Comment on cp-2022-14

Anonymous Referee #2

Referee comment on "Accurately calibrated X-ray fluorescence core scanning (XRF-CS) record of Ti☐/☐Al reveals Early Pleistocene aridity and humidity variability over North Africa and its close relationship to low-latitude insolation" by Rick Hennekam et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2022-14-RC2, 2022

Review of the manuscript entitled « Accurately calibrated XRF-CS record of Ti/Al reveals Early Pleistocene aridity/humidity variability over North Africa and its close relationship to low-latitude insolation" by Hennekam et al.

General comments

The study by Hennekam et al. has two major objectives: (1) investigate how to reliably calibrate core-scanner elemental records using the example of core ODP 967 from the eastern Mediterranean Sea, and (2) discuss changes in North African humidity and aridity over the last 3 Ma (with a special focus on the Mid-Pleistocene Transition (MPT), ~1.2-0.7 Ma), and their drivers (orbital parameters, insolation, ice volume). First, the authors test various numbers of WD-XRF calibration samples and two ways of selecting calibration samples, and discuss how much the calibrated core-scanner Ti/Al record compares statistically to the reference Ti/Al record. Then, the authors discuss calibrated Ti/Al changes in terms of North African aridity changes over the last 3 Ma, in agreement with available high-resolution records. Aridity over North Africa was particularly enhanced after the MPT. They confirm the strong control of orbital parameters (precession, obliquity, eccentricity) on North African humidity. Whereas low-latitude forcing dominates between 3 and 1.2 Ma, North African climate became more sensitive to high-latitude climate forcing when global ice volume increased during the MPT.

The manuscript is concise, well written, easy to read. Figures are clear and well explained. The methodology is sound and generally clear. However, I have two major concerns about the manuscript. First, I find the two parts appear to be rather disconnected one from another. It almost gives the impression two small studies have been merged together to build a manuscript.

Second, I find it difficult to identify the new information this study brings in comparison to previous studies. It is stated (lines 292-293) :"our detailed analysis of the 2.3-1.2 Myr interval and extensive testing of the calibration approach is novel." (NB: Is the Data availability section the right place to make such a statement?) However, even if I find the calibration testing exercise interesting (though frustrating by lack of more detailed discussion), I wonder to which extent it is needed for the interpretation of the Ti/Al record (see specific comment 1 below). Also, even if I am not an expert of North African climate over these time scales, the manuscript gives the impression it confirms previous hypotheses on the control of North African humidity (rather than brings novel ideas). I also have the feeling the study brings more insight on North African changes during the MPT than between 2.3 and 1.2 Ma (as indicated in lines 85-86, and 293).

In conclusion, I think this manuscript deserves publication in Climate of the Past, provided the authors are able to better highlight the added value of this study (compared to already published works), to better link the two parts of the manuscript (calibration exercise and interpretation of Ti/Al in terms of North African humidity) and to better highlight the usefulness of the calibration exercise for the study and the community, by further developing its discussion.

Specific comments

- Calibration testing exercise

I wonder to which extend the detailed exercise of comparing calibrations is really needed for the study, for 3 main reasons, which would all require additional discussion in the text.

(a) Why is the XRF calibration published by Grant et al. 2022 not included in the testing exercise? The study by Grant et al. 2022 is cited in lines 72-74. It uses 42 WD-XRF reference samples (cited as more accurate than ED-XRF samples). So why not include this calibration in the comparison?

I had also been wondering how much the calibrated Ti/Al record published in Grant et al. 2022 differs from the Ti/Al record adopted in this study until I reached the Data availability section, where we can read (lines 290-291): "The calibrated XRF-scanning record of Grant et al. (2022) is essentially the same as the final calibrated XRF-scanning record presented here […]. We recommend to use that record for paleoenvironmental purposes." (NB: Is again the Data Availability section the right place for such a statement?) Above all, is a new calibrated Ti/Al ratio necessary here if it is the same as the one published by Grant et al. 2022?

If the calibration exercise remains in the revised manuscript, I would advice to include the

calibration by Grant et al. 2022 in the comparison, extend the discussion on how much calibrated Ti/Al records differ and clarify the ambiguous statement on the record recommended for paleoenvironmental purposes (the one by Grant et al. 2022 or the newly calibrated one with 1060 reference samples?), and for which reasons one record is preferred if they are essentially the same.

