We thank Reviewer1 for the constructive comments and detailed scrutiny of our manuscript. We agree that there are points that can be further clarified and we are happy to address the main and specific points that the Reviewer has brought up.

General comments

[…] This study goes one step further, as it has the underlying assumption that all climate records are indeed the same (equation in line 228). Spelled out, this means that NGRIP Ca2+ = Speleothem d18O, which is obviously not true as they are very different physical quantities, controlled by different processes. Admittedly, some of the controlling processes may be shared, but the assumption of the applied method is much stronger: It implies the existence of a linear function that relates NGRIP Ca2+ and speleothem d18O. […] the basic assumption underlying the approach presented here (i.e., NGRIP Ca2+ = speleothem d18O) is incorrect and should hence not be used.

We appreciate the Reviewer raising this point, as it allows us to further clarify details of the method. The notation in the equation indicates an approximation (≈) rather than an equality. It is also important to note that this assumption is performed after scaling both data sets to have a strict range of [-1,1], which means $\forall\, t \in T : -1 \leq u(t) \leq 1 \wedge -1 \leq g(\tau(t)) \leq 1$, where T is the window range of the target (this will be clearly explained in the revision).

We agree that a nonlinear alignment between the records is difficult to capture with a single linear function. However, as described in the manuscript, we employ a piecewise linear function to model the alignment between the NGRIP Ca2+ and speleothem $\delta^{18}O$ records. A piecewise linear function is fundamentally different than a simple linear function, as it works by dividing the domain into K separate intervals and fitting a different linear segment within each interval. The linear segments are connected at breakpoints, where the slope can change. This enables the piecewise function to locally approximate nonlinear patterns by using linear segments that capture the essence of more complex functional forms in each region of the domain.

In fact, using this segmented piecewise linear technique provides several key advantages compared to a simple global linear function:

- It allows flexible modelling of nonlinear shapes by adapting the slope in each interval.
- Computationally efficient optimization algorithms can be applied by leveraging the linear segments. A purely nonlinear function would be more challenging to optimize.
- Accuracy can be systematically improved by adding more segments. The nonlinearity is approximated to any desired level.
- The approach balances accuracy with efficiency. More complex nonlinear functions could overfit given uncertainties in the data.

In summary, the piecewise linear methodology enables tractably approximating the nonlinear alignment relationship while facilitating optimization. We understand how these nuanced distinctions may have been unclear in the original manuscript, and we welcome this opportunity to explain our innovative approach more fully in the revised version of the manuscript.

If the authors nonetheless want to follow this approach, they need to i) clearly state that their model assumptions are not fulfilled…

Please see our detailed reply above. We will discuss our approach in more detail in the revised version of the manuscript.

…and ii) discuss these drawbacks and provide additional tests to demonstrate the robustness of the results.

This is a good idea and we are happy to provide two new Δt transfer functions based on NGRIP and GRIP $\delta^{18}O$ records, respectively. The new results are internally coherent and support the findings obtained using NGRIP Ca2+, which ultimately lend strength to our conclusions. These results will be incorporated and discussed in the new version of the manuscript.

1. What determines the inferred timescale shift during the LGM, when there is little co-variability between NGRIP Ca and EASM PC1 (see figure 4)?

Between 18-24ka there is little align-able structure in the timeseries and in fact the model sticks to the information obtained outside this interval, i.e. effectively the Δt does not move much, and the alignment uncertainty grows accordingly. This is to be expected and in line with the design of our alignment model.

2. Which timescale offset is inferred when only the period between 15 – 22 kaBP (or other subsections) is synchronized (and both records are standardized only for this period)?

Unfortunately, the method is not built to align very short timeseries with little structure and low signal-to-noise ratios. We hope that this issue is resolved by providing additional synchronization tests and targets using NGRIP and GRIP $\delta^{18}O$, which demonstrate that the Δt estimates are overall robust during this period. In addition, the Δt during this critical interval is corroborated by independent estimates published by Dong et al., 2022 (see Reviewer2's suggestions/comments). This will be clarified in the new version of the manuscript.

3. How would the results differ if NGRIP d18O was used instead of Ca2+?

Thanks for this suggestion. This is a sensible request and more in line with the premise of our manuscript, i.e. we show the physical relationship between Greenland air temperature and precipitation in the EASM region (see Fig. 1). The results are qualitatively consistent with those based on NGRIP Ca2+, which, again, demonstrates that our method is overall robust (please see our replies above).

4. How would the results (and uncertainties) differ if the uncertainty sigma_ui in the model was increased sufficiently to fulfil the model assumption (NGRIP Ca2+ = speleothem d18O within error).

