

Dear reviewer:

We would like to express our feelings of appreciations to you for your kindly help and professional comments to our manuscript entitled “Summer precipitation reconstructed quantitatively using a Mid Holocene  $\delta^{13}\text{C}$  common millet record from Guanzhong Basin, China”. We have tried our best to modify the weakness and flaws pointing out by you. Now, we believe that we made a better work which would probably satisfy the reviewer and suitable to be published. The answer to the comments is listing in the following paragraph.

Thanks again for your help.

Best wishes!

Sincerely,

Yang Qing and Xiaoqiang Li

**Reviewer 2:**

The authors presented carbon isotope data from fossilized common millet seeds collected from archeological sites in northern China. The carbon isotope record was then used to reconstruct summer precipitation during the mid-Holocene, on the basis of the relationship between millet C isotopes and summer precipitation established in a modern process study by the same group (Yang and Li 2015). The authors then used the reconstructed summer precipitation to discuss East Asian summer monsoon dynamics.

I find the idea is intriguing and potentially promotes the use of abundantly available millet seeds as a paleoclimate archive in this part of the world. The modern process investigation as presented in Yang and Li (2015) is an excellent study that shows a robust relationship between  $\delta^{13}\text{C}$  and summer precipitation, despite that I don't fully understand the mechanism (it is counterintuitive that lower summer precipitation correlates with lower  $\delta^{13}\text{C}$  values).

However, I have issues with both presentation and interpretation of the results.

1. I find that the “conventional” climate reconstruction and interpretation as presented in Figure 4 are unsupported, mostly due to the high variability of the record (including the instrumental climate data!), low resolution, and short, snapshot nature of the record (only cover 8-3.5 ka). The high precipitation variability at present in the study region suggests that 3-5 seeds used in each analysis just captured at most 3 to 5 years of precipitation – too small a sample size to capture mean precipitation. As a result, I don't think the correlation as presented in Figure 6 and monsoon discussion is supported.

However, I wonder if that the data set (especially an expanded data set from the region) can be used to document and understand the summer precipitation variability during the Holocene or part of the Holocene. The science question could be: is there increasing summer precipitation variability from the early Holocene to late Holocene, when the summer monsoon and precipitation decline during that time period? Is it possible that not only summer precipitation decreases during the Holocene but also becomes more and more variable and less and less predictable? The data as presented in Fig. 2a seem to suggest that, though the

number of analysis is still low. I wonder if a future expanded study can analyze a larger number of samples per sample (say 30-40 seeds, preferably single seed analysis) to capture the decadal/centennial (depending time resolution) variability in summer precipitation, even just in a few time intervals (early Holocene, mid-Holocene, and late Holocene). Each individual seed is a product of a single season/year – as clearly argued and implemented during 2008 in Yang and Li (2015). This is similar to a study on oxygen isotope analysis of individual foraminifera from deep marine sediments to document ENSO variability (and annual seasonal cycle) at a few time intervals during the Holocene.

I suggest that the authors should focus on the variability rather than mean climate (precipitation).

Thanks for the reviewer's comment and suggestion. Your comment affirms that  $\delta^{13}\text{C}$  of millet can be used as a new proxy to document variability of summer precipitation for the future study, giving us confidence to analyze a larger number of samples per sample at a few time intervals during the Holocene in the future.

It is worth noting the variability of summer precipitation as the reviewer's suggestion. The summer precipitation indeed becomes more and more variable especially after ~5.2 ka BP, which we have added in the discussion of the manuscript.

Since 3-5 millets of a sample from the cultural layer probably formed in the interval of several decades rather than 3 to 5 years, the reconstructed results can indicate the mean precipitation as we careful consideration. Although it is hard to conclude the summer monsoon and precipitation decreases during the Holocene, the reconstructed precipitation during the Holocene exhibits the characteristics of a systemic increase with significant fluctuations. We have discussed precipitation fluctuated significantly and captured three markedly humid periods, showing the mean precipitation as well as the increasing variability especially after 5.2 ka BP in the manuscript, hoping the reviewer's agreement.

2. The writing in general is clear – I commend the authors' effort to make it an easy read. However, I find there are many superlative words to describe the results, and some of these are overstatement. I will provide examples below in my specific comments.

Specific comments:

Title: -focus on precipitation variability? -change "China" to "northern China", for international readership?

Thanks for the reviewer's commendation and suggestion.

-We have changed "China" to "northern China" following the suggestion. We also consider carefully the precipitation variability in the manuscript but didn't show it in the title.

Abstract -add latitude (34.5 N) and perhaps rounded longitude as well, for international readership?

Thanks for the reviewer's suggestion. We have added rounded latitude and longitude on Guanzhong Basin in the abstract.

