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, 4th October, 2017

Dr. Lixin Wang
Editor
Hydrology and Earth System Sciences

Re: HESS-2017-164

Dear Dr. Wang,

enclosed please find a fully 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 for your consideration in Hydrology and Earth System Sciences.

We have fully revised the paper to take into consideration the constructive comments from the two referees. Given the manuscript required major revision, we have not provided a line-by-line list of the changes since line numbers have been altered significantly. Instead, we summarize here the major revisions we have made:

- Rewriting of the abstract
- Rewriting of the introduction (focusing more specifically on highlighting the novelties and recent literature)
- Rewriting of the study area description (including a discussion about the definition of the dry/wet season)
- Rewriting of the methodology – Section 3.5 (adding a discussion about the uncertainties of trajectory analysis and applied measures to mitigate these uncertainties)
- Rewriting of the results and discussion (based on reviewers' feedback)
- Rewriting of the conclusions (focusing more specifically on highlighting the novelties).

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 both referees for their 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 to reviewers are provided below **in blue** following the individual comments requiring action from both reviewers, followed by a marked up version of the manuscript (all changes in the text are marked **in red**).

Anonymous Referee #1

General Comments:

In recent years, a number of empirical, theoretical, and modeling studies have attempted to identify, characterize, and quantify the dominant controls of the stable isotopic composition of rainfall in tropics, particularly in the Asian monsoon domain. Duy et al manuscript, which at a first glance, seems like yet another manuscript along this line, indeed dives much deeper than the previous studies and attempts to provide more rigorous and quantitative assessments of various climatic factors that control stable isotope composition of rainfall in the Asian monsoon domain. Authors present a robust body of observational precipitation isotope data (weekly to bi-weekly samples over ~1.5 years) collected from Vietnamese Mekong Delta region. This observational isotope data has been examined in the context of both local-and-regional-scale station-based climate data (temperature, precipitation amount, humidity), GNIP data, and finally climate data extracted from GDAS gridded dataset, the latter being used to drive the NOAA's HYSPLIT models. Authors conclude that the influence of the different factors on the isotopic condition is best quantified by multiple linear regressions (MLR) of all factor combinations and that explains up to 80% of the variation of $\delta 18O$ of precipitation. This study, like many previous studies, shows that local rainfall amount and temperature play a minor role in controlling the isotopic composition of the rainfall with upstream precipitation amount emerging as the dominant regional control again a result consistent with previous studies, but the author's conclusion is backed by solid quantitative analysis. The manuscript is well-written, free of excessive jargon, logically structured with high-quality figures and graphics that are instructive and easy to understand. In sum, I did not find any major issues with this manuscript and I highly recommend its publication. I have provided here a few comments, which authors may find useful in further improving their manuscript.

We thank the first anonymous referee for the positive and constructive comments.

Specific Comments:

1. Are results of this manuscript sensitive to the choice of gridded dataset (for example, R1/R2) vs GDAS, which was used to drive the HYSPLIT model?

Yes, we acknowledge that the results of this manuscript might be sensitive to the choice of the climate dataset driving the HYSPLIT model. Moreover, the backward-trajectory simulations by HYSPLIT are also influenced by other parameters that have to be defined for running HYSPLIT, such as starting time and height of the trajectories, trajectory duration, vertical motion options, and number of trajectories. Studying the sensitivity of HYSPLIT backward-trajectory simulations would be an interesting topic, but exceeds the scope of this study.

In order to discuss the sensitivity with regard to the choice of the gridded dataset as well as the uncertainties of trajectory analysis, we included this paragraph to the revised manuscript:

"Single backward trajectory computations by the HYSPLIT model can have large uncertainties. The horizontal uncertainty of the trajectory calculations by HYSPLIT has been estimated to be 10–20 % of the travel distance (Draxler and Hess, 1998). While errors in trajectory calculation computed from analyzed wind fields seem to be typical on the order of 20% of the distance travelled (Stohl, 1998), the statistical analysis of a large number of trajectories arriving at a study site would increase the accuracy of the trajectory analysis (Cabello et al., 2008). Harris et al. (2005) studied trajectory model sensitivity to the input meteorological data (focusing on ERA-40 and NCEP/NCAR reanalysis data) and to the vertical transport method. They pointed out five causes of trajectory uncertainty, expressed as percentage of deviation of the average travel distance: 1) minor differences in the computational methodology: 3–4%; 2) time interpolation: 9–25%; 3) vertical transport method: 18–34%; 4) meteorological input data: 30–40%; and 5) combined two-way differences in the vertical transport method and meteorological input data: 39–47%. However, it would be difficult to prove that in all situations a single meteorological data set or a single method of trajectory modeling was superior to another one (Gebhart et al., 2005;Harris et al., 2005). More details about the uncertainties in trajectory modeling were provided by (Stohl, 1998), later by (Fleming et al., 2012) and references therein."

2. Figure 5 shows backtracking trajectories (only those which produced rainfall). Perhaps I missed reading about it but can authors more clearly elaborate on the criteria they applied to establish when a certain air parcel was considered to produce rainfall?

This paragraph was included to the revised manuscript (in section 4.1) to elaborate on the criteria applied to establish when a certain air parcel was considered to produce rainfall in Figure 5.

"Because there is no daily precipitation data recorded at An Long, we used daily precipitation data at Cao Lanh instead. This is the closest national meteorological station, located approximately 37 km Southeast of An Long. Backtracking trajectories in Fig. 5 are plotted for the days when rainfall was recorded at Cao Lanh. This is based on assumption that days with precipitation at Cao Lanh and An Long coincide."

3. Additionally, I think it will be useful to have another figure that shows major cluster tracks (instead of trajectories) and their relative weights. For example, what percentage of trajectories originate from the Indian Ocean vs continental sources during the rainy season?

Thank you very much for this constructive suggestion. We added Figure 6 to the manuscript.

This paragraph was also included to the revised manuscript (in section 3.5) to discuss the trajectory cluster analysis.

"The trajectory cluster analysis is conducted by the HYSPLIT model to group trajectories with similar pathways. The cluster analysis merges these trajectories that are near each other and represents those clusters by their mean trajectory. Differences between trajectories within a cluster are minimized while differences between clusters are maximized. Computationally, trajectories are combined to decrease the number of clusters until the total spatial variance (TSV) starts to increase significantly. This occurs when disparate clusters are combined. This number of clusters is then selected as the optimal cluster number for sorting and combining similar trajectories. More information about the HYSPLIT cluster analysis can be found at https://ready.arl.noaa.gov/documents/Tutorial/html/."

