

Nguyen Le Duy
Section 5.4 Hydrology
Tel.: +49 331 288 1562
Fax: +49 331 288 1570
Email: duy@gfz-potsdam.de

Helmholtz Centre Potsdam
GFZ German Research Centre For Geosciences
Telegrafenberg, 14473 Potsdam, Germany

Potsdam, 10th January 2018

Dr. Lixin Wang
Editor
Hydrology and Earth System Sciences

Re: HESS-2017-164

Dear Dr. Wang,

enclosed please find a thoroughly revised, original manuscript now titled “**What controls the stable isotope composition of precipitation in the Mekong Delta? A model-based statistical approach**”, which is renamed from the previous title “What controls the stable isotope composition of precipitation in the Asian monsoon region?” (Reference #HESS-2017-164) by Nguyen Le Duy, Ingo Heidbüchel, Hanno Meyer, Bruno Merz, Heiko Apel. We are respectfully submitting our revised manuscript again for your consideration in Hydrology and Earth System Sciences.

We have thoroughly revised the paper to take into consideration the constructive comments from the second referee. The following revisions have been made:

- the abstract was slightly extended
- the introduction was completely reworked (focusing more specifically on highlighting the novelties and recent literature)
- the study area description was extended (based on reviewers’ feedback)
- the methodology section 3.5 was changed (deleting unnecessary parts, but adding a discussion about the uncertainties of trajectory analysis and applied measures to mitigate these uncertainties), but other parts were shortened (LMWL regressions)
- the results and discussion section was rewritten (based on reviewers’ feedback)
- the conclusions were extended (highlighting the novelties and adding some thought on the implications for paleo-climate reconstruction).

This manuscript has neither been previously published in any language nor is it under consideration for publication by another journal. All authors have carefully read the revised manuscript and have agreed to its submission to Hydrology and Earth System Sciences. All results and innovations were developed by the authors using Matlab. Figures were generated using ArcGIS and Matlab. We also published the isotopic data in the open access data repository of GFZ. The data is already available to reviewers under:

<http://pmd.gfz-potsdam.de/panmetaworks/review/9e1af507c8fce65a8d740033e5fea31c2e7c58ade81762c235c6f6bbab91166e/>

Thank you for handling the manuscript during the review process, and to the reviewers for their valuable feedback and edits. We look forward to hearing from you.

Sincerely yours,
Nguyen Le Duy
Corresponding Author

RESPONSE TO THE REFEREES' COMMENTS

We sincerely thank the first referee for the acceptance and the second referee for his/her thorough reviews and most constructive comments on our manuscript (Reference #HESS-2017-164). We fully appreciate the reviewers' efforts in providing these informative reports on our research and their insights have led to an improved interpretation of our results. We have taken into full consideration all of these comments and have prepared responses to these as well as information on how the paper was revised following the referees' suggestions. Our responses are provided below **in blue** following the individual comments requiring action from the second reviewer, followed by a marked up version of the manuscript (all changes as appearing in the text are marked **in red**).

Anonymous Referee #1

We thank the first anonymous referee the suggestion to accept the manuscript for publication.

Anonymous Referee #2

This study evaluates local vs. regional controls on the stable isotopic composition of precipitation in the Vietnamese Mekong Delta by assigning relative weights to multiple linear regression coefficients. As stated in the manuscript, distinguishing local and regional controls on precipitation isotope ratios is a critical concern for accurate interpretation of paleo-proxy records. This study applies a very thorough and novel statistical approach to disentangle these factors. However, it is not entirely clear to me how one would invert this procedure, given a record of precipitation isotope ratios, to reconstruct past climate.

Overall, I find the analysis thorough and compelling, though the methodological descriptions are a bit dense and perhaps lose some clarity in being too detailed. I would like to see the presentation condensed and reorganized in places, as well as a bit more discussion about the broader implications of this work for paleo-proxy interpretations. More specific comments are provided below.

We thank the second referee for the positive and constructive comments. Our answers are also included in the revised version of the manuscript.

Introduction - could better focus on the main story.

1. I'd like to see a strong beginning, emphasizing the scientific question at hand. Why not make the second paragraph the lede?

Thank you for this recommendation. We deleted the first paragraph in the Introduction to show a strong beginning the scientific question.

2. The 4th paragraph suggests "other relevant processes were identified..." presumably for the Monsoon Region. Do all the ensuing publications specifically address the Monsoon Region?

We removed reference (Pausata et al., 2011) of which study sites are not explicitly related to the Monsoon Region. All other references were checked for relevance.

