

Response to reviewer comments: Calibration guidelines and a runoff-isotope module for lake proxy system modeling (PRYSM v2.0)

Rebecca G. Topness¹, Elizabeth K. Thomas¹, Michaela Fendrock^{1,2}, Gerard A. Otiniano¹

¹Department of Earth Sciences, University at Buffalo, Buffalo, NY, 14260, USA

²Alfred University, Alfred, NY, 14802, USA

Correspondence to: Rebecca G. Topness (rgtopnes@buffalo.edu)

We would like to thank one anonymous reviewer (RC1) and Yanbin Lei (RC2) for their helpful comments on our manuscript. Our responses are below, where reviewer comments are in blue and our responses are in black. Line numbers refer to those in the revised manuscript (not the tracked changes document).

Response to RC1:

This paper is an excellent approach to calibrating a proxy system model, and its method is a valuable contribution to the proxy-system model community. The extensive observational dataset used to validate the calibration is impressive. I have mostly very minor comments. My only large-scale ish comment would be that the paper felt slightly unfocused at times. That isn't the end of the world, as it is a methods paper and likely part of a broader project to understand these lake systems throughout different past climates, but I think adding a bit more (a paragraph at most should be more than enough) discussion about a couple things could help focus this manuscript and make your results more relevant to a wider audience:

We thank the reviewer for the positive feedback and clear, constructive comments on our work. We address your comments below and provide detailed responses to changes in the manuscript.

What do your results reveal about lake dynamics/PRYSM? Do the parameter ranges that perform best reveal anything about these lakes, lakes in general, or the proxy system model?

Our results reveal that the PRYSM environment submodel parameters are most sensitive to the warm, ice-free season. During this temporal window, observations are most informative because they can constrain lake dynamics related to evaporation, stratification, and hydrogen and oxygen isotope fractionation and result in parameter values that generalize best to unseen data.

We make this claim more explicit in the Discussion section by adding a header that summarizes that point: "5.2 Observations from the warm, ice-free season are most informative for environment submodel calibration". We also reiterate this finding in the Conclusion in line 694:

When time, resources, or site location limit field season duration, we recommend targeting observations from the warm, ice-free season, which were most informative for calibration of Bear Lake and Red Pond.

We also learned that lake hydrology influences whether model inputs or model parameters dominate when simulating isotope values in lake water. Because Bear Lake is open-basin, the isotope values of the precipitation inputs influence the simulated isotope values in Bear Lake water more than Red Pond, a closed-basin lake. Because Red Pond's isotope values are dominated by evaporation, parameter values are more important for simulating that lake.

We moved discussion of that point, which was in the Results section, to the Discussion section under the subsection "5.2 Modeled temperature outperforms $\delta^2\text{H}$ values of lake water profiles".

Line 592: Modeled $\delta^2\text{H}$ values in Bear Lake and Red Pond water reproduce the approximate magnitude and seasonal timing of our observations (Fig. 4), but errors are large compared to simulated temperature profiles. We explored parameter uncertainty and isotope values in precipitation inputs as two potential sources of error. For Red Pond, thermal and isotope parameters lead to a large spread in modeled $\delta^2\text{H}$ values in the epilimnion during the ice-free season (Fig. 5). This makes sense because those parameters impact evaporation amount and hydrogen isotope fractionation during evaporation, which strongly influence the $\delta^2\text{H}$ values in Red Pond as a closed-basin lake. Because Bear Lake is hydrologically open and receives substantial runoff due to its large catchment, its hydrogen isotope composition is less strongly influenced by thermal and isotope parameters. For open-basin systems like Bear Lake, the modeled $\delta^2\text{H}$ values in the epilimnion are more sensitive to the $\delta^2\text{H}$ values in precipitation inputs. Despite uncertainties with simulating $\delta^2\text{H}$ values in these lakes using the environment submodel, we view them as useful representations of the two lake systems on seasonal and longer timescales.

It seems like your results suggest that deuterium is fairly sensitive to both forcing inputs and parameter choice. This would imply to me that using lake proxy system models for isotope-based proxies doesn't work very well. Is that a fair takeaway? Can anything be done to improve the PSM performance?

