Duan et al. present a new multiproxy speleothem record of East Asian summer monsoon hydroclimate variability in northern China during the 8.2 ka event. The record is of high resolution, has a robust age model, contains multiple proxies, and is well interpreted. The scientific questions being asked are relevant to the journal, and the paper is overall well written. Thus, I recommend the manuscript for publication and have only minor comments to add to previous reviews:

I recommend the authors providing scatter plots and/or correlation matrix showing the relationship and accompanying statistics between the trace elements and stable isotopes.

Re: We thank the reviewer for the positive comments and constructive suggestions on our work. Unfortunately, the Pearson correlation results show that there is no significant and strong relationship (r<0.3, p>0.2) between the trace elements and stable isotopes, although the visual inspection suggests conspicuous excursions at ~8.4 and 8.2 ka BP in all records. Intriguingly, when we focus on the interval of 8.5-8.0 ka BP, the correlations become significant (p<0.05) and the coefficients are higher than 0.3, except the Ba/Ca ratio (p>0.1). This could be attributed to the short duration of climate events relative to the entire record, presumably pointing to the complex local controlling factors when the external climate pressure was not at play.

Might be worth plotting up and citing the recent paper by Wood et al. (2024) that documents the 8.2 ka event in northern Laos (Wood, Christopher T., Kathleen R. Johnson, Lindsey E. Lewis, Kevin Wright, Jessica K. Wang, Andrea Borsato, Michael L. Griffiths et al. "High-Resolution, Multiproxy Speleothem Record of the 8.2 ka Event From Mainland Southeast Asia."*Paleoceanography and Paleoclimatology* 38, no. 12 (2023): e2023PA004675). This record also shows no sign of an earlier (i.e., ~8.4 ka) drought event, which may provide further support for their hypothesis that this precursor event was restricted to the higher latitudes.

Re: Wood et al provide an unprecedented multiproxy 8.2 ka record based on speleothem from Mainland Southeast Asia. They reveal a dry 8.2 ka event for the first time and possible pluvial precursor event, whereas the dry 8.4 ka event unraveled by this study is absent. In our opinion, whether the 8.4 ka event is recorded or not is not dependent on the latitudes, but the climate sensitivity of archive site. For example, Huangyuan and Wuya caves are located in the margin of Asian summer monsoon, and Hoti cave is situated in the margin of Indian summer monsoon (23°5′N). In contrast, the record of Wood et al is from the core region of Intertropical convergence zone belt (1195 mm annual precipitation), thereby the weak climatic signal from the high latitudes during the 8.4 ka event could have been covered by local factors. We cited this work in the discussion section.

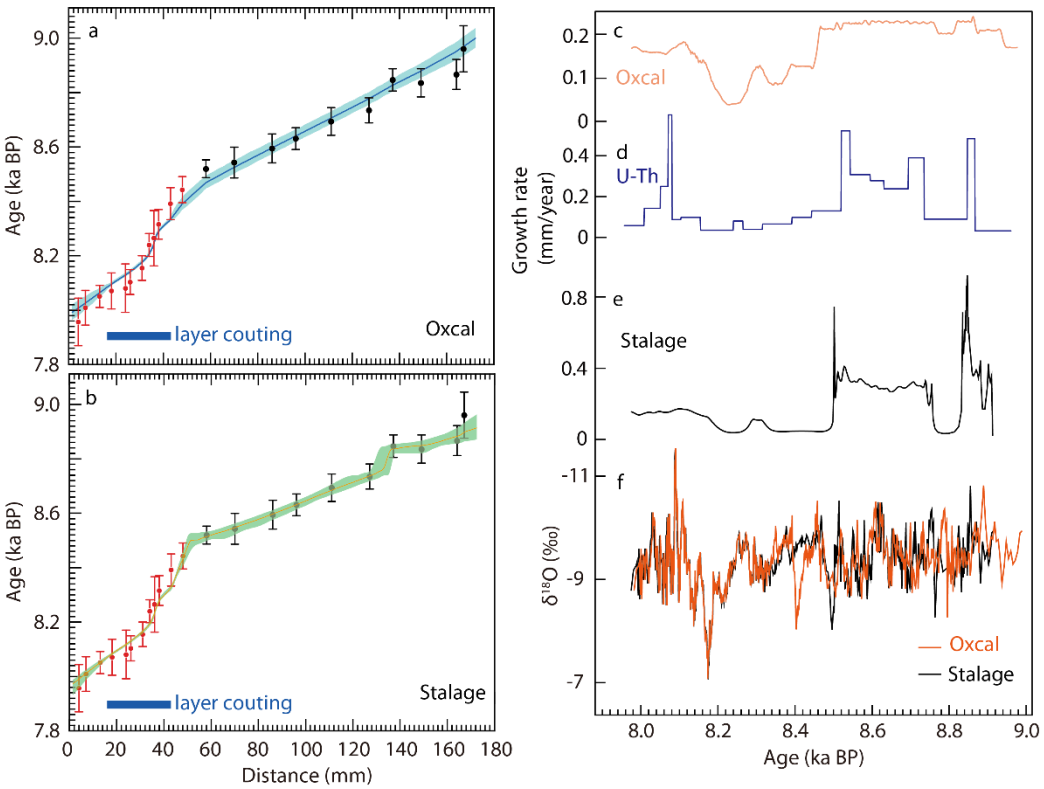
I am curious as to the timing of the 8.2 ka event in the d18O using different age model algorithms. For instance, OxCal shows a very constant growth rate, particularly during the earlier part of the record, that doesn't appear to be matched by the U-Th dates—i.e., the dates suggest more variability in the growth rate than is observed in the age model. Also, beginning at a depth of ~70 mm, the age model passes through the confidence intervals of

most dates rather than the actual date. Is this also the case if another algorithm is used instead (e.g., COPRA, StalAge etc). There was also no mention of the methods for layer counting and the age model derived from this method.