3. The 5th paragraph suggests statistical models are "not able to represent the actual processes..." Some re-wording/re-phrasing here is required. All models are a representation. GCMs, for example, can only approximate many physical processes.

The paragraph was reworded.

4. Limitations and assumptions of paleoclimate reconstructions discussed in the 6th paragraph are nicely described.

Thank you for your nice comment.

5. Also in the 6th paragraph: what is the difference in isotopic signatures of Indian and Pacific Ocean air?

The sentence was re-written as follows:

“The isotopic signatures of air masses originating from the Indian Ocean differing considerably from those of the Pacific Ocean, where the average $\delta^{18}O$ of the latter is about 2.5‰ more negative (Araguás - Araguás et al., 1998)”.

6. Paragraph 7 and onward, some of the narrative flow is lost. What is the purpose of discussing advances and limitations of GCMs? There is a statement about developing GCM code being too daunting a task, but there is code and there are researchers actively developing it, so the argument doesn't quite make sense. Are GCMs and Lagrangian models two different ways of approaching paleoclimate reconstructions? How do these models fit with the methods used in this work? It almost seems as though Paragraphs 12 or 13 could directly follow 6: a monofactorial approach has many limitations...therefore this study suggests a multifactorial one. The multiple factors considered include both local and regional meteorological variables, with back trajectories used to characterize the regional ones. Yes, GCMs also allow one to consider both local and regional factors, but, as stated, their complexity can make interpretation difficult. Perhaps the more detailed GCM discussion could be moved to a proper discussion section. This would help focus and condense the Introduction, which would be desirable.

This part has been shortened considerably. The discussion of the advantages/disadvantages of the Eulerian and Lagrangian approaches has been dropped.

7. Page 5, where the importance of multiple factors in influencing precipitation are discussed, this would be a good place to introduce the need for a multiple linear regression approach and tie this paper's statistical approach to the larger scientific questions at hand.

We followed this suggestion and introduced the MLR here.

8. Page 5, Line 22: LMWLs should be defined for those unfamiliar with isotopic analyses. More broadly, it is not clear to me that the LMWLs play a significant role in this analysis other than to show that re-evaporation may be relevant during the dry season. It seems their presentation could be minimized. More on this below.

We dropped the notion of the LMWL in the introduction, as suggested, because it distracts from the actual focus of the study.

9. The Intro ends by emphasizing the drivers of isotopic variation. But isn't the underlying motivation using the isotopic records from the past to interpret hydroclimate? How do we go from one direction to the other?

We end the Introduction with the mentioning of the study area to test the proposed method for the identification of the primary processes controlling isotopic composition, which is the core aspect of the manuscript.

Study area

10. An Long and its relationship to Cao Lanh should be described here. The best description of this is the first paragraph of Section 4.1. Specifically, the paper should describe why it is okay (or at least necessary) to interchange data from these sites.

We added a paragraph describing the need for using data from Cao Lanh, as well as justifying the usage of the data:

“The local climate of An Long is described by data from Cao Lanh station. Cao Lanh is the closest

national meteorological station to An Long with continuous climate records, located approximately 37 km Southeast of An Long. It is assumed that the climatic conditions of An Long and Cao Lanh are similar. The proximity of the two locations and the similar geographical setting (flat topography, located at the Eastern bank of the Mekong river) justify this assumption.“

Methodology – could be shortened.

11. The section begins with “An overview of the proposed methodology...” Yet this is in fact the methodology used. “Proposed” can be dropped.

“Proposed” was dropped.

12. Describing the LMWL is fairly standard practice, and the comparison of three distinct regression methods seems overkill, particularly since all three give equivalent results. I would suggest moving this sensitivity test to supporting information, which would help shorten the methods and the number of figures.

We present and describe only on regression method in the main manuscript, and moved the other two for reference to the supplementary material.

13. Similarly, the description of HYSPLIT is a bit more detailed than really necessary. I’d like to see Section 3.5 considerably shortened.

We changed this paragraph a bit. Specifically, we removed sentences which are not necessary from the section 3.5 and moved the paragraphs discussing “uncertainties” and “quality control measures to increase confidence of backward trajectories” to Discussion (in section 4.1)

14. “Moving distance” is not clear. I believe what is intended is the distance the air parcel moved. It would be helpful to clarify that this is measured (in km?) along the parcel trajectory (as opposed to the Euclidean distance between start and finish).

