The manuscript addresses climate-driven changes in thermal stratification in shallow polymictic lakes. However, the central research questions, methodology, and conclusions lack sufficient novelty.

We thank the reviewer for comments and criticisms regarding the manuscript. We respond to each comment and question below. We don’t agree with several comments for which we can and will provide justifications.

We refer to several papers in our response, which are listed at the end of the document.

1) Over the past decade, numerous studies have extensively examined the response of lake thermal structure and stratification dynamics to climate change, including in shallow and polymictic systems. For example, Woolway et al. (2022) clearly demonstrated that climate warming leads to earlier onset, prolonged duration of stratification, and altered mixing processes, with important ecological consequences. More recently, Sullivan et al. (2025) further provided mechanistic insights into how climate change regulates lake thermal regimes and ecosystem functioning.

R Iestyn Woolway, Sapna Sharma, John P Smol, Lakes in Hot Water: The Impacts of a Changing Climate on Aquatic Ecosystems, BioScience, Volume 72, Issue 11, November 2022, Pages 1050–1061, https://doi.org/10.1093/biosci/biac052

Sullivan, C.J., Read, J.S. and Hansen, G.J.A. (2025), Climate-driven alterations of lake thermal regimes. Limnol Oceanogr, 70: 2348-2364. https://doi.org/10.1002/lno.70128

Woolway and his coworkers, in their several studies, provided global estimates of expected changes in thermal regimes; however, they didn’t have the opportunity to deepen to diurnal scales. Moreover, they focused mainly on deep lakes and global big-picture issues rather than detailed analyses. Their climate simulations relied on datasets from global climate models (GCMs) in many cases, which are far from able to describe not just diurnal processes but also regional meteorological features. The mentioned papers did not focus on shallow lakes at all.

In shallow and polymictic systems specifically, similar questions have already been explored in depth, see (Török et al., 2025; Zhang et al., 2025). The study seems to emphasize sub-daily processes as a key innovation. However, this aspect is also not new. Previous studies have already investigated sub-daily dynamics in shallow lakes. For example, Frassl et al. (2018) examined sub-daily variability in lake temperature dynamics, and Piccioni et al. (2021) demonstrated that thermal stratification and mixing can alternate on hourly timescales.

Török S D, Torma P. Long-term changes in summer stratification of a shallow polymictic lake by climate change and anthropogenic water level regulation[J]. Science of The Total Environment, 2025, 1009: 181025.

Zhang, M., Leppäranta, M., Heikkilä, M., Weckström, K., Korhola, A., Kirchner, N., … Weckström, J. (2025). The thermal structure of small and shallow Arctic Fennoscandian lakes. Arctic, Antarctic, and Alpine Research, 57(1). https://doi.org/10.1080/15230430.2024.2433829

We can partly agree with the above-mentioned comment. Yes, we did examine diurnal processes of stratification, mixed-layer depths, and thermal dynamics in a large shallow lake in Török and Torma (2025), but only for the past, with no projections for the future, and as a function of climate scenarios.

In our opinion, Zhang et al.'s recently published article is not a good example, either. First, they are not analyzing subdaily variations, but rather the annual cycle. Second, the presented lakes represent a very special climate condition. Third, many lakes are polymictic, but only a few undergo overturn and mixing regularly, like daily. In contrast, a vast amount of shallow lakes are continuously polymictic, characterized by diurnal cycles, especially in the midlatitudes.

In contrast, Frassl et al. (2018) and Piccioni et al. (2021) focused on the diurnal behavior of thermal dynamics; however, they did not provide any projections for the future. In addition, Frassl et al. (2018) highlighted a relevant problem. Lake models for shallow environments require meteorological forcing time series with subdaily resolution. They show that not only climate models struggle with this problem, but that reanalysis datasets can also be highly inaccurate for wind due to spatiotemporal downscaling issues. In their paper, they show that model simulations reproduced the lake's seasonal patterns but failed to reproduce the diurnal patterns. We have overcome this problem.