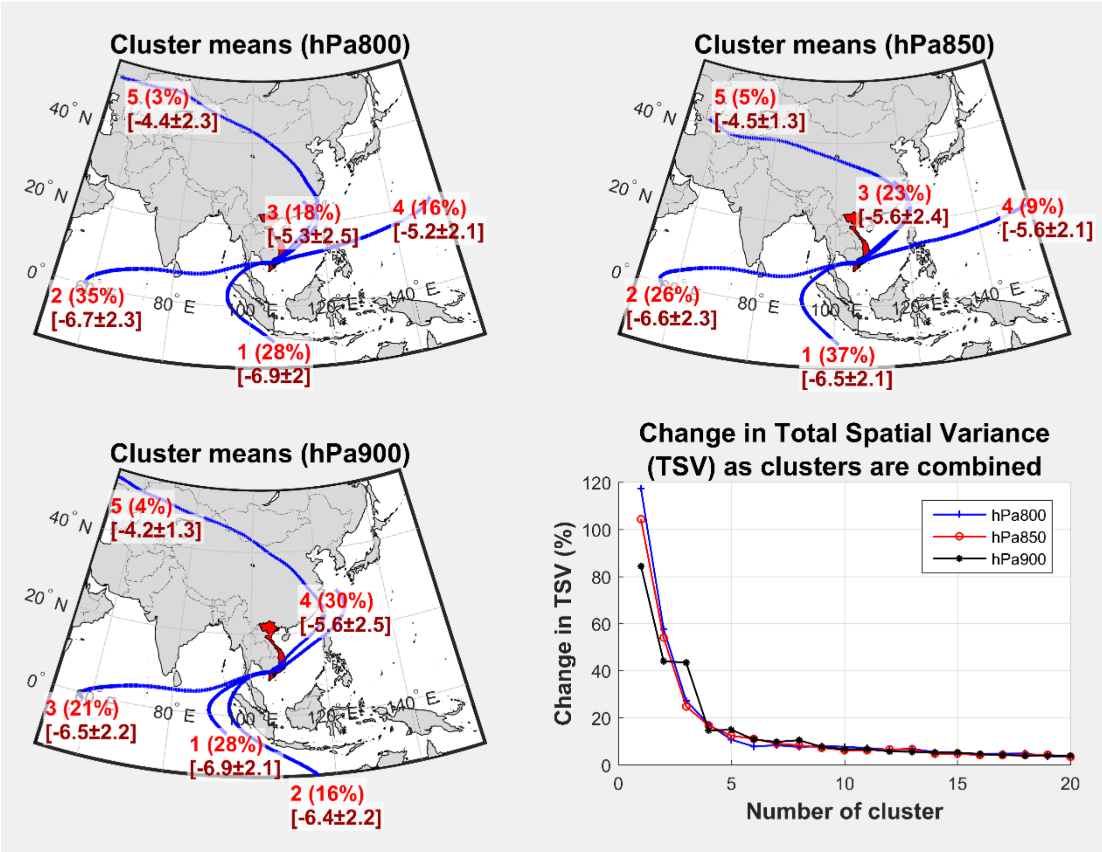


Figure 6: Spatial distribution of vapor trajectories (cluster means) for precipitation days at An Long for 3 barometric surfaces (800, 850, 900 hPa) between June 2014 and December 2015, and change in total spatial variance (TVS) for different cluster numbers. The TSV was used to identify the optimum number of clusters. Red texts indicate the cluster number (1-5) and the percent of all trajectories assigned to each of the five clusters. Brown texts indicate the mean $\delta^{18}O$ values for each cluster plus/minus the standard deviation of each cluster.

Furthermore, can these tracks be fingerprinted with their typical d18O values? I suppose this should not be too difficult given that authors have access to the d18O values of precipitation.

Thank you for this constructive suggestion. This paragraph was included to the revised manuscript (in section 4.1) to discuss how backward trajectories can be fingerprinted with their typical $d^{18}O$ values:

"The mean $\delta^{18}O$ values for the 5 clusters are plotted in Figure 6 (in brown). The mean cluster values are similar for the three pressure levels. Also, the mean values of the two clusters from the Indian Ocean, as well as the two clusters from the Pacific, are similar. For a fingerprinting one also has to consider the variation of the values within the clusters, which partly overlap. This means that the $\delta^{18}O$ values of precipitation in the Mekong Delta cannot be used to uniquely identify the origin of the trajectory. However, they provide a coarse indication of their origin."

4. I think the authors need to be more specific (as opposed to providing generic comments) in suggesting how their conclusions need to be considered in paleoclimate studies. It would be helpful if they can cite some paleoclimate studies where proxy data may have been misinterpreted in light of the results obtained from this study.

The suggestion of citing paleoclimate studies where our findings could have made a difference seems to be appealing, but we have to admit that paleoclimate is not our research focus and that we don't have an encompassing picture about all the past and ongoing research in this field. We thus don't feel qualified to criticize published studies in this field. We rather hope that the paleoclimate community will become aware of our results and model-based statistical approach, and that they might be considered in their future research.

Anonymous Referee #2

General Comments:

In this paper, the authors used their new weekly precipitation isotope dataset in Vietnam's Mekong river delta region for 1.5 years, and they tried to reveal the controls of the temporal variation of the precipitation isotope ratio. To do so, they conducted some statistical analyses, and they concluded that the isotope ratio is controlled by mainly regional scale phenomena (mainly by the previous rainfall activity along the trajectory of air mass) especially during the early rainy season, and the contribution of the control varies by season.

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

In my opinion, even though they conducted multiple methods, nothing is quite new. The control of precipitation isotope had been discussed by many researchers as the authors mentioned, and the authors' findings were already pointed out by many, too. For example, the quantification of the controls was attempted by several model studies including Yoshimura et al., 2003; Risi et al., 2008; Kurita et al., 2011; Ishizaki et al., 2012; etc. Some of these studies do not necessarily focus only Asian monsoon regions, but basically, they tried to reveal more general controls. In these studies, they used GCM or equivalent models to reveal the controls, whereas the present paper used statistical models. Furthermore, by the recent efforts, researchers already began to realize that it is indeed not appropriate to make a simple relationship between precipitation isotopes and climate parameters. The present paper's conclusion of necessity of consideration of multiple climate impacts and temporal (and spatial) dependency on the controls have been explicitly or implicitly stated many times. Therefore, nowadays, more advanced techniques of utilization of isotopic information have been utilized. One of them is data assimilation.

We acknowledge the fact that the results are not new, and that the focus of the paper is the development and testing of the model-based statistical method instead. We also recognize that the title can be quite misleading (as mentioned in major issue #2), and thus may lead to a misunderstanding about the novelty of this study. We therefore modified the title to "What controls the stable isotope composition of precipitation in the Mekong Delta? A model-based statistical approach" and discussed the transferability to the greater region, i.e. SE-Asia. Actually, isotopic data of rainfall has never been collected for the Mekong delta, and therefore the fact that the isotopic variation of the Mekong data is similar to that of Asian monsoon region has never been confirmed before.

We revised the introduction and conclusion to specifically highlight the novelties of the study. Recent literature was also included accordingly. Because the revised introduction and conclusion are too long to present here, please find them in the submitted revised manuscript.

From the above aspect, I have to tell that this paper's methods (multiple regression and trajectory analysis) is no longer insufficient to fulfill the objectives of this study. What I mean is, there is no guarantee that this study's number of 70% regional control can be applied to any other year's temporal variation of precipitation isotopes. In this regard, 1.5-yr long data is not sufficient, too.