We clarified this, writing:

“The HYSPLIT outputs, i.e., precipitation, temperature, relative humidity along the backward trajectories, and the length of trajectories (the distance of moisture sources traveled), were used as regional factors potentially controlling the variation of the isotopic composition of precipitation at An Long. Accumulated precipitation, mean values of temperature and humidity of the hourly HYSPLIT output were calculated along the trajectory, as well as the length of the trajectory. All these factors were used as predictors in the MLR.”

15. I would suggest removing the clause “In order to derive figures representative for each trajectory...” from Line 13 on Page 9, as it is not clear.

Thank you for this suggestion. We deleted “In order to derive figures representative for each trajectory” from the sentence.

16. Some of the remaining paragraphs on Page 9 related to HYSPLIT assumptions can be shifted to a Discussion section.

As mentioned above, the remaining paragraphs on Page 9 were move to the discussion.

17. The first paragraph of Section 3.6 is quite clear and helpful in describing the paper’s methodology.

Thank you for your nice comment.

18. Equations 2 and 3 should follow immediately after they are mentioned.

Equations 2 and 3 were moved to follow immediately after they are mentioned.

19. The number of ML regressions considered is quite impressive and reflects the thoroughness of the paper’s approach.

Thank you for your nice comment.

20. I appreciate the fact that the paper openly acknowledges the correlations among predictor variables and address multicollinearity using relative weight analysis. This method will be somewhat new for many readers and should be given a bit more description. (This is one of the only sections where I would recommend expanding the text!)

We re-wrote the section. Please find it in the revised manuscript.

21. I had assumed all weights described in the results are relative weights. Is the RPSS used as well? If so, this is not clear. Similar to my suggestion for LMWLs, I would recommend emphasizing one method and simply stating that other methods did not provide qualitatively different results. This will help streamline the methodology tremendously and help give other researchers a roadmap for conducting a similar statistical analysis for their region(s) of interest.

Thank you for this suggestion. We removed RPSS and used only Johnson's relative weights for the relative importance analysis in the study.

Results – could be reorganized.

22. I might suggest a bit of reorganization (and condensing!) here: what if the section began by describing the local data, contextualized it within the larger region, then discussed the distant moisture sources to the region? This would give some additional motivation for evaluating local vs. regional controls on precipitation as the final, most important segment of this section.

We followed the suggestion and re-structured the result part. However, we did not separate it into Results and Discussion, because this would unnecessarily lengthen the manuscript. Some results need to be repeated in order to understand the discussion, if separated from the results.

23. The first paragraph really belongs in the Methods, as does description of TSV.

This paragraph was moved to Methodology with some editing (see section 3.1).

24. Line 30, Page 12: the d18O values are “noted” or “written” not “plotted.” How about an isotopic bar chart to actually plot them? This would be much easier to “read” than the text.

Thank you for the good ideal. The word “plotted” was replaced by “noted”. We also added isotopic bar charts to the Figure 6 and re-color the text for a better view.

25. As written, it is not clear how section 4.2.1 (LMWLs) answers the local vs. regional control question. See previous comments about shortening the presentation and discussion of LMWLs. The seasonal LMWLs do provide some evidence of secondary fractionation (re-evaporation), which is presumably a local process. But that's really the only message I took from their inclusion in this work (and it's not clear that this is the intended use of the LMWLs in the paper.)

The presentation and discussion of the LMWL were shortened to the minimum, showing only one regression result and illustrating the evidence of some secondary fractionation process.

26. It's not clear from the Methods that the GNIP data will be used to set this paper's measurements within a larger regional context. This could be stated earlier in the paper so that the reader knows to expect this and to understand how the GNIP data will be used.

We underlined the use and purpose of the GNIP data in the method section.

27. Top of Page 14: the paper highlights differences between An Long and Bangkok, but the figure doesn't really show substantial differences. Moreover, wouldn't an unusually dry period tend to enrich An Long compared to Bangkok's climatology? I don't see this in the data. Lastly, it doesn't really make sense that one would use the sites to “represent or complement each other.” Perhaps one could rephrase to say the overall similarity suggests

an important role for regional or larger-scale controls on An Long precipitation isotope ratios.

We rephrased this part to clarify the rationale for the comparison with the variability of the isotopic composition with Bangkok: we want to show that the results obtained are likely representative not only for An Long, but also for a larger region.

28. The Levene test description can be moved to Methods.

The Levene test description was moved to section 3.1.

29. Page 15 first sentence: we can't yet know that precipitation is "mainly controlled by large-scale circulation." What we infer is that it is influenced by other factors such as the large-scale circulation.