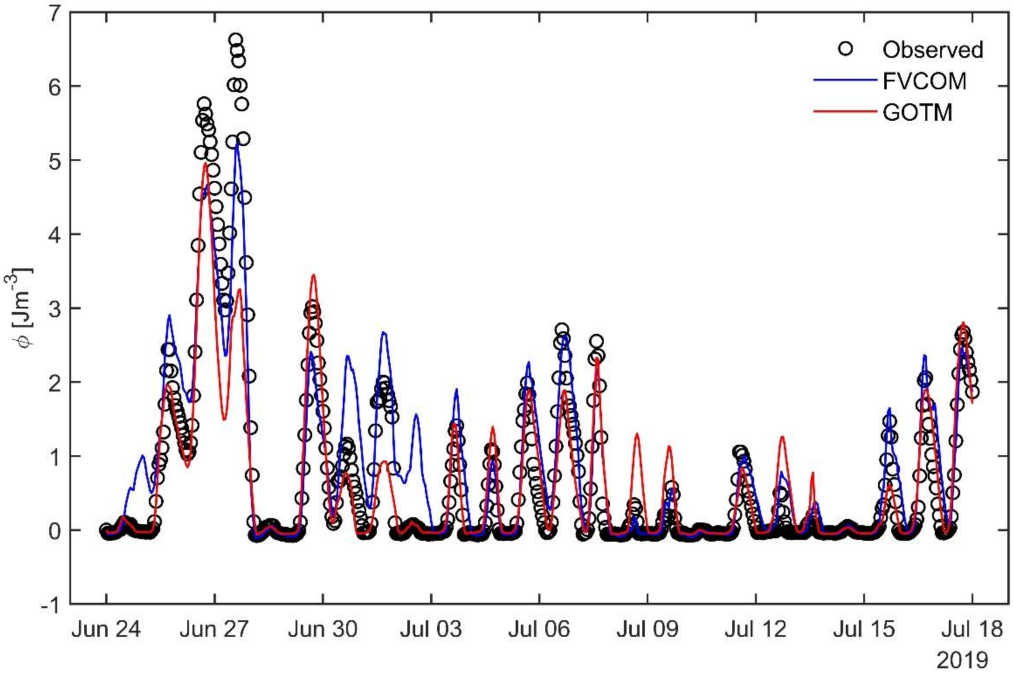
In our paper, (i) we provide a methodology that can deal with the temporal downscaling problem to provide reliable subdaily forcing data, and (ii) perform an analysis for the first time to estimate stratification changes for the future on a subdaily scale. Our weather generation method can be applied to regions where future regional climate datasets are uncertain for some meteorological variables (e.g., wind) and only past observations are available.

2) While the authors devote substantial effort to describing the weather generation and overall modeling workflow, the actual lake model setup, calibration, and validation are insufficiently documented. These aspects are critical for any numerical simulation study, yet the authors only briefly states that "model setup, including the calibration and validation procedures and their results, are provided in Török and Torma (2025, 2024)." This is not adequate. Moreover, In Section 3.1, the validation focuses primarily on the temporal downscaling procedure. However, it should also include validation of the lake model itself, particularly its ability to reproduce thermal dynamics, including both stratified and mixed conditions. Without such validation, the reliability of the projected results remains uncertain.

We can just partly agree with this comment. We have performed detailed model calibration, validation, and evaluation, which has already been published (Török and Torma, 2025). In that paper, we compare and assess model simulations with measurements for several years. In that evaluation, we not only provide metrics but also analyze how well the model captures stratification dynamics when slight deviations occur in the depth-averaged temperature. Furthermore, we also evaluate the modeled vertical profiles. In addition, the evaluation included the modeled turbulent heat fluxes at the air-lake interface and the applicability of reanalysis data. Since these results have already been published, we believe it is neither necessary nor appropriate to repeat them in another article. Nevertheless, we will provide a brief summary of the model's accuracy in the revised manuscript, noting the type of details available in the referenced article.