Yes, that's a fair takeaway. Depending on lake hydrology, the site may be more or less sensitive to forcing inputs vs. parameter choice. Because input for the isotope values in precipitation is a monthly climatology (which is the case for most lake sites given short or no available observational datasets), the set-up may not be suited for investigating lake response to synoptic-scale precipitation events, but useful for answering paleoclimate questions dealing with seasonal or longer timescales.

To highlight which timescales this model and our framework would be most useful, we added some context for the magnitude of error when calibrated to all available $\delta^2\text{H}$ profile observations and in what cases this model would be useful for simulating lake water isotopes in the Results section:

Line 426: *RMSE ensemble averages for $\delta^2\text{H}$ values in the surface water are 13.5 ‰ and 6.9 ‰ for Bear Lake and Red Pond (Table 3). Given a total observed seasonal range of approximately 40 ‰ for both lakes at the surface, the errors account for 34 % and 17 % of the seasonal range for Bear Lake and Red Pond, respectively.*

Line 600: *Despite uncertainties with simulating $\delta^2\text{H}$ values in these lakes using the environment submodel, we view our simulations as useful representations of the two lake systems on seasonal and longer timescales.*

Line 674 (Conclusion): *Given limitations for isotope values in precipitation inputs, our modeling framework may not be well-suited for simulating the response of $\delta^2\text{H}$ values in lake water to high frequency precipitation events, but useful for investigating paleoclimate questions over seasonal or longer timescales.*

We offer two suggestions to improve PSM performance regarding simulating isotopes in lake water: more monitoring of isotope values in precipitation (to generate better model inputs) and parameterization of the inputs to characterize year-to-year variability and uncertainty, when observational datasets are unavailable. We moved this paragraph to a new subsection called “5.1 Modeled temperature outperforms $\delta^2\text{H}$ values of lake water profiles” so that the ideas for model improvement follow the discussion and evaluation of the model performance (lines 606-616).

Otherwise I think this paper is pretty much ready to go! Great work.

Thank you!

Minor comments:

Line 36: ‘i.e.’ should be ‘e.g.’ (unless I am misunderstanding what is meant by this parenthetical)

Thank you for pointing out this typo. We have corrected the line to use ‘e.g.’ instead of ‘i.e.’ because the parenthetical gives examples of climate variables. We have also corrected the same typo in line 32, which has a similar parenthetical.

Line 46 when you say ‘individual changes’ do you mean the changes are applied individually, and not all at once?

Correct; in the studies we cited, changes are applied to individual variables in the climate forcing one at a time to test the model’s sensitivity. We have clarified this phrase in the text to say “changes to individual variables in the climate forcing.”

Line 46-47 this should either be ‘increased/decreased’ or ‘increase/decrease in’ for each of these examples

We updated this line to use “increase/decrease in”.

Lines 55-58 the sentence ‘resulting in a proxy time series that is translated into climate variables’ is confusing to me, because my understanding of forward models was that they translate climate variables into proxy units rather than the other way around. But more broadly I am not sure you need to go into too much detail about Bayesian modeling schema given that you are not using one. You can just say something like ‘some proxies, like terrestrial pollen, don’t require PSMs that incorporate lake dynamics but, x and y that use water isotopes, do’

Yes – forward models translate climate variables into proxy units. The Bayesian application we describe is an inverse approach, in which the goal is to constrain the possible climate conditions that resulted in a real, observed proxy time series. This is done by comparing a forward-modeled proxy time series with an observed proxy time series. When the time series agree, the climate forcing that resulted in the agreement represents plausible climate conditions for the proxy data.

We removed the following sentence:

Climate forcings are perturbed until simulated and observed proxy data agree, resulting in a proxy time series that is translated into climate variables, including uncertainty.

