This paper presents result of model projections of the temperature and mixing under past and future climates (scenarios ssp125, 370 and 585) for four peri-alpine lakes, using the IPSL-CM6-LR atmospheric model from ISIMIP3b with the 1D lake model MyLake. It also considers future changes in thermal habitat and oxygen solubility of these lakes, making inferences about their suitability as a habitat for some fish species. Good projections, eg as used with the new ISIMIP3b data, of the potential effects of warming on these important lakes will be very valuable and an important and novel research result.

My main issue with the paper concerns the style. I found it quite confusing and the rationale behind the study in the introduction as well as the methods should be more clearly stated. For instance, I did not properly understand why the authors chose to use only shortwave radiation and air temperature projections for model hindcasts and future projections, and use historical means for the other model forcing variables. Also it was not so easy to quickly understand how this was done. I think you show show the comparison between the model configurations with just shortwave and aitemp, and with all forcings from the GCM more clearly and make a short justification for your chosen configuration. I also didn’t understand why the authors began with an ensemble of climate and lake models, but selected only one of each for the analysis, or what the value is of performing a long hindcast. This all might sound provocative, and I am not saying this is unjustified at all, only that it was not explained sufficiently to the reader. The scientific results seem sound, but I struggled to interpret the plots (which have a resolution and font size that are too low for me to read the axes), and I think the discussion needs to be reworked, focusing on the interpretation rather than the description of the results. There are some interpretations to explain the differences between the lakes, but I think they need to go deeper, like if simply wind fetch is really causing deeper mixing in Geneva, or the effect of earlier stratification onset timing for deep water warming trends. Several formulations in the manuscript describe complex things, but need to be sharpened to make this complexity more understandable. In the sections below I have tried to point out clearly the parts that I am referring to.

Abstract
General: in this modelling context, try to use “parameters” only to refer to model parameters, and use “variables” for things like wind (eg line 62) or e.g. “substances” when talking about concentrations (eg line 20).

Introduction

You go into some detail to describe the approach, but not quite enough to make it clear, mainly in lines 75-86. Later in the methods you go into a little more detail, but I found it unclear when starting with the introduction whether you use observed radiation and airtemp to force the model, and then estimated the other variables? It would be helpful to describe the rationale behind this. The introduction should be focused more to justify and define the main research question. For instance, there is no mention of the thermal habitat and I don’t recall that you reported or discussed any results based on the 4 different model configurations.

Line 129-130: did you use the historical scenario for each model (I assume)? Please state explicitly. The historical scenario from the IPSL model is available back to 1850 for all required forcing variables for the lake model. As mentioned above, please explain why you use only shortwave and airtemp.

Line 132-133: sentence not clear (350 data)

Line 148-150: It is unclear what is meant by the sensitivity test – could you please briefly elucidate, even if the description is in the supplement? This hypothesis needs to be more clearly stated in the introduction

Line 151-160: it would be simpler if you described the configurations in order i), ii), iii), iv), rather than beginning with ii and iv.

Table 2: I don’t completely understand the table from the legend. Do the performance metrics represent the lake model temperature errors (units degrees Celsius)? The configurations are not really understandable in this context – when you say only air temperature and shortwave, do you mean that these variables come from the climate model IPSL-CM6-LR, while the other input variables are observed data?

Line 181-184: I didn’t understand this bit – you should add more specific details (eg what is the instrumental period and why do you need to identify similar climate conditions, what do you mean by the RMSE between winter air temperatures, etc?)
Line 201-202: Be more precise about the formulation and units – Ts at 5 m (not Ts=5m), and add units for Tb (Tb at 60 m, 60m, 299 m, and 140 m depth for ...)

How did you define stratification duration – is it the total number of stratified days per year (where each period > 5 days), or the duration of the longest uninterrupted period?

Results

Note: the resolution of the figures is very low and the labels are very small, so that I am unable to read the axis labels on any of them except Fig 1. Please bear this in mind if there are potential interpretation differences.