Of course, due to the limited length of the time series we cannot be 100% sure that the identified contribution of local and regional factors will be the same in other years. However, as shown in figure 7, the long term monthly isotopic values in Bangkok and the values of our two rainy seasons in the Mekong delta are quite similar. Considering also the climatic similarities between the two locations, this indicates that the recorded isotopic variation is likely to be representative for a longer period and a wider area. This suggests in turn that the identified contribution of the factors could also be the same in other years. Also, the fact that our findings agree with the ones of Ishizaki et al. (2012) supports this assumption.

Figure 7 (in the old-version of manuscript) was edited to include the short-term mean monthly isotopic signature of precipitation of Bangkok, and renamed to Figure 8 (in the revised manuscript). The number of the other figures was edited accordingly.

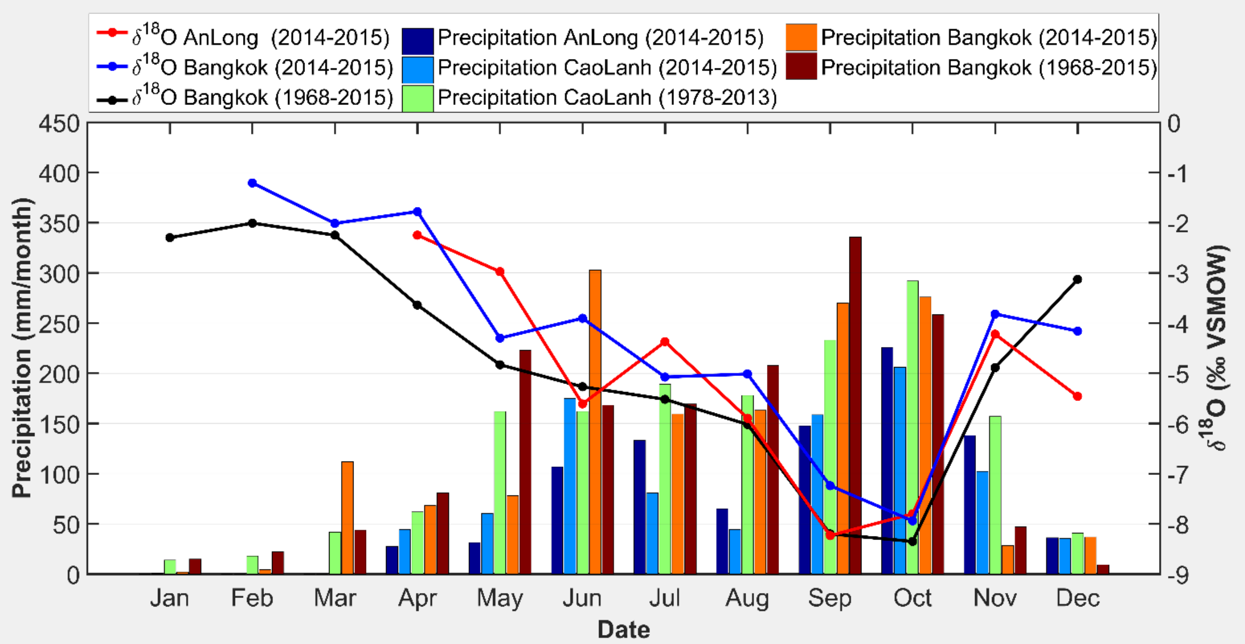


Figure 8: Seasonal variation of the average monthly precipitation for An Long and Cao Lanh and $\delta^{18}O$ values of precipitation for An Long (for the period of observation (red)) and Bangkok (both for the period of observation (blue) and the long-term mean (black)).

Major issue:

1. Drop unnecessary and unrelated analyses. Especially the parts with local meteoric line is not directly related to the conclusion of the study. It is too simple analysis. Even global meteoric line is just conceptual idea (slope of 8 and intercept of 10 is not certain). There maybe some physical reason to have smaller slope, especially by kinetic effect, but in this study, it is out discussed enough. It’s better to drop the part.

You are right that the derivation of a local meteoric water line is a very simple analysis. We still think it provides valuable information for the following reasons:

- From our point of view the analysis of isotopic data by means of a meteoric water lines is a standard for such kind of data and should always be conducted, just as descriptive statistics of other data.
- Up to now, there is neither a LMWL for the Vietnamese Mekong Delta (VMD) nor for the Indochinese Peninsula, which could be used as a baseline for other studies using isotopic data to investigate hydrological processes in this area.
- The close fit of all considered regressions is one piece of evidence indicating that secondary fractionation processes, e.g. sub-cloud evaporation, are insignificant in the study area. This provides support for the discussion of sub-cloud evaporation in Sec. 4.3.1.

2. One point data cannot represent Asian monsoon. Perhaps Mekong river delta data had some similarity with Bangkok, but with only 1.5-yr long data, the authors cannot reject possibility of “by chance”. Furthermore, such similarity is nothing related to that Mekong data represent all Asian monsoon region. The title is quite misleading.

We acknowledge that the title is too generic. We changed it to “What controls the stable isotope composition of precipitation in the Mekong Delta? A model-based statistical approach” and discussed the transferability to the greater region, i.e. SE-Asia. Actually, isotopic data of rainfall has never been collected for the Mekong delta, and therefore the fact that the isotopic variation of the Mekong data is similar to that of Asian monsoon region has never been confirmed before.

We also went at length to illustrate that the variability of the isotopic data is similar to the long term data from Bangkok in order to provide evidence that the derived results might be representative for SE-Asia. This was already discussed in section 4.2, but we added some critical discussion of the issue of representability in the discussion and conclusion of the revised manuscript.

3. Organize the previous literature with focused temporal and spatial scales. The authors listed many previous studies, which partly investigated on precipitation isotope controls, and (implicitly) stated that there is still huge discussion on the controls. However, it is misleading and not true. What is confusing is the controls can be different dependent on temporal and spatial scales. For example, daily variation of precipitation isotopes in some parts of the world is quite likely determined by synoptic-scale moisture circulation, in which previous rainfall activity along the trajectory matters a lot, rather than local precipitation or temperature, and nowadays there is consensus on this in the research community. However, even in the same place, the controls of monthly or interannual time series can be different. It is simply because those smaller scale impact can be offset each other in those scales, so that local signal only remains.

We completely agree that scales matter. This is fundamental to hydrology. What we present is the result for daily variation (or bi-weekly, to be exact) in rainfall, in a monsoonal climate region with a strong seasonal variation. We stressed this more in the discussion and conclusion, and sorted the cited literature according to the scales considered.

4. Limitation of statistical approach with such short-term data. The conclusion of the study is based on the statistical regression using all samples. The authors should validate their statistical model(s) with different independent samples. In this regard, the observation data is perhaps too short.

As described in section 3.6, we use PRESS for selecting the best model. Within PRESS the model is fitted to all data except one, and the missing value is predicted with the fitted model, i.e. not all data is used for fitting the models at once. This procedure is repeated for every data point. Thus PRESS is equivalent to a so called leave-one-out cross validation (LOOCV), as described in section 3.6. LOOCV is the cross validation procedure appropriate for a limited data set, when a standard split sample validation cannot be applied. There are numerous papers available employing this method in different fields of environmental sciences. LOOCV is actually a split sample validation of the regression, where the data is split as often as data points are available. This means that our results are in fact validated.