And instead summarized in the previous sentence (Line 53):

*Another application of PSMs is within Bayesian inference frameworks, in which a calibrated PSM converts a wide range of climate scenarios into simulated proxy data **until simulated and observed proxy data agree** (Guiot et al., 2009; Parnell et al., 2015).*

We simplified “the lake system’s response to climate” to “lake dynamics” in Line 59: *Yet, application to other lake sediment paleoclimate proxies (such as aquatic plant leaf waxes and branched dialkyl glycerol tetraethers) that grow within the water column may require simulation of lake dynamics.*

Line 230 could you explain a bit more how this creek is relevant for your runoff model? Does it runoff into either of your lakes? Does the proxy system model include catchment-wide area information?

We use streamflow observations from Canadaway Creek to assess the performance of the runoff model and select appropriate runoff parameter ranges before lake model calibration. Canadaway Creek does not runoff into either Red Pond or Bear Lake, but Canadaway Creek is the only monitored stream that is not dammed within the region of our lake sites. Therefore, catchment-area normalized streamflow observations from this site provide a reasonable model target for the amount and timing of runoff entering the catchments of our lake study sites.

We added a sentence to describe the logic behind using these streamflow observations to calibrate the runoff module:

Line 225: *Streamflow observations, which reflect runoff from rainfall and snowmelt, provide a reasonable target for the amount and timing of runoff entering the catchments of our lake study sites.*

Catchment area is a parameter within the proxy system model (as noted in line 117). When the model runs, runoff amounts in the meteorological forcing file are multiplied by the prescribed catchment area to get volume of runoff. Catchment information such as vegetation is not included in the runoff or lake models used in our study.

We moved the sentence in line 171 (*These runoff module outputs force the environment submodel, which converts and transforms input runoff amounts into volume in meters cubed by multiplying amounts by the prescribed catchment area.*) to the first paragraph in the section called “2.2 Runoff module” so the functionality of the runoff module is clearer for the reader from the beginning (now at line 134).

Table 2 could you include the distance of each source of observational data from to your lake sites? Or does this not apply if they are within the same grid cell for ERA5?

The observational datasets from Jamestown airport (which we use to validate ERA5 data for Red Pond) and Dunkirk airport (which we use to validate ERA5 data for Bear Lake) are within the area of the ERA5 grid cells for their respective lakes.

We added this detail to the text to clarify this point (line 432):

*Quantile-mapping is often used to remove bias from reanalysis data (e.g. Morrill et al., 2019), but we do not correct model inputs due to minimal differences with **nearby meteorological***

observations. We compared ERA5 variables with observations from weather stations that are located within the same ERA5 grid cells as their respective lakes (Table 2.

Canadaway Creek and Amherst, New York (site for observations of isotope values in precipitation) lie outside of the ERA5 grid cells but are the nearest available sites for those observation types.

Because the locations are available in Table 2 (latitudes and longitudes) and on the location map (Figure 1), we feel including the distance of each source of observational data from the lake sites is unnecessary.

Line 388 are these RMSE values the average of all 1000 simulations or the 30 you selected? Could you add text about why Bear lake may have such a high error compared to red pond to the discussion?

Those RMSE values are the average of the 30 simulations we selected. In the text, we clarified that the “30 best-performing” simulations represent the “ensemble average” RMSE values we describe:

Line 386: *We select the 30 best-performing (lowest RMSE) water temperature profile simulations for each lake as our calibrated ensemble.*

Bear Lake likely has a higher error compared to Red Pond because Red Pond is meromictic and the water column is largely stable throughout the year, while Bear Lake mixes twice a year. We elaborate on water temperature profile performance for the lakes when calibrated to all observations in the section of the Discussion called “5.1 Modeled profiles of water temperature outperform isotopes”:

Lines 580 to 590: *Bear Lake and Red Pond thermal dynamics are well-represented in our simulations, as indicated by low RMSE scores when calibrated to all available water temperature profile observations (Table 3). Yet, the ensemble average RMSE for Bear Lake is more than two times higher than for Red Pond. This likely stems from Red Pond’s greater stability as a meromictic lake, in which water column temperatures are largely unchanging year to year below 8 meters depth. Meanwhile, Bear Lake mixes completely twice a year, meaning that the temperature of the entire water column responds to atmospheric conditions (Fig. 3). The shortwave extinction coefficient is a sensitive parameter for simulating temperature profiles in these lakes (Table 1; Fig. S12). All parameters in the environment submodel are static and do not change throughout the simulations. However, Secchi disk depth observations from Bear Lake illustrate seasonally changing water clarity, and therefore shortwave extinction coefficient (Fig. S7). Allowing seasonal or yearly flexibility in this parameter could be useful for improving model*

fit further; given that water clarity can vary over time and this could be reflected in multi-year observational datasets used for calibration.