Thank you for the comment. We replaced "mainly controlled by" by "influenced by other factors such as".

30. Page 16, Line 7: the correlations can only show a correlation, not that P_hysplit is the dominant control. Our physical understanding of isotopic responses to precipitation is what suggests precipitation is the control.

Thank you for this comment. We removed the sentence "Thus, P_hysplit is likely the dominant factor controlling the isotopic composition of precipitation." from the manuscript.

31. Section 4.4: I'm a bit confused how the MLR models are evaluated. Aren't all factors, including met variables at various heights and for various trajectory lengths considered all at once to select the best model? The section almost suggests the height and length are picked first, and then the best met variables are identified second, which wouldn't make sense. Some re-phrasing is needed.

We clarified this: all factor combinations are used for all travel distances and pressure levels.

32. Page 17 is really quite compelling and well written.

Thank you for your nice comment.

33. Up to 7 predictors for seasonal regressions with 42, 18, and 14 data points is not ideal. Some discussion of this potential limitation would be useful in a proper discussion section.

We now explicitly acknowledge the limitation/uncertainty introduced by the low seasonal sample size, but also give reasons for the validity of the findings.

34. Moreover, it would be useful to see the final best model (and which predictors are included!) for both the annual and seasonal analyses.

We added the Table 5 which show the final best models (both for annual and seasonal analyses) for $\delta^{18}\text{O}$, $\delta^2\text{H}$, and d-excess as the response variable in MLR.

35. Some discussion of why dxs seems to reflect regional processes more than the individual isotope ratios would be useful. Again, this could go in a proper discussion section.

We extended the discussion on the d-excess and provided reasons for the different factors controlling the d-excess in different seasons. We also included a statement about the uncertainty stemming from the limited sample size in the dry season regression.

Conclusion

36. Page 20, Line 20: Perhaps "play a smaller role in influencing" rather than "modulate."

Thank you. We replaced "modulate" by "play a smaller role in influencing" as your suggestion.

37. Page 21, Line 8: scratch "without a priori knowledge or assumptions." The method of

course makes a priori assumptions when picking variables like P and T as predictors of the isotope ratios. Also, assumptions are made about the importance of both local and regional factors.

Thank you. “without a priori knowledge or assumption” was removed.

38. Page 21, Line 20: Where are the LMWLs of all stations compared? Perhaps this statement should just be eliminated as the LMWLs don’t seem to add much to the analysis.

We deleted this statement and refer to the test for similarity in variance instead.

39. Last paragraph: again, how can we go from understanding controls on isotopes to using isotope ratios to reconstruct climate?

We extended this part accordingly. It now reads:

“The results have direct implications for the interpretation of paleorecords of stable water isotopes in terms of past climate conditions for Southeast Asia. This study shows that the factors controlling the isotopic signature of precipitation are changing between and even within seasons, and that regional factors have substantial impacts on the local isotopic composition of rainfall. This needs to be considered in the reconstruction of past climates based on isotopic records: for the presented study area $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values are likely to be representative of the local climatic conditions during the dry season. However, regional factors dominate during most of the rainy season receiving the bulk of the total annual rainfall. In this case, reconstructions of past climates based on paleo isotopic records would have to be carefully interpreted. The proposed approach might open a pathway for an improved reconstruction of paleoclimates based on isotopic records. It may e.g. be used for identifying suitable variables to improve the performance of proxy data assimilation in paleoclimate reconstruction by circulation models. Moreover, assuming that the general circulation is stable over the period considered in paleoclimate reconstruction, which is reasonable for e.g. the Holocene, the presented findings can be used to infer moisture source regions and the strength of the two monsoonal regimes influencing SE-Asia from paleo isotopic records.”

Tables and Figures

40. Table 3: d^{18}O - d^2H order should be swapped in first column, second row, to be consistent with other rows.

Agree. We edited the table as your suggestion.

41. Figure 6, in addition to the isotopic bar chart suggested above, the brown text could be re-colored so it is more distinguishable from the red text.

Thank you. We added the isotopic bar chart and re-colored the text for a better view as your suggestion.

42. Figure 10: Consider plotting the arithmetic vs. amount-weighted means as a difference for faster viewing and interpretation.

Agree. We changed the figure as your suggestion.

43. Figure 11 caption: the best model is “marked” or “annotated” with red text.

Thank you. The word “highlight” was replaced by “marked”.

44. Figure 12 caption: the dots and bars in the top panel should be identified.

Thank you. We re-colored the figure.