5. Most importantly, what is new in this study? As I wrote above, it is well known that precipitation isotope is not controlled by a single factor and the relationship can be different in time and space. The finding in this paper is nothing more than these.

We revised the introduction and conclusion to highlight more specific the novelties of this study. As we have stated previously, we acknowledge the fact that our methods (trajectory analysis, multiple linear regression and relative importance analysis) are relatively simple and easy to apply, but we would like to stress again that the combination of these methods to investigate factors controlling isotopic composition in precipitation has never been applied before.

Moreover, our study focuses on the quantification of the impact of the various factors controlling isotopic composition in precipitation. This has not been performed in such an exhaustive way as presented here (as reviewer 1 actually points out particularly). Of course, the qualitative outcome of the study is not novel in itself, but the way we achieved these results constitutes a novel approach. Furthermore, this approach is easily reproducible and contains a rigorous analysis and quantification of the interplay of the different factors. Thus we argue that the manuscript indeed goes beyond just stating that regional factors are more important than local factors for the daily rainfall isotopic composition of the study region. It rather supports this finding by a thorough and reproducible method that combines trajectory modelling and statistical analysis.

In order to stress the novelty of this study, we also included this paragraph to the conclusion:

“The validity of the approach is confirmed by similar, but mainly qualitative results obtained in other studies. The comparable results provide a strong indication that the method is able to identify the dominant factors responsible for the isotopic composition of rainfall without a priori knowledge or assumptions. In contrast to previous studies, the presented approach and results provide, however, a quantitative assessment of the impact of different factors, and thus information about the dominant processes of isotopic fractionation. It can support the interpretation of processes responsible for observed patterns of isotopic composition. The rather simple approach can, of course, not provide detailed information about atmospheric dynamics, but it provides a relatively simple and easy to apply approach supplementing or preceding more complex studies of isotopic composition with circulation models. Due to the simplicity, any scientist can easily apply this method in order to investigate factors controlling isotopic composition in precipitation at any given study area around the world without the requirement of setting up and in-depth knowledge about running a complex numerical atmospheric circulation model. Furthermore, the approach is easily reproducible and contains a rigorous quantitative analysis of the interplay of different driving factors. Moreover, the analysis can easily be extended to other factors and processes of importance in order to capture particularly the d-excess better, e.g. the sea surface temperatures at the source regions.”

Minor issues:

P2L17: what is “circulation effect”? Describe.

The term “circulation effect” (Tan, 2009;Tan, 2014) is used to describe the changes in isotopic composition in precipitation that appear because arriving moisture is coming from different areas of the ocean. The revised manuscript now includes this explanation.

P2L23: what is difference between “distillation during vapor transport” and “upstream rainout”. Aren’t they essentially the same?

Yes, thank you for pointing this out. We used only the term “distillation during vapor transport” in the revised manuscript.

P2L22-P3L3: Different temporal scales are mixed.

We sorted the references according to scale. The paragraph was revised as follows:

“Recently, many studies have presented evidence that large-scale monsoon circulation is the primary driver of variations in precipitation isotopes instead of local controls (e.g. local precipitation amount or temperature) in some parts of the Asian monsoon region. This evidence has been found at different temporal scales including daily isotopic variability (Yoshimura et al., 2003;Yoshimura et al., 2008), seasonal isotopic variability (Araguás-Araguás et al., 1998;Kurita et al., 2009;Dayem et al., 2010;Peng et al., 2010;Baker et al., 2015), and/or interannual isotopic variability (Vuille et al., 2005;LeGrande and Schmidt, 2009;Ishizaki et al., 2012;Tan, 2014;Kurita et al., 2015).”

P3L21: Before the authors’ conclusion, there are many studies which state necessity of consideration of multiple parameters.

Yes, the paragraph is misleading. We replaced the whole paragraph in the introduction with:

"It has been frequently stated and agreed to that local and regional factors should be considered simultaneously to explain the isotopic variation in rainfall (e.g. Johnson and Ingram, 2004). Hence, it can be hypothesized that using multiple factors in a single linear model is able to explain a larger share of the observed variance in isotopic composition. We aim at developing and testing a model-based statistical approach for the quantification of the contribution of isotopic separation processes for explaining the isotopic variation of precipitation. Such a model-based statistical method could also be applied in paleoclimate studies, separating and quantifying the impacts of local and regional factors on the isotopic composition of local precipitation (Sturm et al., 2010), thus overcoming the shortcomings of single factor analysis."

P3L27: For quantification of the controls, usually researchers try to develop a physical simulator. Any statistical model principally cannot explain the real control.

Physical models are one way to address this problem. But statistical models are an alternative way and have in fact be applied many times in all sorts of environmental studies. Both approaches have their advantages and disadvantages, and they coexist, respectively supplement each other. And while statistical models are not able to represent the actual process causing a phenomenon, they are able to detect results of a process. And this is what we actually are aiming at. Therefore we are arguing that the proposed model-based statistical approach is valid and accepted by the majority of researchers, as long as the limitations are clearly taken into consideration. We underlined this point in more detail in the introduction (P3L3-P4L33) of revised manuscript.

P4L20: There are many other definition of dry/wet season. What is the impact?

We included this paragraph in the revised manuscript (Section 2. Study area) to discuss the impact of the definition of dry/wet season.

"The definition used here is particularly developed for the local climatic conditions, the problem to be solved, and the data available. Other definitions could cause some data points to be assigned to the other season. However, those data points will most likely be from the transition period from one season to the other, i.e. other definitions would affect samples that have the least explanatory value for the actual dry and wet seasons."

P5L5: "three methods" are not really regarded as different "method".

Thank you for this point. "three methods" was changed to "three regression methods"

P6L4-L20: drop

As discussed in the 1st comment under 'major issues', we consider this part relevant and important for the manuscript. Therefore we would like to keep it.

P7L18: what is TRATIO?

We modified the sentence from P7L17-L19 in the old version of the manuscript as follows:

"Secondly, we use the shortest possible integration time step (i.e. 1 h) and a small value for the parameter TRATIO (0.25), which is the fraction of a grid cell that a trajectory is permitted to transit in one advection time step. Smaller values of TRATIO help to minimize the trajectory computation error using the HYSPLIT model".

P7L20: The uncertainty of trajectory analysis is not quantified. Perhaps it is minimized in the suggested framework, but how large is the "minimized" uncertainty and what is its potential consequence?