Line 600 does the turbidity (and thus shortwave extinction coefficient) of the lakes change over the course of the year? Does the parameter related to this value also change month to month, or is it a single number?

Yes, based on our observations, the turbidity in Bear Lake changes over the course of the year (see Supplement Figure S7). We only have one Secchi disk depth observation from Red Pond, but the turbidity likely varies in that lake as well.

The shortwave extinction coefficient parameter is a single number and remains constant throughout the entire simulation. We suggest in the Discussion that allowing the shortwave extinction coefficient to vary might improve model fit:

Line 585: The shortwave extinction coefficient is a sensitive parameter for simulating temperature profiles in these lakes (Table 1; Fig. S12). All parameters in the environment submodel are static and do not change throughout the simulations. However, Secchi disk depth observations from Bear Lake illustrate seasonally changing water clarity, and therefore shortwave extinction coefficient (Fig. S7). Allowing seasonal or yearly flexibility in this parameter could be useful for improving model fit further; given that water clarity can vary over time and this could be reflected in multi-year observational datasets used for calibration.

In addressing this comment, we noticed that both the terms “light extinction coefficient” and “shortwave extinction coefficient” were used throughout the manuscript to refer to the same parameter. We edited the text to use “shortwave extinction coefficient” consistently throughout.

We also standardized “environment submodel” throughout (instead of “lake model” or “environment model”) and adjusted in-text units and numbers that did not match GMD writing guidelines (e.g. use a space between number and unit when written in the text).

Response to RC2, Yanbin Lei:

Lake sediments preserve valuable information about past climate change, such as air temperature and precipitation, but how to interpret them properly is challenging because it is affected by multiple climate signal as well as lake dynamics. The authors integrated a simple runoff-isotope module by using multi-year observational datasets from the two lakes with contrasting morphometry and hydrology. This study about the modern process of runoff-lake water isotope is very useful to interpret proxies in lake sediment. It is worthy of being published after minor revision. Here are some comments:

Thank you for the helpful and encouraging comments on our manuscript. We address the comments below.

1, Language is difficult to understand for me. It would be good to improve it and made it more readable.

Thank you for this feedback. We made several changes to the manuscript to make it more readable.

- We used both the terms “light extinction coefficient” and “shortwave extinction coefficient” throughout the manuscript to refer to the same parameter. We edited the text to use “shortwave extinction coefficient” consistently throughout.
- We standardized “environment submodel” throughout (instead of “lake model” or “environment model”) and adjusted in-text units and numbers that did not match GMD writing guidelines (e.g. use a space between number and unit when written in the text).
- We made changes throughout the text to follow proper terminology when describing hydrogen and oxygen isotope values in water following Sharp (2017). This framework helps clarify that elements, rather than water, have isotopes (i.e. “stable oxygen and hydrogen isotope values in water” vs. “stable water isotopes”) and that the specific isotope should be included when describing fractionation (i.e. “isotope fractionation” vs. “hydrogen isotope fractionation”).

For example, in the Abstract, instead of:

First, lake water isotopes are the target of many proxies preserved in lake sediments, but the PSM does not have a built-in catchment model to aid simulation of lake water isotopes.

We amend this sentence to:

*First, **stable oxygen and hydrogen isotope values in lake water** are the target of many proxies preserved in lake sediments, but the PSM does not have a built-in catchment model to aid simulation of **isotope values**.*

Additionally, RC1 provided similar feedback to make specific sections of the manuscript clearer, which we address in our response to RC1. We are grateful for this feedback that enhanced the readability of the manuscript.