Furthermore, Lake Balaton is a large (surface area ~596 km²) and elongated lake, where spatial heterogeneity, especially differential heating driven by meteorological forcing, is likely to play an important role in thermal dynamics. The current modeling approach appears to neglect the spatial variability of meteorological inputs, which may limit its ability to accurately represent lake-wide processes.

The division of the distinct basins and the elongated shape is irrelevant given the lake's simple bathymetry. The depth increases quite quickly around the shorelines, and most of the lake's open water areas possess very mild slopes. In addition, the lake's throughflow is negligible. The area of the lake is 596 km^2 with a mean depth of 3.5 m, while the mean flow of the largest tributary is only 4.5 m^3/s, and the outflow is almost zero due to water level regulation. This results in a very high residence time of more than 2 years. Altogether, this means that meteorology uniquely drives the hydrodynamic and thermal dynamics in the case of Lake Balaton. Due to the lake's simple bathymetry – the lake is like a pan – spatial variability plays a weak role in long-term climatic variations, including both temperature evolution and stratification. A good, convincing example of this is that stratification intensity strongly correlates with bathymetry, as shown in Fig. 11 of Lükő et al. (2026), who conducted a 3D FVCOM model analysis of stratification using varying wind-forcing models for the same period that we used for model calibration. We show here how well the 1D (GOTM) model can capture stratification compared to a 3D model (FVCOM) with spatially varying wind forcing:



The 3D model provides only slightly better results, even though its spatial wind forcing includes micrometeorological features such as internal boundary-layer development, which RCM models cannot capture. There are also mixed and stratified events, which are better simulated by GOTM, e.g., June 24th, June 29th, or June 30th.

We agree that we make a simplification by neglecting the spatial variability of meteorological inputs. But there are two reasons to do so. First, Li et al. (2024) analyzed long-term surface water temperature data from the lake over the past decades using remote-sensed data. They showed that there is no significant variation due to meteorological forcings; rather, water temperature correlates with bathymetry. In this sense, the chosen monitoring location represents the entire lake well, as it is located in the middle of a basin and has a water depth very close to the lake's average depth. Second, spatial variability is important for wind, whereas the other forcing factors exhibit little spatial variability at our site. Furthermore, it cannot be expected from climate models (not even from RCMs) that they can describe micro- and mesoscale variations caused by surrounding orography or by the lake itself, given their spatial resolution, which characterizes the study site.

In addition, although the authors considers several meteorological variables (e.g., shortwave radiation, air temperature, and relative humidity), the treatment of wind forcing is unclear and appears insufficient despite some discussion of the limitation of wind in Discussion. Wind is a primary driver of mixing and thermal structure in open-water systems, and its long-term variability is well documented. For example, surface wind speeds have increased by 10–20% in some regions (e.g., over Lake Superior; Desai et al., 2009), while decreasing by similar magnitudes elsewhere (Vautard et al., 2010). Neglecting or inadequately representing wind forcing raises concerns about the physical realism of the model and the robustness of the conclusions.

Desai A R, Austin J A, Bennington V, et al. Stronger winds over a large lake in response to weakening air-to-lake temperature gradient[J]. Nature Geoscience, 2009, 2(12): 855-858.

Vautard R, Cattiaux J, Yiou P, et al. Northern Hemisphere atmospheric stilling partly attributed to an increase in surface roughness[J]. Nature geoscience, 2010, 3(11): 756-761.

Wind plays a crucial role in the mixing process, but we also need to emphasize that the diurnal cycle of stratification in shallow lakes is dominated by daytime warming and nighttime surface cooling, the latter resulting in a daily overturn. Wind leads to continuously mixed periods and is not the driver of the diurnal behavior. Nevertheless, we tried to emphasize wind generation and the related uncertainties.