In order to discuss the sensitivity with regard to the choice of the gridded dataset as well as the uncertainties of the trajectory analysis, we included this paragraph to the methodology (Section 3.5) in the revised manuscript:

"Single backward trajectory computations by the HYSPLIT model can have large uncertainties. The horizontal uncertainty of the trajectory calculations by HYSPLIT has been estimated to be 10–20 % of the travel distance (Draxler and Hess, 1998). While errors in trajectory calculation computed from analyzed wind fields seem to be typical on the order of 20% of the distance travelled (Stohl, 1998), the statistical analysis of a large number of trajectories arriving at a study site would increase the accuracy of the trajectory analysis (Cabello et al., 2008). Harris et al. (2005) studied trajectory model sensitivity to the input meteorological data (focusing on ERA-40 and NCEP/NCAR reanalysis data) and to the vertical transport method. They pointed out five causes of trajectory uncertainty, expressed as percentage of deviation of the average travel distance: 1) minor differences in the computational methodology: 3–4%; 2) time interpolation: 9–25%; 3) vertical transport method: 18–34%; 4) meteorological input data: 30–40%; and 5) combined two-way differences in the vertical transport method and meteorological input data: 39–47%. However, it would be difficult to prove that in all situations a single meteorological data set or a single method of trajectory modeling was superior to another one (Gebhart et al., 2005;Harris et al., 2005). More details about the uncertainties in trajectory modeling were provided by (Stohl, 1998), later by (Fleming et al., 2012) and references therein."

P8L4: PRESS is essentially the same as root mean square error (RMSE), which is more popular in the community.

RMSE is calculated from the residuals of the model fitted to all data, while PRESS is based on the residuals resulting from the model fitted to all data except one, for which the residual is calculated. Repeating this for all data points and summing the calculated residuals results in PRESS. PRESS is therefore a cross validation method. See also our comment above, and for example the definition in WIKIPEDIA as reference (https://en.wikipedia.org/wiki/PRESS_statistic).

P8L5: what is "leave-one-out cross validation"? and what does it mean by "equivalent to" it?

See our reply to major comment 4 and to the previous comment.

P8L16: what is physical meaning of using "mean values of their combinations"? Combination of 800hPa and 850hPa represent 825hPa level (somehow the precipitation was formed at that level at that time)? In this regard, what is meaning of 800/850/900hPa combination?

In P7L4-L7 we discuss that the three levels at 1000, 1500, and 2000 m above ground are corresponding to barometric surfaces of approximately 900, 850, and 800 hPa. These barometric surfaces were chosen because the 850 hPa vorticity is highly indicative of the strength of the boundary layer moisture convergence and of rainfall in regions away from the equator (Wang et al., 2001). Hence rainfall is expected to mostly originate from these altitudes. We included this paragraph in the revised manuscript (in section 3.5) to elaborate on the physical meaning of using "mean values of their combinations" as follows:

"Consequently, the combination of 800 hPa and 850 hPa barometric surfaces accounts for the fact that rainfall is expected to mostly originate between 1500 and 2000 m above ground level. Correspondingly, the combination of the barometric surfaces of 800, 850 and 900 hPa means that rainfall is expected to mostly originate between 1000 and 2000 m above ground level."

P10L4-L27: drop

As discussed in the 1st comment under ‘major issues’, we argue that this part is relevant for the manuscript. We would like to keep it.

P11L23-L24: I don’t agree with this statement. More evidence is needed.

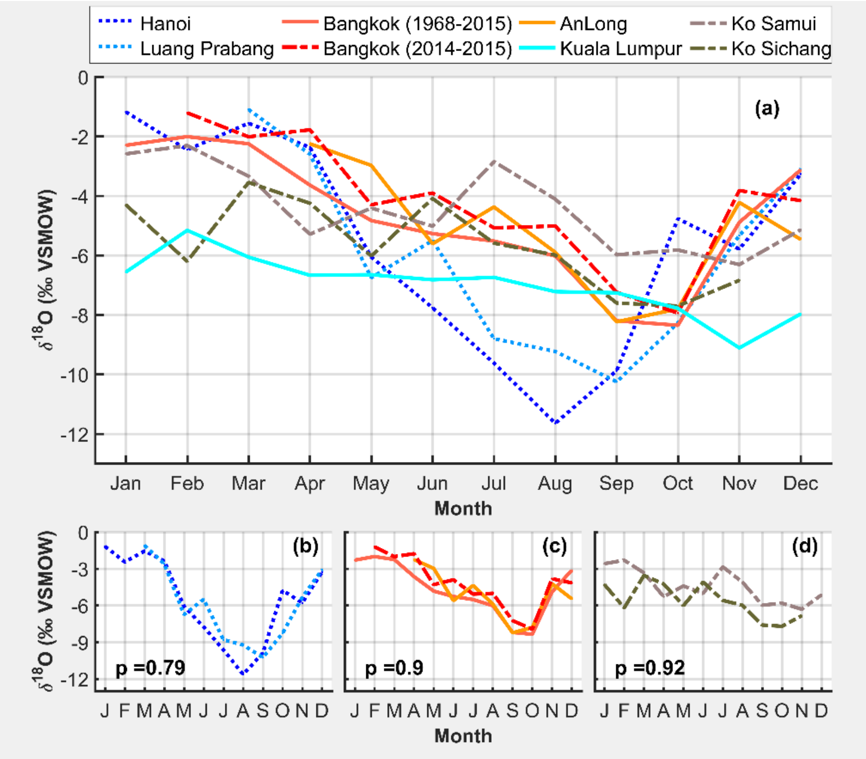
The Levene test (Levene, 1960) for equality of variances was used to compare the data of the different stations across the Indochinese Peninsula. We argue that the observed similarity of the isotopic values and their seasonal variances between An Long and the long term time series of Bangkok (Fig. 8c) (of which the visible similarity is also confirmed with high significance by the statistical Levene test) provides sufficient evidence for our statement. In order to substantiate this finding we added the time series of Bangkok covering the same time span as our data collected in the Mekong Delta to the analysis (new figure 8c, shown below). This time series is even more similar to the one of An Long, resulting in a highly significant Levene test statistic of 0.98. This means that the isotopic variation of the An Long time series is almost identical to the one from Bangkok, and that the variation of the short term time series of Bangkok and An Long is also very similar to the long term time series. In turn, one can infer from this that the data collected in An Long are likely to be representative for the area (i.e. the southern part of SE-Asia). This evidence was included in the revised manuscript (Section 4.2.2) as follows:

“In addition, the short-term time series of Bangkok and An Long (i.e. 2014-2015) show similar variances, resulting in a highly significant Levene test statistic of 0.98. The variation of the short-term time series of Bangkok and An Long is also very similar to the long-term time series, again shown by a highly significant Levene test statistic of 0.90 (Fig. 9c). This indicates that the isotopic variation of the An Long time series is almost identical to the one from Bangkok.”

We also modified the statement acknowledging the remaining uncertainty to:

“In summary, the analyzed GNIP data suggests that the data and results from this study are likely to be representative of the Southern continental part of the Indochinese Peninsula.”

Figure 8 (in the old version of the manuscript) was replaced by the following figure (Figure 9 in the revised manuscript), where the time series of Bangkok for the same period as our observation is added:



P14L9: Why was 124th model chosen as best?

Because the PRESS value of the 124th model is smallest. The sentence provides this information. We also stated this in the methodology section (P8L13) in the old version of the manuscript. In revised manuscript, this evidence is at P11L4.

P15L2: It is good idea. Why don’t you do this trial?