(b) The comparison between the different tests of calibration is based on five statistical tests comparing the calibrated core-scanner Ti/Al and the reference WD-XRF dataset (Table 1). I remain highly frustrated by the currently limited discussion (lines 141-147) on how many reference samples are recommended or suitable for the calibration, and which type of selection of reference samples should be preferred (even spacing or Xelerate automatic selection). I find it very difficult with this limited discussion to draw inferences for other calibration studies. In light of the exercise, what is the minimum recommended number of reference samples? (What about the recommendation by Weltje et al. 2015 (equation 21.15a) of having as number of calibration samples at least 3 times the number of elements to be calibrated?) How should reference samples rather be selected: evenly spaced, manually or automatically with Xelerate? Indeed, I am quite surprised to see that even spaced samples seem to give a more robust calibration than automatically selected samples and would have liked to read a more extended discussion in lines 146-147. In summary, I would strongly advice to develop the discussion on the comparison of calibrations to make it more meaningful and useful to the community beyond the case study of core ODP 967.

Similarly, I would advice to add a direct comparison of the differently calibrated Ti/Al records and discuss their possible differences in the text. Indeed, at first sight from Figure 2, there do not seem to be major differences between the various Ti/Al records. Thus, the reader wonders why a detailed calibration exercise is included in the manuscript if all tested calibrations provide relatively similar calibrated log-ratios.

(c) So far, I thought that calibration of core scanner intensities was a requirement in provenance studies (where absolute values of elemental ratios are compared to the composition of source material) and a bonus in classical paleoenvironmental studies, as it is the case here. In my own experience, the calibration modifies the absolute values and amplitude of change of elemental ratios, but not so much their downcore variations. It does not seem to be the case here (Figure 2, mostly below 30 m) and I am curious to know why. Thus, I would also recommend to develop the discussion on how much the calibration modifies the uncalibrated Ti/Al record. I think it would make more convincing and better illustrate the statement of the necessity of the calibration for paleoenvironmental purposes (lines 158-159, 292). It would also reinforce the usefulness of having the exercise comparing the various calibrations within the manuscript and strengthen the link between the two "parts" of the manuscript.

Finally, I think the information provided on the calibration would deserve clarification at two places. First, I would state more clearly in lines 122-124 that 10 elements are calibrated, give the name of calibrated elements and provide (as a supplement?) an illustration of the retained calibration for all elements (e.g. the Xelerate figure with reference vs. predicted concentrations). Second, I wonder how the authors managed to run the Xelerate software with 22 reference samples only for 10 calibrated elements,

when I think the software requires as number of reference samples at least 3 times the number of elements (equation 21.15a in Weltje et al. 2015).

- Changes in North African humidity

(a) As a non-expert on these long time scales, I would have liked to have more information on the chronologies and related age uncertainties. In particular:

What is an estimate of age uncertainties in core ODP 967 (lines 95-99)?

How were constructed the age models of sites ODP 659 and 721-722 (lines 192-194)? What is the related age uncertainty? What is an estimate of the age offset that is expected between the records of these sites and core ODP 967 for the period of interest?

How was estimated the small lead of 2 ka of obliquity over Ti/Al (line 230) and how does it compare to age uncertainties for the period?

(b) I would also have liked to read more detailed information (lines 235-239) on the climate model results (Bosmans et al. 2015a, b). Which type of simulations was run? Climate sensitivity experiments? What was exactly tested? Which results are observed?

(c) I am not fully convinced by the values of the coefficients of correlation between Ti/Al and the Gibraltar relative sea level record (Figure 5c). Can we speak of a high correlation when the absolute value of the coefficient reach 0.3-0.4?

Also in Figure 5 I think the method how "Sapropel intervals are removed in this data set and data accordingly interpolated" (line 475) should be explained.

Technical comments

Line 27: I would write "the longest period and highest amplitude"

Line 45-46: I would write "throughout the Pleistocene"

Line 59: I would write "during 0-1.2 Ma and 2.3-3.2 Ma"

Line 112: Please correct the reference "Zhan, 2005".

Line 131: Please indicate which version of Analyseries has been used.

Line 146-147: I would rather write "(i.e. all calibrated elements)".

Line 459: Is the Ba/Al ratio shown in Figure 3f also calibrated? Please specify.