2, This study simulates the runoff-lake water isotopes on a very high resolution (hourly or daily) by using comprehensive in-situ observations of climate variables and lake water isotope and temperature. However, lake sediment is usually inter-annual, decadal or centennial time scales. Factors influencing lake water isotope may differ on different time scale. How to balance the time scale should also be considered.

We agree that bridging the gap between using high-resolution simulations from the PSM to interpret paleoclimate data from lake sediment records that integrate longer time scales is an important concept to highlight in the manuscript.

To clarify time scale aspects of our modeling framework:

- Model forcing: Water isotopes in runoff are calculated based on air temperature and precipitation amount (daily resolution) and hydrogen and oxygen isotope values in precipitation (amount-weighted monthly average climatology). Because input for the isotope values in precipitation is a monthly climatology (which is the case for most lake sites given short or no available observational datasets), we conclude that our modeling framework may not be suited for investigating lake response to synoptic-scale precipitation events, but useful for answering paleoclimate questions dealing with seasonal or longer timescales.

We clarify on what timescales the PSM is most useful for simulating isotope values in lake water in the Discussion and Conclusion, and how model outputs may be used to compare with proxy time series:

Line 598 (Discussion): *Despite uncertainties with simulating δ^2H values in these lakes using the environment submodel, we view our simulations as useful representations of the two lake systems on seasonal and longer timescales. Proxy time series in lake sediments often integrate years to decades within a single sample. Therefore, daily model outputs can be aggregated to interannual or interdecadal time scales to obtain values that are directly comparable to proxies (e.g. Morrill et al., 2019).*

Line 674 (Conclusion): *Given limitations for isotope values in precipitation inputs, our modeling framework may not be well-suited for simulating the response of δ^2H values in lake water to high frequency precipitation events, but useful for investigating paleoclimate questions over seasonal or longer timescales. Because paleoclimate data from lake sediments often integrate years or decades within a single sample, daily proxy system model outputs can be aggregated for direct comparison with proxy time series.*

- Model calibration: Using modern, in-situ observations (typically daily measurements or profiles) to calibrate PSM parameters is necessary to check that the model can reproduce the lake system as it is today before using the calibrated model to answer paleoclimate questions. This approach is standard in proxy system modeling studies (e.g. Acharya et al., 2025; Corcoran et al., 2023; Dee et al., 2021; Holtzman et al., 2025; Marshall et al., 2025; Morrill et al., 2019).

We edited text to highlight this point in the Introduction (line 59): *A comprehensive understanding of the lake system based on modern observations and meteorological forcing is foundational for applying the PSM to interpret paleoclimate data.*

- Parameter values over time: While using modern observations to calibrate the PSM is standard, this assumes that parameter values for the study lake do not change over time. We have added a paragraph to the section **5.3 Limitations** in the Discussion to address that assumption and provide suggestions (lines 674-683):

Another limitation when applying the PSM to interpret paleoclimate data is whether environment submodel parameter values remain constant over time. While modern observations are critical for selecting the best-performing parameter sets that reproduce the lake system, some processes that these parameters represent may change across paleoclimate time scales. For example, a lake catchment that hosted glaciers in the past but lacks them today may have experienced shifts in water clarity over time. This may render a shortwave extinction coefficient calibrated using modern observations inadequate for paleoclimate simulations. Similarly, past climate regimes could alter runoff-to-precipitation ratios, shifting both the amount and isotope values of runoff. Our runoff module offers some flexibility in this regard because several parameters control runoff-to-precipitation ratios and they are each only activated when user-specified air temperature thresholds are met. Testing the sensitivity of environment submodel parameters using the framework developed here could evaluate the impact key parameters have on a lake system prior to using the PSM to compare modeled values with paleoclimate data.

References

Morrill, C., Meador, E., Livneh, B., Liefert, D. T., and Shuman, B. N.: Quantitative model-data comparison of mid-Holocene lake-level change in the central Rocky Mountains, *Clim. Dyn.*, 53, 1077–1094, <https://doi.org/10.1007/s00382-019-04633-3>, 2019.

Sharp, Z: *Principles of Stable Isotope Geochemistry*, 2nd ed., <https://doi.org/10.25844/h9q1-0p82>, 2017.