The FORESEE climate database for Central Europe and the Carpathian Basin contains downscaled and bias-corrected datasets at 0.1° resolution for several variables from 12 GCM-RCM model chains. It doesn't include wind, and it is not incidental or unique. Unfortunately, climate models show high uncertainty regarding wind. Not just climate models but also reanalysis data are uncertain, as Frassl et al. (2018) pointed out. For example, Jung and Schnidler (2023) reported low R2 values (~0.5) for wind speeds in Germany using the CORDEX models. Even wind atlases were more reliable than climate models. Studies highlighted that, instead of individual model chains, projections should rely on ensemble averages (Jung and Schnidler, 2023; Carvalho et al., 2021; Kjellström et al., 2011). Carvalho et al. (2021) also reported that CMIP6 future wind resource projections for Europe show significant differences compared to CMIP5. Given these studies, we cannot expect to obtain reliable spatial variation across the lake surface, as it remains at a small scale compared to the GCM models' resolutions. Dynamic downscaling products by RCMs and the following bias corrections based on CMIP6 GCM results are still underway.

Having the predicted uncertainties and the lack of regional downscaled data, we have chosen a method to relate wind speed to other meteorological variables. On the one hand, the weather generator provided reliable time series for the present period, as simulated lake-temperature dynamics are well aligned with those observed as we compared lake simulations for the present period (Fig. 3). On the other hand, with this method, we assume that the relationship between wind and other variables will be the same for the future. Other temporal downscaling techniques also cannot overcome this assumption. Ayala et al (2020) used a neural network for temporal downscaling of the wind. Besides, it requires daily wind data for the future; it has been trained to present conditions similar to our methodology. The Teddy tool (Zabel and Poschlod, 2023) faces the same problem because it relies on daily climate analogs from the training (present) period. Furthermore, both techniques showed the worst accuracy for wind. Our approach can be applied to other regions where downscaled climate data are not available. Finally, the statistics of the generated wind data are well aligned with the wind speed predictions for Central Europe, showing a decrease of 1-3% when compared with the available results in the Discussion section.

3) The Results section appears to lack direct presentation of the simulated thermal structure. In particular, there is no clear visualization of depth–time thermal dynamics showing the evolution of stratification and mixing. Instead, the manuscript immediately proceeds to climate change projections and derived metrics. This makes it difficult to assess the physical realism of the model simulations. For a modeling study of this type, it is essential to first demonstrate that the model can reproduce realistic thermal behavior, including both stratified and mixed conditions, before moving on to climate-driven changes. The absence of such direct simulation outputs (e.g., temperature contour plots over depth and time) significantly limits the reader's ability to evaluate the robustness of the results.

Considering the visualization of depth-time dynamics (or vertical profiles), we agree that it is missing from the manuscript and could help improve understanding. We will provide these types of plots and the accompanying analysis in the revised version. Regarding the model's realistic reproduction of dynamics (i.e., model validation), we still don't intend to replicate the results of our earlier studies.

In addition, regarding the climate forcing, the analysis appears to focus primarily on precipitation and air temperature. It is unclear why other key meteorological drivers—particularly wind—are not explicitly considered in terms of their long-term trends. Wind forcing plays a dominant role in regulating mixing and stratification dynamics, especially in shallow polymictic systems. Neglecting its projected changes raises concerns about the completeness and physical consistency of the forcing framework.

We agree that the wind trend should be more explicitly given in text, not just in the Discussion. Precipitation and air temperature play key roles as the two main inputs to the weather generator. Furthermore, climate model development in general focuses primarily on these two parameters, as do regional climate data bias-correction efforts. As a result, these two parameters can achieve the highest accuracy, which is why we are primarily relying on them.

4) The Discussion section remains largely descriptive and does not sufficiently interpret the results within a mechanistic or conceptual framework. Wind forcing is a key driver of

thermal stratification and mixing dynamics, especially in shallow polymictic lakes. However, the modeling framework appears to lack an explicit treatment of wind variability, particularly its long-term trends under climate change. This omission raises significant concerns regarding the reliability of the projections, as changes in wind forcing can fundamentally alter stratification and mixing regimes. Furthermore, the discussion of ecological implications is largely generic and reflects well-established textbook knowledge (e.g., increasing temperature leads to reduced dissolved oxygen, and stronger stratification promotes hypoxia). To enhance the scientific significance of the work, the authors should provide quantitative estimates—for example, how much dissolved oxygen is expected to decline under different climate scenarios?