We actually did this. The result are shown in Figure 12 and discussed in section 4.4 (from P15L6 to P16L8) in the old version of the manuscript. In revised manuscript, the result are shown in Figure 13 and discussed in section 4.4 (from P18L6 to P19L7).

References

Cabello, M., Orza, J., Galiano, V., and Ruiz, G.: Influence of meteorological input data on backtrajectory cluster analysis? a seven-year study for southeastern Spain, Advances in Science and Research, 2, 65-70, 2008.

Conroy, J. L., Cobb, K. M., and Noone, D.: Comparison of precipitation isotope variability across the tropical Pacific in observations and SWING2 model simulations, Journal of Geophysical Research: Atmospheres, 118, 5867-5892, 2013.

Curio, J., and Scherer, D.: Seasonality and spatial variability of dynamic precipitation controls on the Tibetan Plateau, Earth System Dynamics, 7, 767, 2016.

Fleming, Z. L., Monks, P. S., and Manning, A. J.: Untangling the influence of air-mass history in interpreting observed atmospheric composition, Atmospheric Research, 104, 1-39, 2012.

Gebhart, K. A., Schichtel, B. A., and Barna, M. G.: Directional biases in back trajectories caused by model and input data, Journal of the Air & Waste Management Association, 55, 1649-1662, 2005.

Harris, J. M., Draxler, R. R., and Oltmans, S. J.: Trajectory model sensitivity to differences in input data and vertical transport method, Journal of Geophysical Research: Atmospheres, 110, 2005.

Ishizaki, Y., Yoshimura, K., Kanae, S., Kimoto, M., Kurita, N., and Oki, T.: Interannual variability of H218O in precipitation over the Asian monsoon region, Journal of Geophysical Research: Atmospheres, 117, 2012.

LeGrande, A. N., and Schmidt, G. A.: Sources of Holocene variability of oxygen isotopes in paleoclimate archives, Clim. Past, 5, 441-455, 10.5194/cp-5-441-2009, 2009.

Levene, H.: Robust tests for equality of variances, Contributions to probability and statistics, 1, 278-292, 1960.

Sánchez-Murillo, R., Birkel, C., Welsh, K., Esquivel-Hernández, G., Corrales-Salazar, J., Boll, J., Brooks, E., Roupsard, O., Sáenz-Rosales, O., and Katchan, I.: Key drivers controlling stable isotope variations in daily precipitation of Costa Rica: Caribbean Sea versus Eastern Pacific Ocean moisture sources, Quaternary Science Reviews, 131, 250-261, 2016.

Stohl, A.: Computation, accuracy and applications of trajectories—a review and bibliography, Atmospheric Environment, 32, 947-966, 1998.

Tan, M.: Circulation effect: climatic significance of the short term variability of the oxygen isotopes in stalagmites from monsoonal China—dialogue between paleoclimate records and modern climate research, Quaternary Sciences, 29, 851-862, 2009.

Tan, M.: Circulation effect: response of precipitation δ18O to the ENSO cycle in monsoon regions of China, Climate Dynamics, 42, 1067-1077, 2014.

Tindall, J., Valdes, P., and Sime, L. C.: Stable water isotopes in HadCM3: Isotopic signature of El Niño–Southern Oscillation and the tropical amount effect, Journal of Geophysical Research: Atmospheres, 114, 2009.

Vuille, M., Bradley, R., Werner, M., Healy, R., and Keimig, F.: Modeling δ18O in precipitation over the tropical Americas: 1. Interannual variability and climatic controls, Journal of Geophysical Research: Atmospheres, 108, 2003.

Vuille, M., Werner, M., Bradley, R., and Keimig, F.: Stable isotopes in precipitation in the Asian monsoon region, Journal of Geophysical Research: Atmospheres, 110, 2005.

Wang, B., Wu, R., and Lau, K.: Interannual variability of the Asian summer monsoon: contrasts between the Indian and the Western North Pacific-East Asian Monsoons*, Journal of climate, 14, 4073-4090, 2001.

What controls the stable isotope composition of precipitation in the ~~Asian monsoon region?~~Mekong Delta? A model-based statistical approach

Nguyen Le Duy[1, 2], Ingo Heidbüchel[1], Hanno Meyer[3], Bruno Merz[1, 4], Heiko Apel[1]

[1]GFZ German Research Centre for Geosciences, Section 5.4 – Hydrology, Potsdam, Germany
[2]SIWRR Southern Institute of Water Resources Research, Ho Chi Minh City, Vietnam
[3]Alfred Wegener Institute for Polar and Marine Research, Research Unit Potsdam, Germany
[4]University Potsdam, Institute of Earth and Environmental Science, Germany

Correspondence to: Nguyen Le Duy (duy@gfz-potsdam.de)

Abstract

This study analyzes the influence of local and regional climatic factors on the stable isotopic composition of rainfall in the Vietnamese Mekong Delta (VMD) as part of the Asian monsoon region. It is based on 1.5 years of weekly rainfall samples. ~~Their~~In a first step, the isotopic ~~content~~composition of the samples is analyzed by local meteoric water lines (LMWL) and single-factor ~~regressions~~linear correlations. Additionally, the contribution of several regional and local factors is quantified by multiple linear regression~~s~~ (MLR) of all possible factor combinations and by relative importance analysis~~, a novel~~. This approach is novel for the interpretation of isotopic records~~.~~ and enables an objective quantification of the explained variance in isotopic records for individual factors. In this study, the local factors are extracted from local climate records, while the regional factors are derived from atmospheric backward trajectories of water particles. The regional factors, i.e. precipitation, temperature, relative humidity and moving distance of the backward trajectories, are combined with equivalent local climatic parameters to ~~predict~~explain the response variables $\delta^{18}O$, $\delta^{2}H$, and d-excess of precipitation at the station of measurement.

The results indicate that (i) MLR can much better explain the isotopic variation of precipitation ($R^2 = 0.8$) compared to single-factor linear regression ($R^2 = 0.3$); (ii) the isotopic variation in precipitation is controlled dominantly by regional moisture regimes (~70%) compared to local climatic conditions (~30%); (iii) the most important climatic parameter during the ~~early~~ rainy season is the precipitation amount along the trajectories of air mass movement~~s~~; (iv) the influence of local precipitation amount and temperature is not significant during the early rainy season, unlike the regional precipitation amount effect; (v) secondary fractionation processes (e.g. sub-cloud evaporation) take place mainly in the dry season, either locally for $\delta^{18}O$ and $\delta^{2}H$, or along the air mass trajectories for d-excess. The analysis shows that regional and local factors vary in importance over the seasons~~,~~ and that the source regions and transport pathways, and in particular the climatic conditions along the pathways~~,~~ have a large influence on the isotopic composition of rainfall. While the general results have been reported qualitatively in previous studies (proving the validity of the approach), the proposed ~~methods thus proved to be valuable~~method provides

quantitative estimates of the controlling factors, both for the ~~interpretation of the~~ whole data set and for distinct seasons. Therefore it is argued that the approach constitutes an advancement in the statistical analysis of isotopic records in rainfall ~~and the factors controlling it.~~ that can supplement and precede more complex studies utilizing atmospheric models. Due to its relative simplicity, the method can be easily transferred to other regions, or extended with other factors.

