The manuscript investigates historical and future impacts of climate change on peri-alpine lakes by applying vertical 1D hydrodynamic lake models to the lakes Geneva, Annecy, Bourget and Aiguebelette. Here, the focus is on quantifying changes in the hydrothermal environment, i.e., water temperature, water column stability, stratification timing and thermocline depth, which the authors also use to link to changes regarding oxygen concentrations by calculating oxygen solubility as a function of water temperature. A novelty of this study is the experimental approach to only force the 1D lake model by air temperature and short-wave radiation data while relying on historical dynamics of other variables like wind speed and cloud cover. The study is of interest to limnologists and modelers alike, but the visualization of results needs to be revised (i.e., low resolution of current figures) as well as methodological aspects of the study.

Major points:

- The abstract (L25: "[...] several 1D lake model’s [...]”) as well as the introduction and methods part make the impression that the study will be mainly about running an ensemble of lake models on peri-alpine lakes, but in the manuscript only MyLake results are shown and interpreted (all other ensemble model output are in the Supplementary). Here, I think several points are missing: (1) an explicit statement explaining why MyLake was used instead of the other ensemble members, (2) a revision of the text to make it clearer that the study is highlighting MyLake, and (3) a discussion paragraph looking at other ensemble model members. To (1), table S4 is showing that MyLake has the best model fits, and it makes sense to use it. I think this should be stated instead of L115 "[...] as this model is well adapted to Northern and alpine regions” as this is also true for Simstrat which was explicitly developed and applied on Alpine Lake systems. Further, Simstrat also incorporates TKE production by seiche which was mentioned as a shortcoming of the current study in replicating thermocline depth dynamics (L612). Further, the Supplementary is missing information on which model parameters were calibrated for FLake, GLM, GOTM and Simstrat. Also,
as the title suggest the focus on oxygen dynamics, why wasn’t for example GLM chosen to simulate these dynamics directly instead of MyLake which in the LakeEnsemblR version does not quantify dissolved oxygen dynamics at all?

- Resolution of figures: the resolution of most figures (2, 3, 4, 5, 6, 7, 8, 9, 10, 11) is very low and their axes labels are not readable. Please improve the resolution.
- L479: This sentence seems wrong to me. First, wouldn’t increasing density gradients especially in deep lakes reduce vertical mixing and prevent hypolimnetic water temperatures from increasing. Further, Lake Geneva’s surface area is big (which I suppose you mean here by wind exposure) but the wind energy itself as well as the reach of TKE is supposed to be decreasing in the future (see Castro et al 2021, Nature Comm Earth&Env). Second, how should enhanced deep winter mixing (assuming you refer to convective overturn) lead to plunging of cold water from melting glaciers in deep layers? Is this referring to the Rhone River receiving increased meltoffs? But in that regard, increased mixing would lead to increased deep water temperatures and that would cause a colder Rhone inflow to entrain in to deeper water layers.
- L502: The statement that the simulated warming is different from current observations due to high wind exposure is not enough in my opinion. Couldn’t this be a model shortcoming and bias. Are you sure only the wind mixing pathway is the main difference here? Please focus more on discussing why the future projections are different from recent observations.

Minor points:

- L80: It’s unclear how oxygen solubility is a biological indicator. I agree that dissolved oxygen itself is one, but solubility (also as represented here) is merely a function of temperature.
- L85: I’d change “dissolved oxygen” to “dissolved oxygen solubility”
- L97: You’re describing all lakes as monomictic but I recall that Lake Geneva did not experience a full turnover for over a decade now. Could you provide references for that statement please?
- L138-142: In most parts of the manuscript, you refer to the scenarios as ssp126, ssp370 and ssp585 only, but here you also refer to them as SSP1, SSP3 and SSP5. Please stick to one naming convention.
- L148: Which sensitivity test was carried out and how?
- L152: The mentioning of (ii) and (iv) without (i) and (iii) is slightly confusing. Would it be possible to name these scenarios differently, e.g., (i) airT and SW no correction, (ii) all no correction, (iii) airt and SW with correction, (iv) all with correction?
- L155: Wouldn’t especially Lake Geneva have more meteorological data since the launch of LeXPLORE?
- Table 2: Are all fits calculated for all observational data of temperature?
- L183: I agree with your methodology, and I love how you also compare the fits to long-term time periods, but nonetheless it seems strange that you first mention that L70 “Models are in large extent calibrated against very few years of limnological records so far” but then you are also focusing the calibration/validation on ten years each only. What is the reason for this?
- L207: “[…] epilimnion extent and temperature, hypolimnion extent and temperature, metalimnion upper and lower depths, […]”
L221: “When normal distribution of residuals […]”

L223: What were the conditions for either choosing Mann-Whitney or Kolmogorov-Smirnov?

L252: The thermocline depth fits are quite bad (would they be better if you would only quantify them during stratified summer periods?), but what is the reasoning to quantify them here as percentage of lakes total depth? Although these values sound low, I don’t see any reasoning for expressing them this way as an error of 20 m in thermocline depth is still significant and it does not matter if the lake is 300 m deep as the thermocline depth will be in the upper 50 m part anyway.

L253-260: Why are these fits (DOYs) not included in Table 7?

L345: I recommend deleting “delay” here

3.3 Water volumes: habitat: This paragraph is unclear to me. Where these temperature volumes chosen as average temperature output per year, or are these daily data? Also the comparison of non-overlapping distribution areas is quite confusing between two different decades. Could you add more information please?

L409: Please keep an eye out for the different digits (. or ,) as here you wrote 12,21 (also the same in the Supplementary tables, please change to dots)

L442: Please add references to the ecological thresholds of 10 and 11 mg/L

L474: The formulation of delta raise per decade of epi-hypo is very confusing. Could you please state this differently (like just write delta T decade-1)?

L489: Could you please discuss why hypolimnetic temperatures are increasing although water column stability is also increasing? How do these systems still have enough energy for vertical mixing? The same also for ssp126 results in L504: why has that scenario such different results here for epilimnetic and hypolimnetic temperatures?

L519: I’d argue that lateral flow paths and production in littoral zones would also be important for deep and large lake ecosystems.

L547: I’d argue that Lake Geneva was only the most extensive due to its high depth. Could you check Schmidt Stability trends for the first 50 m of each lake and see if Geneva is still an outlier?

L591: Please add information to which approach you are referring here, I assume it’s the focus on air temperature and short-wave radiation, right?

L645: The sentence about “Differences in the oxygen solubility response to climate change are also observed” seems very vague to me.