Re: We thank the reviewer's concerns about the chronology reconstruction and growth rate.

We compare the growth rate results based on U-Th dates and Stalage and Oxcal age models. As can be seen, the growth rate is more variable, especially in the stage before 8.5 ka BP, in U-Th age and Stalage modelled result, whereas is smoothed in the other age model. The smoothed curve could result from the algorithm of age model, which is based on the principle that the growth rate of speleothem unlikely changes within short time. To produce a fitting curve, the growth rate among contiguous U-Th ages is smoothed within or even sometime a little beyond the dating errors. It seems that the chronology established by Stalage program is more suitable relative to the Oxcal method because the former one has fewer out-of-confidence interval U-Th ages. Therefore, we replace the BH-2 records established by the Stalage age model in the revised manuscript. However, it is worth emphasizing that the growth rates based on all three methods consistently display lower excursion between ~8.5 and 8.1 ka BP, not contradict with our view in the manuscript that the drought conditions, indicated by stable isotopes and trace elements, induced slow growth rate of speleothem. In other words, the choice of reconstruction method of growth rate does not affect our conclusions of a series of climate events during 8.5-8.0 ka BP rather than only one. Even though the possible uncertainty on the accurate age for the 8.4/8.5 ka event, we are quite confident with our 8.2 ka event because of the negligible offset in comparison with floating chronologies for 8.324–8.077 ka BP and ^{230}Th dates within uncertainties.

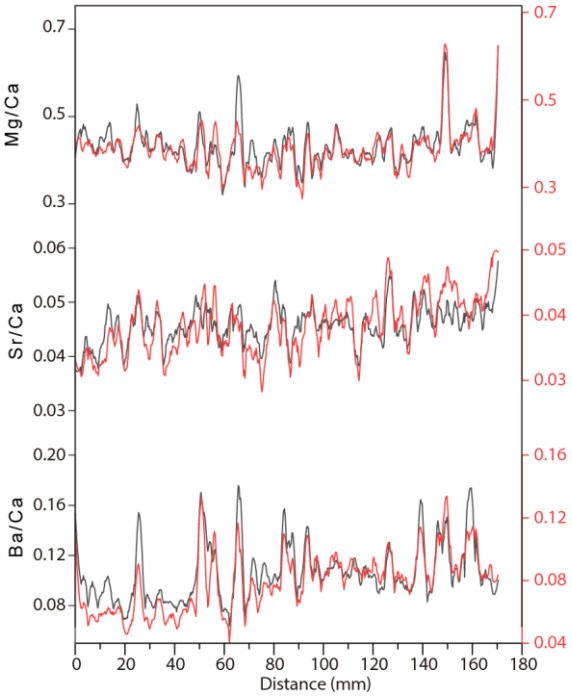
The results of layer counting and relative age model methods were cited from the published paper by Duan et al. (2023), and we didn't revise the data in this study. (*Duan, P., Li, H., Ma, Z., Zhao, J., Dong, X., Sinha, A., et al. (2023). Interdecadal to centennial climate variability surrounding the 8.2 ka event in North China revealed through an annually resolved speleothem record from Beijing. Geophysical Research Letters, 50, e2022GL101182. https://doi.org/10.1029/2022GL101182*). In brief, the least square method was used to establish the chronology through anchoring annual lamina counting to the encompassed seven ^{230}Th dates in the speleothem section of 15-43 mm. To establish the consecutive chronology for the entire record, all above fitting results for each lamina in 16–43 mm (corresponding to 8.077–8.324 ka BP) with uncertainties and the other ^{230}Th dates in the remnant study section are input to age model algorithm. The out-of-confidence interval at ~70 mm could be the smoothed result in order to fitting the massive age constrains at 15-43mm.



Additional methods on the Laser Induced Breakdown Spectroscopy (LIBS) are needed. For example, what standards were used for calibration and correction for instrumental drift? What were the RSDs for standards?

Re: We provide more information about our LIBS method. “Trace element ratios (Mg/Ca, Sr/Ca, Ba/Ca), of which the intensity ratio of emission lines are 285.2 (Mg), 407.8 (Sr), and 493.4 (Ba) nm relative to 373.7 nm (Ca), were measured using Laser Induced Breakdown Spectroscopy (LIBS) following the detailed description in Li et al. (2018). In brief, analyses were performed by pulsing and focusing yttrium-aluminum-garnet-Nd laser beam to 0.1 mm. Emitted plasma from the stalagmite surface was collected by optical fibers and sent to a four-passage spectrometer (Ocean Optics MX500+) to obtain a spectrum within the 200-to 580-nm range. These data were determined through the intensity of characteristic spectral line for each element, and then the intensity ratio of each trace element signal to Ca element was calculated and output as the final result for each point. The obtained record is the median intensity ratio based on 20 pulses at each sampling site after 5 laser shots for pre-cleaning the surface. The measurements were performed continuously along the speleothem's growth axis at 0.3 mm increment and a total of 565 data were obtained. The accuracy of data was ensured through the excellent replicability between two-time measurements instead of inset standard materials because of the overwhelming amount of Ca relative to trace elements in speleothem. The original spectral data were processed using an interface created in MATLAB (2020a). The typical standard deviation for the average signal intensity is less than 0.02 (without unit).”

As can be seen in below figure, trace element ratio records between two-time measurements are broadly similar despite of the discrepancy of absolute values in some points, suggesting the stability of LIBS device and measuring approach.



Line 268: "*presumably were definitely correlated on a broad pattern but did not necessarily exactly follow each other.*" This needs to be quantified as stated above.

Re: please see the response to the first comment.