We will revise and extend the Discussion section. The weather generator model does not explicitly link wind trend to a climate model wind result; however, the generated wind speed decreases very similarly to the results of the above-mentioned and the referred studies. In our opinion, this strengthens the weather generator's reliability.

5) There are also many concerns about Figs.

Figure 1 is overly simplified and lacks essential information. The figure should be structured into clearly defined panels (e.g., (a), (b), and (c)), each with explicit descriptions in the caption. It is also unclear whether the single measurement site, located in the western corner of the lake, is representative of the entire Lake Balaton system. Given the elongated morphology and known spatial heterogeneity of the lake, one site may not adequately capture basin-wide thermal dynamics.

In addition, the figure lacks geographic coordinates. The colored lines should also be explicitly defined in the caption (e.g., as bathymetric contours or isobaths).

We revise the figure, providing defined panels and explicit descriptions for those. However, we think this figure is very easy to understand, and the inner panels are also very clear. The geographical situation can be determined from them, while the legend shows the horizontal scale instead of the coordinates. Nevertheless, we will revise and complete the figure.

We have discussed above the representativeness of the location and the spatial heterogeneities of the forcings.

Figure 2 is overly complex and difficult to follow. The figure contains too many elements, including multiple GOTM simulations (1 h and 6 h), different meteorological datasets, and repeated bias correction steps, which together reduce clarity. In particular, the logic within the "Lake modeling" section is unclear. The relationship between the 1-hour and 6-hour simulations is not well explained, and it is not evident why both are required. The use of dashed lines further adds ambiguity, as their meaning is not clearly defined.

Moreover, the caption is overly descriptive but does not clearly communicate the purpose of the figure. It would benefit from an opening sentence that concisely states the figure's objective, for example: *"Figure 2. Schematic overview of the modeling framework used to simulate lake thermal dynamics under present and future climate conditions."*

We agree that the figure is complicated, as we added the sensitivity analysis of the downscaling regarding the time resolution. We will revise the figure to make it more easily understandable.

Figure 3 provides an important validation of the weather generator; however, the comparison remains purely qualitative. The agreement between datasets is assessed visually without any quantitative metrics (e.g., RMSE, MAE, or statistical tests), which limits the strength of the validation. In addition, the use of probability density functions may not be the most clear way to present the comparison. More direct and interpretable plots (e.g., scatter plots, time series comparisons, or bias plots) could improve clarity and make differences between datasets more transparent.

The recommendation to present the comparison using scatter plots or time series is pointless, since the black lines show model simulations forced by the meteorological time series generated by the weather generator. The weather generator provides synthetic time series for the required variables, as climate models do for the future. The key of weather generation is that we can create time series with subdaily resolution from daily datasets. Still, the generated high-resolution time series don't correspond to the given days. The weather generator generates time series with the same climatic features. With this figure and table, we aim to show that, for the present period, the simulated lake climate conditions in terms of temperature and stratification are well captured by the generated subdaily meteorological time series based on daily inputs, indicating that the weather generator can perform temporal downscaling while retaining the climate. Given this comment, we will revise the description of the weather generator and the general methodology.

Figures 5–10 appear to be repetitive and provide limited additional insight. While multiple figures are presented, they largely convey similar information (e.g., temperature trends, distributions, or stratification metrics) without offering new perspectives. For example, Figure 5 shows temperature trends, Figure 6 presents distributions, and Figures 7–9 repeat similar analyses for related variables. This redundancy reduces the overall impact of the results. The authors are encouraged to consolidate these figures and focus on presenting the most informative and non-redundant results, potentially by combining variables or highlighting key relationships rather than repeating similar plots.

