Response to Reviewer 1

- This manuscript by Magee & Wu is based on an extraordinary data set of 104 years and focuses on effects of changing air temperature and wind speed on water temperature and stratification patterns of lakes with differing morphometry. The lakes are situated close to each other which is a great asset in this kind of research. The long data sets on basic variables and drivers is a good argument for publication and the results based on these data are fairly convincing. They are also logical and actually so logical that they very often leave a feeling that 'I already know this'. This may at least partly be due to simplification of morphometry to lake depth and surface area, but also due to lack of deep discussion; big part of 'Discussion' actually belong to 'Results' and to certain extent to 'Methods'. Thus restructuring and extending the real discussion (starting from 4.4.), the paper would certainly improve.

  The authors thank Reviewer 1 for taking the time to review and providing helpful comments to improve the manuscript. Following the suggestion, the authors are restructuring the paper and extending the discussion section. We have addressed additional comments in a point-by-point reply and carefully address the issues raised in the revised manuscript.

- Some parts of the paper are also technically challenging for the reader since they are based on listing the numerical results one by one; a good example is section 3.5.

  The authors thank the reviewer for this comment. We will address the structure of the writing that become technically challenging for readers. We have reduced the presentation of trends and related thermal characteristics within the text. We avoid listing the numerical results one by one without any meanings and incorporate these numerical results into the figures and tables.

- The authors should also think about leaving Lake Wingra out completely; I suggest this because this paper has strong focus on lake stratification and Lake Wingra is a polymictic lake. The problem with Lake Wingra becomes obvious in Tables 2, 3 and 4 – lots of N/A markings.

  The authors thank the reviewer for this comment. Lake Wingra does stratify on daily or weekly timescales during the summer months (Kimura et al, 2016). Summer Schmidt stability was calculated at daily timescales, and then averaged for each year before comparing coherence among the lake pairs. Higher average stability for one year on Lake Wingra would indicate that the lake experienced more days of stratification during the period. This phenomenon can be coherent with changes in stability for the other two lakes.

  Reference:

  Kimura, N., Wu, C.H., Hoopes, J.A., and Tai, A. 2016, Diurnal thermal dynamic processes in a small and shallow lake under non-uniform wind and weak stratification, Journal of Hydraulic Engineering-ASCE, 142(11), 04016047,

- I also find it strange that in a paper where models are such an elemental part, they are not properly described; besides the equation for light extinction (eq 1), the authors only use references to published articles.

  The authors thank the Reviewer 1 for this comment. Indeed, the description of the model have been described in great detail in the papers (Magee and Wu, 2016 in *Hydrological Processes* doi/10.1002/hyp.10996/full, Magee et al, 2016 in *Hydrology and Earth System Sciences*, DOI:10.5194, 20(5), 1681-1702. As a result, we did not intend to repeat the information and included the detailed description in the manuscript. Based upon the concerns raised by the reviewer, we will add proper description of the model as necessary. The equation for light extinction is included in the manuscript as it is a new updated component. We will provide additional details on the model subroutines that directly affect horizontal processes in the lake. We will also detail parameterizations and describe how they influence the results of this analysis.

- More emphasis should also be given to description of gap filling and calibration data; both are now somewhat superficial.

  More detailed description of gap filling and calibration data can be found in other manuscripts (Magee et al, 2016 in *Hydrology and Earth System Sciences*, Magee and Wu, *Hydrological Processes* doi/10.1002/hyp.10996/full in press). We agree that this information is pertinent to the discussion and results. The authors will provide more details of these analysis in a supplement to the paper, which would address the concern raised by the reviewer.

Besides these more general comments, I list here some more detailed ones:

1. I found it a little bit strange that sediment heat fluxes were hardly mentioned in this paper. Although there may have been no data on this or these fluxes were not included in models, they should have been tackled somehow at least in 'Discussion'.

   Sediment heat flux is included in the model. Details of this information can be found in other published papers (Magee et al, 2016, Magee and Wu, 2016). We agree that sediment heat fluxes are an important component to the overall heat budget of the lake. We will discuss sediment heat fluxes in the revised manuscript. Additionally, we will include more detail on the parameterization and inclusion of sediment heat flux in the revised manuscript.

2. The readers would benefit from some more information about the lakes. Especially information on lake clarity (water colour etc; cf. Table 1) would have been useful in a paper with such a strong focus on lake stratification.

The authors thank the reviewer for this comment. Additional information on the lakes has been added to Table 1, including fetch, shoreline development, landscape position, Secchi depth, surface water chlorophyll concentration, and DOC in each lake. Specific values of lake water color is not collected by the NTL-LTER program as other data were.

3. As a reader I would also appreciate information on fetch for each lake; now the word 'fetch' and importance of fetch is mentioned several times, but the reader is left with the bathymetric maps to figure out the fetch

The authors apologize for neglecting to include this value explicitly in the manuscript. Information on lake fetch for each lake has been added to Table 1.

4. It is said that water level in Fish Lake has raised considerably and this has probably affected some of the results. However, nothing is said about the possible reason behind this phenomenon. Related to climate, human activity or what?