The results illustrate that the interpretation of the isotopic composition ~~in~~of precipitation as a recorder of local climatic conditions, as for example performed for ~~paleo records~~paleorecords of water isotopes, may not be adequate in the Southern part of the Indochinese Peninsula, and likely ~~also not~~neither in other regions affected by monsoon processes. However, the presented approach could open a pathway towards better and seasonally differentiated reconstruction of paleoclimates based on isotopic records.

1 Introduction

The analysis of stable water isotopes ($\delta^{18}O$ and $\delta^{2}H$) and their use as tracers have become an effective tool in hydrology. They are widely used to characterize water resources in a given region and to understand dynamics of hydro-geo-ecological processes such as precipitation, groundwater recharge or groundwater-surface water interactions – from the plot to the catchment scale.

Precipitation is typically composed of regional contributions where atmospheric moisture has been transported over large distances and local contributions, where the moisture has been provided by evapotranspiration within the close vicinity. Understanding the sources of precipitation and their relative contribution is critical for basin-wide water balance studies (Ingraham, 1998). Stable isotopes offer the possibility to identify the sources of precipitation and to quantify the contribution of regional and local sources (Gat, 1996). Furthermore, they can be used to investigate hydrological processes such as mechanisms responsible for streamflow generation (e.g. Kendall and Caldwell, 1998), in groundwater studies (e.g. Gonfiantini et al., 1998) ~~or~~and in rainfall-runoff studies (e.g. Genereux and Hooper, 1998).

Isotopic variation in precipitation ~~at a given location~~ has been correlated with climatic parameters such as precipitation amount, air temperature, and air mass history (Dansgaard, 1964;Rozanski et al., 1992;Gat, 1996), termed amount effect, temperature effect (Dansgaard, 1964), and circulation effect (Tan, 2009;Tan, 2014), respectively. The circulation effect describes the changes in isotopic composition in precipitation that appear because arriving moisture is coming from different areas of the ocean.

Delineating the present-day relationship between climatic factors and stable isotop~~ic~~e variation in precipitation ~~may~~can also help to understand past climatic conditions at regional and global scales. However, the factors controlling isotopic variation of precipitation are numerous and complex; hence a better understanding of the climatic influences on isotopic values would improve the use of precipitation isotopes as a proxy to reconstruct paleoclimates (Yang et al., 2016).

In the Asian monsoon region, the isotopic ~~variation~~signature of precipitation has been found to correlate with ~~many other~~large-scale climatic parameters ~~and processes,~~ such as ~~distillation during vapor transport , the source of water vapor and changes in its temperature , upstream rainout ,~~ sea surface temperature and relative humidity of the air masses (Dansgaard, 1964;Merlivat and Jouzel, 1979;Clark and Fritz, 1997;Lachniet, 2009), ENSO (Ichiyanagi and Yamanaka, 2005;Tan, 2014;Yang et al., 2016)~~,~~ and the vertical wind shear index (Vuille et al., 2005)~~, and convective~~. Other relevant processes ~~such~~were identified as distillation during vapor transport (Araguás-Araguás et al., 1998;Yoshimura et al., 2003;Vuille et al., 2005;Dayem et al., 2010;Pausata et al., 2011;Lee et al., 2012;Liu et al., 2014) , re-evaporation and rain~~–~~-vapor interactions (Risi et al., 2008b;Chakraborty et al., 2016). ~~Recently, many studies have presented evidence that large-scale monsoon circulation is the primary driver of variations in precipitation isotopes instead of local precipitation amount in some parts of the Asian monsoon region . However, the influence of the different factors has been described qualitatively only, with the exception of the study of . Hence, a better quantitative understanding of the climatic influences on isotopic values over the Asian monsoon region is required.~~

Relations between climate and water isotopes have been analyzed by univariate statistical regression methods (e.g. Araguás-Araguás et al., 1998;Bowen, 2008), isotope-enabled global climate models (GCMs) (Yoshimura et al., 2008;Risi et al., 2010b;Yoshimura et al., 2014;Okazaki and Yoshimura, 2017), isotope-incorporated Lagrangian models (Pfahl and Wernli, 2008;Sodemann et al., 2008), or the combination of GCMs (or Lagrangian models) with statistical analysis (Vuille et al., 2003;Vuille et al., 2005;LeGrande and Schmidt, 2009;Tindall et al., 2009;Ishizaki et al., 2012;Conroy et al., 2013). While statistical models are not able to represent the actual process causing a phenomenon (e.g. the physical controls of isotope variations in precipitation), in contrast to physical models (e.g. GCMs or Lagrangian models), they can, however, detect the results of a process, and thus help to identify the responsible processes. Both approaches have their advantages and disadvantages and hence coexist supplementing each other. We argue that clearly taking into consideration the limitations and advantages of both statistical and physical models (discussed in next paragraphs) can enhance their power to interpret the relations between climate and water isotopes.

As illustrated in previous studies (e.g. Noone and Simmonds, 2002) and discussed in Sturm et al. (2010)~~Firstly, the interpretation of isotopic data using only local climatic factors, such as temperature or rainfall amount, can lead to incorrect paleoclimate reconstructions . For example, using $\delta^{18}O$ as a tracer of local precipitation is usually based on~~, the inherent limitations of empirical (or statistical) climate reconstructions from precipitation isotopes can lead to incorrect paleoclimate reconstructions. A major limitation is the assumption that the isotopic signal is controlled by a single climatic factor and that the stationary relationship (e.g. between temperature and $\delta^{18}O$) remains valid over the entire proxy record. This mono-factorial relationship does not consider the interplay of different climatic factors and is possibly biased. Another limitation is the assumption of a constant precipitation source or similar isotopic signatures of different moisture sources~~. However, both~~ throughout the study period when using only local parameters (e.g. local precipitation) to interpret precipitation isotopes. In real cases, these assumptions are rarely fulfilled and often unrealistic~~.~~ because of the changes in seasonality and atmospheric circulation patterns. This is particularly true in those parts of the Asian monsoon region located in the transition zone between the Indian and Western North-Pacific monsoon where precipitation originates from both the Indian and Pacific Oceans (Delgado et al., 2012a), with the isotopic signatures of air masses originating from the Indian Ocean differing considerably from those of the Pacific Ocean (Araguás-Araguás et al., 1998)~~.~~. Seasonally varying sources of precipitation have also been observed in China (Tan, 2014, and references therein), India (e.g. Breitenbach et al., 2010;Chakraborty et al., 2016), Korea (Lee et al., 2003), Thailand (Ishizaki et al., 2012), and elsewhere (Araguás-Araguás et al., 1998). ~~The isotopic signatures of air masses originating from the Indian Ocean differ considerably from those stemming from the Pacific Ocean, where the average δ^{18}O of the latter is about 2.5‰ more negative .~~

Since the pioneering work of Joussaume et al. (1984), GCMs have been frequently used for isotopic studies with at least a half-dozen GCMs (Risi et al., 2010b;Sturm et al., 2010). For a more detailed discussion about advances in the development of GCMs, the reader is referred to Galewsky et al. (2016) and references therein. Although GCMs could provide the physical links between climate and water isotopes (Yoshimura et al., 2008), the model parameterizations are still far from perfect due to downscaling issues and intrinsic atmospheric variability (Sturm et al., 2010). Modeling isotopic composition in precipitation by GCMs has some limitations stemming from the model uncertainties, e.g. the frequently reported biases in precipitation or temperature simulation (Mathieu et al., 2002;Lee et al., 2007;Yoshimura et al., 2008), and/or numerical inaccuracies in transport processes (Noone and Sturm, 2010). For example, the moist bias persisting in many GCMs in the tropical and subtropical middle and upper troposphere is due to excessively diffusive vertical advection (Risi et al., 2012). These limitations have obvious consequences (e.g. low correlation between simulated and observed δ^{18}O) for the simulation of isotopic variations in precipitation.