Line 277-281: rather report change/decade only for trends over time. If you make a comparison of period means (eg mean 1990-2020 with mean 2070-2100), then report the change as absolute, eg an increase of 1.5 degrees Celsius from the baseline period to the end of the century.

Figure 4: it looks like you applied some smoothing to the time series – please describe what you did. Also, I suggest to label the panels only with letters, not combined letters and numbers

Line 297: I am curious about the negative temperature trends you report here, and also interpret in the discussion. I cannot see any trend in Fig 5 – the epilimnion temperature always looks to be increasing in ssp126.

Line 316-318 (and corresponding methods section): the non-overlap in the temperature density distributions is not sufficiently described to understand what it represents. Which temperatures are being compared exactly? They are obviously averaged, but how (eg across the water column, but then they may need to be volume weighted)?

Line 339-351: I think it would make more sense to analyse the trend over the whole period, or if you compare the baseline 1990-2020 with the future 2070-2100, then it would be more informative to report the absolute change rather than the trends today and in the future. This would put into perspective the fact that ssp126 initially warms but then cools slightly towards 2100 reflecting emissions reductions. So although there is a cooling trend at the end of the century in this scenario, this should not be confused with the overall long-term warming trend compared to today. Also, in Table S4, MyLake has an R2 of practically zero for stratification phenology in many instances – if you trust there
results, you should mention the reliability in the discussion. Also, I wonder if these results would be better in a table, which makes comparison and overview of the changes much easier.

Figure 7: I suggest to reverse the colours – blue for the coolest scenario and red for the warmest.

Line 368: do you mean that temperature exceeded 7 degrees in all scenarios in the future period? Are these daily temperatures or a mean (Methods say averaged daily data, but this is not quite clear)? The background to my question is that it is quite different if the lake never cools below 7 degrees, even in winter, or if there are only a few days in summer when this threshold is exceeded.

Line 371: Better say 100% non-overlap rather than 100% increase.

Lines 440 ff: Section on oxygen solubility density overlap: I feel that it does not make sense to report the non-overlap of the density distribution of oxygen solubility. This is hard for me to meaningfully interpret, and I suggest to remove this part.

Discussion

The discussion section contains a large repetition of the results, and there should be more integration of the findings into the context of the literature, with the aim of detecting some of the important mechanisms driving the thermal changes. You may like to consult Kraemer et al (2021) and literature therein to interpret lake thermal habitat (Nature Climate Change 11, 521-529). Plus there is a lot of literature on the thermal properties of these peri-alpine lakes eg Geneva that may help identify the mechanisms why these lakes behave differently. Altogether, I think the discussion and conclusion should be heavily revised and especially remove the results descriptions.

Line 469: check year

Line 473-475: you can remove this repetition of results and just concentrate on the interpretation. If you want to mention this, it would be simpler to say that surface water is warming faster than deep water at 0.07 degrees/decade in Lake Bourget and at 0.15-0.16 degrees/decade in the other lakes.

Line 481-485: as I mentioned above, you should be careful with reporting ssp126 in this
way. Although ssp126 shows cooling air temperatures at the end of the century (negative trends due to emissions reductions), it still represents a warmer future, so make sure your formulations are not misleading. I find the formulation as it is now sounds like the surface is cooling compared to today. Also, I could not see the negative trends you are talking about in Fig 5, perhaps you could comment on this?

Lines 489-497: this is all just a repetition of the results and is unnecessary.

Line 511 ff: You shouldn’t call this section Oxygen Trends because you didn’t model oxygen, just temperature, on which the oxygen solubility is only calculated. This section is essentially a repetition of the results. You should focus on interpreting the results, explaining their significance, drivers, and context.

Lines 542-551: You need to account for the differences in lake morphology when comparing Schmidt stability. Lake Geneva will need much more energy to mix it than Lake Annecy. Perhaps comparing relative changes would be more meaningful than absolute changes.