Referee's Report on Revised Submission of Determining lake surface water temperatures (LSWTs) worldwide using a 1-dimensional lake model (FLake, v1)" by A. Layden et al.

This paper is improved over the previous submission, particularly in terms of proper usage of units. None of my comments is really major, but the sum of the ones that are substantive leads me to recommend major revisions again.

Its long length is largely necessary for giving the type of readers who will benefit from this paper the information that they need. However, even though I am not providing specific instances, I ask the authors to comb through the manuscript again to find redundancies that can be removed.

A frequent issue in modeling is the idea of what happens in the virtual world of the model, and how does that compare to the real world, not just in terms of output, but in terms of process. I am not intimately familiar with the FLake model, but in looking at your tuning in terms of adjusting the lake depth in the model, I have to think about two limiting cases in the real world: 1. A very shallow lake, in which the entire depth constitutes a well-mixed epilimnion nearly all the time. 2. A very deep lake, in which the metalimnion is so separated from the lake bottom that its depth, strength, and heat flux are independent of the overall lake depth. I guess what I am most curious about is how FLake formulates metalimnion/thermocline depth, and how this meshes with these two limiting cases, as well as intermediate cases. For instance, is it so insistent on having essentially two layers that it will not collapse them into one for the shallow limiting case? And will it allow the lower layer to be much thicker than the upper layer, to capture the deep lake limiting case?

I am referring to line numbers as given in the double-spaced PDF version that I reviewed.

P. 3, line 7: Use a semi-colon rather than a comma after "feedbacks".

P. 3, lines 11-12: The cause and effect is definitely more complex than stated here, and, I think largely opposite, with heat and moisture fluxes from lakes causing lake-effect storms.

P. 3, line 26: "Compare" seems to make more sense here than "use both".

The term "daily mean absolute difference" is throwing me. Does this mean that you have multiple data within each day and calculate the mean across each day? Or does it mean that you start with daily data and average the absolute difference over, say, a month? To make matters worse, p. 4, line 6 has "average" in front of this term and line 8 has "mean" in front of it. Please clarify what data is the starting point and what period it is being averaged over.

P. 4, line 20: You spelled out the names of these model parameters just three lines earlier, so no problem to abbreviate them now.

P. 6, lines 28-29: "An average…" is singular, while "…are used" is plural. It's not really clear to me which was intended but make them agree.

P. 6, line 16: This is where a problem with units is still present. The formula for lakes with length and breadth dimensions seems quite intuitive and is roughly proportional to the square root of area, but this formula has a linear term based on area and is tougher to fathom. It seems like you're supposed to assume that area is in square km, but this isn't explicitly stated. If it really were, though, you'd need to then divide by km to get the final unit of km.

This also got me to wondering whether there is any special class considered for lakes of highly irregular shapes. One that I think of is Lake of the Woods, which has one large basin, but is also connected to many narrow necks and small basins separated by trees clinging to rocks on a shield terrain. How does this affect fetch independently of area?

P. 11, line 9: I think you mean "spectrum" rather than "spectre".

P. 11, line 10: I recommend adding "bands represented by" or "bands centered on" before "three wavelengths".

P. 12, line 4: Is there a rule or formula for what makes ice blue or white in the model?

Sub-sect 2.3.5: Are we always assuming that we're talking about wind speeds at 10 m height?

P. 13, lines 28-29: This refers back to my comments about the term "daily mean absolute difference". The statement here seems tautological in that it is trying to define "MAD" mainly by spelling it out, with "difference between the modeled and observed LSWTs" barely needing to be said. P. 14, lines 11-12 is a nearly exact repeat of this sentence.

P. 14, line 19: Should this be the variance of LSWT, rather than the mean?

P. 14, line 29: "x" seems to be missing, leaving just the superscript "obs_jas", and this should have the description "for each individual year" at the end.

The equations in sub-section 2.4.3 can all benefit from citations, but especially the most complex one on p. 15, lines 11-13.

P. 16, lines 12-13: The equation needs units, or it might be easier to refer back to eq. 4, which has units in it. Same on p. 31, lines 11-12.

P. 17, line 9: Change "from" to "to".

P. 21, line 8: Change "Hemispheric" to "Hemisphere".

P. 21, line 28: I think "simulated" captures this meaning better than "represented".

P. 22, line 23: Fig. 16 seems to be referenced before Figs. 14 and 15.

Sub-section 5.1: Be careful about cause and effect. When adjusting parameters, the root cause of any change is always the adjusted parameter. So changes in 1 deg C warming date may be a link in the chain, but aren't the ultimate cause. The schematic in Fig. 16 is a key to this, but you need to be careful about confusing causes in the real world with causes in the world of tuning model parameters. This enters in more in the mention of Austin and Colman. The earlier warming may be one aspect of what is going on there, but Sun et al., 2015, J. Climate, 4373-4389; and Foster and Heidinger, 2014, J. Climate, 6687-6697 seem to show that the root cause is more likely cloud albedo.

Sub-section 5.2: Again, I am not intimately familiar with FLake's formulation, but it seems very likely that it would be formulated so that it is inevitable that hypolimnion temperature is very closely tied to the annual minimum surface temperature, due to static stability constraints.

P. 26, line 31: Change "aren't" to "is not". Same on p. 30, line 28.

P. 33, line 29: The usual transliteration of Kirillin's first name is "Georgiy".

Fig. 4: Again, not an expert on FLake's formulation, but my understanding is that it uses an s-curve shape to represent the temperature profile. This makes it somewhat equivalent to a 2-layer model, but it does not literally consist of two discrete layers, each with a constant temperature. If I've misunderstood that, go ahead and correct that. However, this figure confuses me even more about how this works. The left side seems to depict it using the language of two layers, with the "upper mixed layer" depicted as a range of depths, while the "bottom layer" has the arrow pointing only at one depth. Then the one on the right has the hypolimnion labeled. Is this exactly the same as the bottom layer, even though it seems to be depicted with non-zero thickness? The epilimnion is not labeled, but seems to correspond to the upper mixed layer. How correct is this? Then the thermocline (largely synonymous with "metalimnion") is shown as having a finite thickness, but how does it fit into this idea of a two-layer model, and how is its depth and thickness determined?