 21	What is meant by best? Please clarify how this is judged.
Minor22	4	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.
Minor23	4	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.
Minor24	5	1-2	As with the hydrodynamic model, the reader needs to know why this approach is used. What is the
Minor25	5	5-7	These lines are unclear, please consider rewording.
Minor26	5	8-13	The authors state that an algorithm is used in the parameterisation. What is the stopping criteria?
Minor27	5-6	28; 1-3	Please explain to the reader why they should care about these metrics - why are they important?
Minor28	6	12	Did you test for autocorrelation? Was it all autocorrelated? Please be clearer.
Minor29	6	13-14	Please correct Figure 3 accordingly - the time-series should not extend beyond the point for which it
Minor3	1	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.
Minor30	6	21	Please consider moving this line to the start of the section. Please also include the package version
Minor31	9	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?
Minor32	9	2	In the discussion, please explain to the reader why this matters, what it indicates etc. - It is not
Minor33	9	5	Please define your terms, e.g. epilimnetic.
Minor34	9	9-14	Please be consistent in the number of significant figures for temperature.
Minor35	9	14	Please start a new paragraph before discussing thermal stratification.
Minor36	9-10		As a decadal mean, it would be useful to see the reporting of confidence intervals for these values.
Minor37	11	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.
Minor38	11	12	I do not agree that it indicates the reliability, the wording is too strong. It could be described as a
Minor39	12	1-11	Much of this text appears to be results.
Minor4	1	10/11/15/16	State the months associated with your seasons.
Minor40	12	14-17	The provision of a confidence interval would help to expand upon this further (it could also improve
Minor41	12	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.
Minor42	13	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 is relevant - the implications are not made clear.
Minor43	14	8-12	The assertion of "accurately" cannot be made whilst there is no robust consideration of model
Minor44	14	13-19	Suggest that a dedicated conclusion would help to wrap up the paper and reassert the
Minor5	1	23-25	Abstract does not necessarily make clear why this matters - what is the need for the work?
Minor6	1	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.

Minor7	2	1-2	Again, repetition - it is self-evident that a rise in lake water tempeature increases water
Minor8	2	5-7	It would be helpful to explain what some of the conclusions of these studies are/were - it makes it
NA	6	4-7	I would like to highlight that the level of description here is excellent and represents the level that