Thanks for bringing this up. It should be noted that the input and target are scaled between -1 and 1, so by using a 0.1 stdev we are effectively covering 30% of the observable window, and on top of that we are using a heavy tailed distribution (t-distro) which means that we are assuming an uncertainty that fulfils the NGRIP Ca2+ = speleothem $\delta^{18}O$ assumption. We should also mention that we are employing overly conservative estimates for sigma using a multiplying factor of 2 (this will be clarified in the revised version of the manuscript).

Further, the results need to be evaluated more critically with respect to previous studies:

1. Please include the timescale differences by Corrick et al. into figures 4 & 5.

We will include Corrick's Δt estimates in the new version of the manuscript.

2. It appears that most other studies (Buizert et al. / Corrick et al. / Martin et al) found systematically smaller timescale differences then the results presented here. Why?

We can certainly mention this in the manuscript. The small differences may stem from the fact that previous studies used only Hulu Cave $\delta^{18}O$ data, whereas here we use a more comprehensive approach that incorporates several EASM spelothem records.

Specific comments

L205 (eq. 1): Maybe I got this wrong but looking at this equations and trying to put in [units]: m must be [years/m]; so t must be [m] not time; so tau is defined on depth? If so, please use a different symbol as t is time later on.

We appreciate the reviewer raising this insightful question about the units in equation 1. It allows us to clarify that in our case, m is dimensionless, representing an expansion/compaction parameter in units of years/years. Meanwhile, τ0, δ, and ci have units of years. This will be clarified in the revised version of the manuscript.

L180: "in response to changes in accumulation" since you're not modelling accumulation, maybe better "in response to miscounting"? L213-214: "…distinct depositional environments…" But you only model the ice core alignement and their accumulation rates are certainly autocorrelated? It is ok to do it like that, but I am not sure I agree with the explanation. L214-216: Isn't it that: You are not modelling the timescale (or ice accumulation) but only minor modifications of it (counting errors), which do not need to be autorcorrelated.

This is correct. Thank you for point this out. Instead of using the term "accumulation rates" we will discuss the model results in terms of compaction/expansion of the original timescale.

L228: "u(ti) = g(tau(ti))" See major comments. This is obviously not true and should be discussed.

As we mentioned in our response to the previous question regarding equation 1, we agree it is crucial to use clear and consistent notation to avoid confusion. The reviewer is correct that "u(ti) = g(tau(ti))" is imprecise shorthand and could be misinterpreted. Nevertheless, we use the notation u(ti) ≈ g(τ(ti)), which means ∃ τ(z) | u(zi) ≈ g(τ(zi)). Note that we use ≈ and not =.

L244 (Eq. 3): This was defined for comparing 14C-dates to a 14C-calibration. I.e., similar physical quantities. Because your sigma_ui is too small to fulfil the model (u=g) the vast majority of the data is essentially treated as outliers in the gamma-distribution. See major comments.

It is important to note that u(ti) and g(τ(ti)) represent the NGRIP Ca2+ and aligned speleothem δ18O records, respectively, after rescaling the data to the interval [-1,1]. This rescaling means that the uncertainty σ_ui used in the t-distribution becomes a conservative estimate around the rescaled record u(zi). However, the use of the t-distribution has proven robust against outliers. Therefore, any data points that become outliers due to the rescaling assumptions do not significantly affect the resulting inferences of the alignment function τ(z). The t-distribution's heavy tails downweigh the influence of extreme values. In summary, rescaling the records to [-1,1] provides a simple standardized domain for comparing the data, while the t-distribution likelihood protects against artifacts from this transformation when inferring the optimal τ(z). This will be clarified in the new version of the manuscript.

L331-332: The agreement between the records is not very convincing. Please discuss critically. What is the correlation coefficient? What is the error of the model (u=g) after alignment?

We appreciate the reviewer's feedback on discussing agreement between the aligned records. However, we deem applying traditional metrics like a correlation coefficient unnecessary in our Bayesian alignment framework. The optimization process inherently identifies the maximum likelihood alignment given the uncertainties around each data point u(ti) and g(τ(ti)).

Specifically, in each step of inferring the posterior distribution for the piecewise linear function τ(t), the likelihood is calculated based on the t-distribution residuals between u(ti) and g(τ(ti)). The Bayesian approach thus quantitatively determines the optimal nonlinear alignment that maximizes the joint likelihood. Therefore, rather than introducing additional metrics, we believe the optimal uncertainties around τ(z) themselves demonstrate the credible alignment between the records.

Other specific comments

All the other minor comments and suggestions raised by the Reviewer will be respected in our revised manuscript.