The text has been changed to read "the water level of the lake rose by 2.75 meters due to an increase in regional groundwater recharge causing increased groundwater flow to the lake (Krohelski et al., 2002). Krohelski et al. (2002) hypothesized that the increase in regional groundwater recharge may be the result of increased infiltration from snowmelt after increased snowfall and less frost-covered soil."

5. The authors state that Fish Lake does not always turn over completely in spring. This is an important piece of information, since in small, dark coloured boreal lakes this is a fairly common observation and it is believed that it is weather/climate driven change. It would be nice if the authors could dig deeper in this observation, especially since they have such a long time series.

The lake does mix each spring, however, low water temperatures of only ~5°C in the hypolimnion may be due to shortened spring mixing durations compared to years where the hypolimnion temperature reaches ~8-11°C. The authors hypothesize that this phenomenon may be related to late ice-out in some years. We will investigate the phenomenon further in the revised version of the manuscript. Ice cover does impact the timing of stratification and hypolimnion temperatures. We propose to add a section in the discussion that describes this interaction and its influence on the results presented here while not reproducing content already presented in Magee and Wu (2016). This section would address our hypothesis that colder hypolimnion temperature is related to years with later ice-off.

6. Fish Lake and Lake Wingra have Secchi-depth results only from 1995 onwards. This appears problematic; could you give more explanation on this.

   Fish Lake and Lake Wingra became part of the NTL-LTER program in 1995; and regular Secchi depth measurements were taken starting then. The authors agree with Reviewer 1 that it is not ideal to use seasonal averages for the historical period before 1995; however, given the strongly seasonal dynamic of water clarity and light extinction in the lakes, using seasonal averages of Secchi depth to estimate light extinction are preferably to a constant light extinction for the lakes, which is not representative of observed phenomenon in the lakes. The authors will perform additional model analysis to quantify the uncertainty that may be caused by using seasonally-average Secchi depth instead of the measured values.

7. Data on below-ice Secchi-depth were used which I to certain extent understand, but since it is not that common practice to measure Secchi under the ice, it would be useful to have some more information.

   Light extinction, which can be estimated from Secchi depth, greatly influences water temperatures and overall temperature profile. Including light extinction in winter more reliably reproduces under-ice water temperatures and as a result, water temperatures at the time of ice-off. Temperature profiles at ice-off impact the timing of stratification and the hypolimnetic water temperature through the summer. Properly characterising and capturing these phenomena in the model will enable accurate reproduction of water temperatures during the historical period. For this study, the authors choose to utilize the available data from previous ecological and water quality studies conducted on the lakes to better inform the model and more accurately reproduce water temperature profiles.

8. Figure 3 shows that in general simulations resulted in slightly lower temperatures in comparison to observations. Did you make this clear also in text?

   We will improve the text to make this clearer and expand on why this bias exists.

9. The possible importance of internal waves is mentioned only on general level and not properly discussed in relation to the study lakes

   Internal waves are parameterized in the model, however, the reviewer is correct that we did not explicitly explain how the model deals with internal waves nor how internal waves affect hydrodynamics in each of the lakes. Based on other research, the effect of internal waves may be large in Lake Mendota and much smaller in Fish Lake. We will clearly expand upon this issue and discuss the validity of modeling results in relation to the parameterization of internal waves in the model.

10. Using wording ‘increasing (decreasing)’ is clumsy for the reader

The authors thank the reviewer for this comment. We have revised the text to make the writing clearer for readers.

11. Throughout the text there is repetition, e.g. in ‘Results’ sentences which belong to ‘Material’ and are already tackled there. Check the whole manuscript for that

The authors have removed and moved sections of the manuscript to address this issue.

12. Table 1: The meaning of the row ‘Groundwater’ is not clear to me

The row ‘groundwater’ describes the groundwater inflow type of the lakes. For example, “discharge” lakes are those which have a net groundwater discharge into the lake. “Flowthrough” lakes are those which have small net inflow or outflow. As for Lake Wingra have high groundwater inflows and outflows which result in small net discharge into the lake.

13. In Figure 4, the legend contains some description of results

The authors thank the reviewer for pointing out this comment. We have removed the description of the results out of the legend.

14. In Figure 6, results on Lake Wingra should be left out (= zero line). And in general, the stability index is somehow funny in this context since the lake was known to be polymictic

As suggested by Reviewer 1, we will leave Lake Wingra out of this figure. To address stability, please see our previous response concerning the inclusion of Lake Wingra.

15. The real discussion starts in 4.4. and all before that should be merged with ‘Results’. An indication of that is the fact that for instance in 4.3.1 and 4.3.2 there are no references in the text.

The authors thank the Reviewer 1 for the comment and suggestion. We will restructure the discussion by moving Sections 4.2 and 4.3 to the Results section and emphasize the results in terms of physical mechanisms that are influencing the simulated and observed responses.

16. There are some spelling mistakes in the text, please check.

The typos have been fixed within the manuscript. Thank you to the reviewer for pointing out these mistakes to the authors. We will additionally review the manuscript carefully for typographical errors before resubmitting a revised manuscript.