-there are many superlative descriptors here in the abstract, such as "accurate" (line 17), "robust" (line 19), "reliably" (line 26), "precise" (line 28). It seems to me none of these is needed and justified. Most of similar words should be deleted throughout the text.

Thanks for the reviewer's suggestion. To be more justified, we have removed all the mentioned words above in the abstract as well as some similar words throughout the text.

-the abstract needs to refocus if the authors accept my suggestion above.

Thanks for the reviewer's suggestion. We partly accepted your suggestion above, that the summer precipitation variability from the early Holocene to late Holocene was increasing, and have refocused the abstract in the manuscript.

Introduction -It is unnecessarily too long. In particular, the general discussion on Holocene climate in the first 3 paragraphs on page 1-2 is not really needed. Delete or shorten.

Thanks for the reviewer's suggestion. We have shortened the first 3 paragraphs following the reviewer's suggestion.

-superlative word examples: "accurate" (l 41), "more completely and accurately" (l 60), "robust" (l 135).

Thanks for the reviewer's suggestion. We have removed the words referred above from the manuscript.

Line 91: change "between 5.2-4.3 ka BP", to either "between 5.2 AND 4,3 ka BP", or "at 5.2-4.3 ka BP" (there are other cases of matching "between: : and: : ." in the text) Methods This section reads well.

Thanks for the reviewer's suggestion. We have changed "between 5.2-4.3 ka BP" to "at 5.2-4.3 ka BP" following the suggestion.

Line 156: I wonder if a single seed is large enough for C isotope analysis, but multiple seed analysis still can be used for the variability study as suggested above (but it will be "conservative" reconstruction of precipitation variability, due to averaging of multiple years growth in one sample).

Thanks for the reviewer's question. The millet individuals are very tiny and single millet is not enough for C isotope analysis. Therefore, multiple millets were used as a sample for  $\delta^{13}\text{C}$  measurement, which averaging the multiple years growth. So we consider that the reconstruction result can indicate a decadal averaging precipitation. However, the variability can also be concluded from comparison of series of results at different intervals.

L 169: "1r" = 1 sigma? (67% probability?)

Thanks for the reviewer's kind remind. Yes, it should be 1 sigma (67% probability) and it's our error writing. So, we have corrected "1r" into "1 $\sigma$ ".

L 179: change to "the sampled culture layers" ("section's" is awkward usage)

Thanks for the reviewer's suggestion. We have changed to "the sampled culture layers" following your suggestion.

L 181-186: unclear how it was done.

Thanks for the reviewer's expression. To make the readers understand, we have separated these sentences from the preceding as another section entitled "processing data" and added a figure (Figure 3a) with all  $\delta^{13}\text{C}$  and calibrated age range versus depth as well as the groups we built, hoping it helps the reader to understand.

L 189: delete "ref."

Thanks for the reviewer's suggestion. We have deleted "ref" following your suggestion.

L 190: delete " ," before "demonstrated"

Thanks for the reviewer's suggestion. We have deleted " ," before "demonstrated" as you suggestion.

Results L 209: change "eliminating" to "without considering"

Thanks for the reviewer's suggestion. We have changed "eliminating" to "without considering" following the suggestion.

Line 207-214: I'm confused here. You describe carbonized Neolithic seed remains and modern common millet, but you compare "modern seeds" in the last sentence. Also, millet is more negative than seeds, rather than "positive" as described. Check.

Thanks for the reviewer's kind remind. It should be millets remains and modern millet in the paragraph. To avoid confusion, we have rewrite the sentence as follow: Common millet remains sampled from cultural layers of Guanzhong Basin in our study.....It can thus be seen that the  $\delta^{13}\text{C}$  values of common millet remains are more positive than those of modern millet by  $\sim 2.9\%$ .

L 233-235: "slightly higher" and "a much more humid" is contradictory. Overstatement/over-interpretation?

Thanks for the reviewer's kind remind. To avoid contradictory, we have removed "slightly" and "much" from the sentence and the sentence is now as follow: The  $\delta^{13}\text{C}$  values yielded by ancient common millets are higher than those of modern common millet seeds, suggesting that these ancient plants grew in a more humid environment than today's.

Discussion As I commented above, the mean precipitation reconstruction doesn't allow for much comparison and discussion on summer monsoon, while precipitation variability is potentially a novel aspect of paleoclimate research. Although your current data are not robust enough, it seems to me that it holds great promise for the future project: even just 3 or 4 horizons, with large analysis per horizon.

Thanks for the reviewer's suggestion. As we answered above, we consider that  $\delta^{13}\text{C}$  of millet can reconstruct mean precipitation although the resolution of current data is low. However, we also accept the reviewer's viewpoint that variability becomes higher from the early Holocene to late Holocene. So we weakened the discussion on the absolute value of the mean precipitation reconstruction and added discussion about the increasing variability of summer precipitation as well as provided the markedly humid periods.