We cannot agree that these figures don't provide additional insights, even though they are of the same type. Figures 5 and 7 show the temperature and stratification increments, while Figures 6 and 8 show their sensitivity to water depth, respectively. Their behavior is also different. Figure 9 presents the number of those type of events when stratification survives nighttime cooling and persists for more than a day. Unfortunately, this latter information is missing from the caption. These events are crucial as anoxic conditions can develop with significantly higher probability and lead to an algal bloom. We want to note that this type of projection has not been done before. Similarly, heatwave events for diurnally stratified lakes were not estimated for the future (Fig. 10). As we plan to add a plot on depth-time dynamics, we will reduce the number of these figures.

Given this context, the scientific question addressed here and the some simple results are not particularly novel. Overall, the study appears to be incremental rather than providing a substantial conceptual advance, and may be more suitable for a regional or more specialized journal.

Even though some of the obtained results are not highly surprising, the results are novel because, to our knowledge, there doesn't exist such a detailed analysis for a large shallow lake's future condition, and especially its diurnal behavior. The applied lake model is well-validated. The suggested weather generator and downscaling methodology can be applied to many regions where climate models exhibit high uncertainty. Furthermore, climate databases – with a few exceptions – don't provide subdaily data. As a result, we believe our study is not lacking in novelty, and the central research questions have not yet been analyzed in sufficient detail; therefore, our study cannot be described as incremental.

Naturally, upon reviewing the questions and criticisms about our methodology, we will revise the methodological description and, especially, the discussion to support our methods, simplifications, and assumptions.

References

Ayala, A. I., Moras, S., & Pierson, D. C. (2020). Simulations of future changes in thermal structure of Lake Erken using daily and synthetic hourly climate model data. Hydrology and Earth System Sciences, 24, 3311–3330. https://doi.org/10.5194/hess-24-3311-2020

Carvalho, D., Rocha, A., Costoya, X., deCastro, M., & Gómez-Gesteira, M. (2021). Wind energy resource over Europe under CMIP6 future climate projections: What changes from CMIP5 to CMIP6. Renewable and Sustainable Energy Reviews, 151, 111594. https://doi.org/10.1016/j.rser.2021.111594

Frassl, M. A., Boehrer, B., Holtermann, P. L., Hu, W., Klingbeil, K., Peng, Z., Zhu, J., & Rinke, K. (2018). Opportunities and limits of using meteorological reanalysis data for simulating seasonal to sub-daily water temperature dynamics in a large shallow lake. Water, 10(5), 594. https://doi.org/10.3390/w10050594

Jung, C., & Schindler, D. (2023). Comprehensive validation of 68 wind speed models highlights the benefits of ensemble approaches. Energy Conversion and Management, 286, 117012. https://doi.org/10.1016/j.enconman.2023.117012

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G., & Ullerstig, A. (2011). 21st century changes in the European climate: Uncertainties derived from an ensemble of regional climate model simulations. Tellus A: Dynamic Meteorology and Oceanography, 63(1), 24–40. https://doi.org/10.1111/j.1600-0870.2010.00475.x

Lükő, G., Krámer, T., & Torma, P. (2026). Effects of spatial wind variability on currents and temperature dynamics of shallow lakes. Journal of Great Lakes Research, 52(2), 102775. https://doi.org/10.1016/j.jglr.2026.102775

Piccioni, F., Casenave, C., Lemaire, B. J., Le Moigne, P., Dubois, P., & Vinçon-Leite, B. (2021). The thermal response of small and shallow lakes to climate change: new insights from 3D hindcast modelling. Earth System Dynamics, 12, 439–456. https://doi.org/10.5194/esd-12-439-2021

Zabel, F., & Poschlod, B. (2023). The Teddy tool v1.1: temporal disaggregation of daily climate model data for climate impact analysis. Geoscientific Model Development, 16, 5383–5399. https://doi.org/10.5194/gmd-16-5383-2023