L 341: change “1961-2011” to “1951-2011”? which makes 60 years and also is consistent from description earlier.

Thanks for the reviewer’s kind remind. It should be 1951 and we have corrected it to “1951-2011”.

Conclusions L 392-393: “low resolution” and “convincing” are contradictory.

Thanks for the reviewer’s kind remind. To avoid contradictory, we have changed “convincing” to “innovative”.

Tables Table 1 -move latitudes and longitudes from Figure 1 to new columns here.

-what “source” means here? “12” here means “12 culture layers”? if so, spell out.

-change “No.” to lower case and italic “*n*” (to indicate number of analyses or samples)

Thanks for the reviewer’s suggestions.

-We have added latitudes and longitudes in a new column.

-The “sources” means “sample source” and we have added “sample” before “sources”. “12” here was placed in an error box and we have moved it to the right box.

-We replaced “No.” by “*n*” and gave a footnote “*n* means the number of remnant common millet samples derived from the section”.

Table 2 -Change heading to “AMS 14C dates” (“dating data” is unusual”)

-why does a 250-cm-long section (such as BN) have the same or reversed ages? Very rapid accumulation of these layers? I hope it is discussed elsewhere in archeological literature.

-change the heading of column 4 to “AMS 14C date (yr BP)” – it is wrong to say 14C date as “cal yr BP”

-maybe a footnote to indicate the dating lab for OZM

Thanks for the reviewer’s suggestions.

-We have changed heading to “Accelerator mass spectrometry (AMS) dates from Baijia (BJ), Huiduipo (HDP), Manan (MN), Beiniu (BN), and Nansha (NS)”.

-Since all sections selected were cultural layers and deeply affected by human activities, the sections especially from the ash pits accumulate rapidly, which may result in the same ages, and have reverse layer due to disturbance of human activities. The thickness of the sections and the cultural types were discussed in the archeological literature “Atlas of Chinese Cultural Relics: Shannxi Municipality”, which was edited by State Cultural Relics Bureau in 1998, in Chinese. However, the <sup>14</sup>C dates was not adopt in the literature, due to archaeologists usually hold the view that archeological periodization are more reliable. Our radiocarbon dates showed the cultural layers are different from the natural layer. That is why we cannot apply linear interpolation and extension for age verse depth as which usually applied to the natural section.

-We have changed the heading of column 4 to “Radiocarbon age (<sup>14</sup>C yr BP)”.

-We have added a footnote for the dating lab of OZM in the manuscript as follow: All assays were run on the STAR Accelerator, ANSTO, Australia.

Table 3 -maybe indicate ages as “calibrated ages” to avoid confusion

-add footnote to indicate “d13Cre” and “Ps”

Thanks for the reviewer's suggestions.

-We have changed "<sup>14</sup>C age" to "calibrated ages" following your suggestion.

-We have changed " $\delta^{13}\text{C}_{\text{re}}$ " to "corrected  $\delta^{13}\text{C}$ " and also added footnote to indicate corrected  $\delta^{13}\text{C}$  and Ps.

Figure 1. -move latitude and longitude to Table 1

-CorelDRAW12 is not needed to mention, as it is just a map.

Thanks for the reviewer's kind remind. We have moved latitude and longitude to table 1 and removed the sentence mentioned CorelDRAW12 from the footnote.

Figure 2. -need more explanation about panel a (raw data points) and b (box plot) in figure caption

-again Fig. 2a kind of shows increase in precipitation variability from 8 ka to 3 ka. Have you tried a regression of all the data to see if there is a significant decline (in precipitation) during that period as well? (perhaps the number of data points are still low)

Thanks for the reviewer's suggestion.

-We have added more explanation about panel a and panel b in figure caption in the manuscript following your suggestion.

-To be more specific, we redrew the figure 2a, showing all raw data points including  $\delta^{13}\text{C}$  and calibrated age range versus depth. However, a significant decline in  $\delta^{13}\text{C}$  or precipitation from 8 ka BP to 3 ka BP wasn't concluded. But the increasing variability of precipitation is visible, which we have discussed in the manuscript.

Figure 3 -indicate reference in figure caption "Modified or data from Yang and Li (2015)"

Thanks for the reviewer's suggestion. We have added "which data from Yang and Li (2015)" in the figure caption.

Figure 4 -Again, I don't think it is a good way to present the data as groups to get mean climate/precipitation – considering the large variability almost nothing can be concluded here (a.k.a. the pattern is not robust/convincing, because of uncertainty).

Thanks for the reviewer's suggestion. To better present, we have added another panel as figure 5a to display all reconstructed precipitation data, hoping it is helpful to understand the mean precipitation and variability. It can be seen from the panel a that there are three markedly humid periods, which have the mean precipitation higher than the other periods, and the variability of precipitation from 8 ka BP to 3 ka BP becomes increasing obviously.