For paleoclimate reconstruction, the proxy data assimilation method has been proven to obtain adequate results (Yoshimura et al., 2014;Okazaki and Yoshimura, 2017). This approach, however, requires in-depth knowledge of the atmospheric modeling and/or data assimilation algorithm (Sturm et al., 2010). In any case, it takes a lot of effort to establish such a system if it is not already present. Particularly, even though the underlying physics is relatively simple, it would be a daunting task to develop a GCM source code which requires tens of thousands of code lines to simulate the hydrological cycle (Sturm et al., 2010). Generally, the complexity of GCMs impedes their interpretation.

While GCMs are typically Eulerian in the sense that mass is exchanged between fixed discrete volumes, Lagrangian models (e.g. the HYSPLIT model used in this study, mentioned in section 3.5) are used to calculate the composition of infinitesimal air parcels in the atmosphere according to the mean wind field data (Galewsky et al., 2016). The transport pathway along which air parcels travel is called a trajectory. In contrast to Eulerian models, Lagrangian models are not subject to numerical diffusion, hence they are computationally cheaper to simulate moisture sources. Because of their relative simplicity, these models are suitable to study the influences of different processes along transport trajectories (Helsen et al., 2006).They also more explicitly retain information about the history of the air parcels, which is useful to investigate controls on the isotopic composition of vapor arriving at a site of interest (Galewsky et al., 2016). In spite of their high suitability for exploring stable isotopes in paleoclimate reconstructions, using GCMs for simulating single meteorological events is more difficult due to their coarse spatial resolution (Pfahl and Wernli, 2008). Moreover, GCMs cannot capture the seasonal cycle of water isotopes on local scales (Angert et al., 2008). For these reasons, Lagrangian models are more suitable than GCMs to investigate controls on precipitation isotopes at a given location.

Although relationships between atmospheric circulation patterns and precipitation isotopes are frequently acknowledged and applied to reconstruct past climates, the actual causes of these relationships remain unclear (Ishizaki et al., 2012). Similarly, even if GCMs or Lagrangian models could provide much more detailed information about the fractionation processes along the transport pathways of water in the atmosphere, they cannot be used in a straightforward way to extract the impact of dominant factors and weight their relative importance for the variability of the observed isotopic signal. Statistical techniques are required to quantify the correlation between observed isotopic signal variability and regional climate change patterns (Sturm et al., 2010). Statistical analysis techniques such as principal component analysis (PCA) (Vuille et al., 2003;Curio and Scherer, 2016), sensitivity experiments (Ishizaki et al., 2012), or machine learning techniques like random forests (Sánchez-Murillo et al., 2016) have been used to investigate dominant factors ~~in~~ controlling isotopic composition in precipitation.

Recently, many studies have presented evidence that large-scale monsoon circulation is the primary driver of variations in precipitation isotopes instead of local controls (e.g. local precipitation amount or temperature) in some parts of the Asian monsoon region. This evidence has been found at different temporal scales including daily isotopic variability (Yoshimura et al., 2003;Yoshimura et al., 2008), seasonal isotopic variability (Araguás-Araguás et al., 1998;Kurita et al., 2009;Dayem et al., 2010;Peng et al., 2010;Baker et al., 2015), and/or interannual isotopic variability (Vuille et al., 2005;LeGrande and Schmidt, 2009;Ishizaki et al., 2012;Tan, 2014;Kurita et al., 2015). However, the influence of the different factors has been described qualitatively only, with the exception of the study of Ishizaki et al. (2012), in which the quantitative analysis of the controls is limited ~~the analysis~~ to two factors ~~only~~ (local precipitation amount and distillation of the moisture along its transport trajectories). That means, to our best knowledge, there is no study considering quantitatively the interplay of several local and regional factors.

~~Hence, we conclude~~ It has been frequently stated and agreed to that local and regional factors should be considered simultaneously to explain the isotopic variation ~~of precipitation~~ in rainfall (e.g. Johnson and Ingram, 2004). Hence, it can be hypothesized that using multiple factors in a single linear model ~~explains~~is able to explain a larger share of the observed variance in isotopic composition. We aim at developing and testing a model-based statistical approach for the quantification of the contribution of ~~regional and local factors controlling~~isotopic separation processes for explaining the isotopic variation of precipitation. Such a model-based statistical method could also be applied in paleoclimate studies, separating and quantifying the impacts of local and regional factors on the isotopic composition of local precipitation (Sturm et al., 2010), thus overcoming the shortcomings of ~~the regression factors, in combination with a regression~~single factor ~~importance~~ analysis. Th~~e~~is study uses the Vietnamese Mekong Delta (VMD) as a test case~~. During a field campaign, isotopic~~ , for which isotopic data in precipitation has been collected for the first time. The rainfall samples ($\delta^{18}O$ and $\delta^{2}H$) were collected ~~with~~ comparatively

~~high frequency~~frequently (bi-weekly to weekly) ~~f~~over a period of 18 months. This ~~data set~~dataset enables a better analysis of the temporal dynamics of the isotopic composition as compared to the typical~~ly~~ monthly Global Network of Isotopes in Precipitation (GNIP) data (IAEA/WMO, 2016). The collected data was used to characterize the isotopic composition for the Mekong Delta ~~for the first time. This was achieved by~~by means of local meteoric water lines, which were ~~set in relation~~compared to other locations in South-East Asia. The local meteoric water lines (LMWLs) developed in this study can be used as a baseline for other studies using isotopic data to investigate hydrological processes in the Mekong Delta. Furthermore, the data was used to test the proposed approach for the identification and quantification of the controls on ~~isotopic composition in rainfall in South-East Asia.~~the isotopic variation of precipitation.

The main objective of this study is to develop a model-based statistical approach that quantitatively estimates the relative contribution and the interplay of regional and local factors in controlling the isotopic variation of precipitation for a given study site. The proposed approach is based on backward trajectory analysis exploiting the benefits of a Lagrangian model (the HYSPLIT model mentioned in section 3.5), in combination with multiple linear regression (MLR) of all factor combinations specifically considering the widespread issue of multicollinearity of the regression factors, and relative importance analysis. The effort in this study is not meant to develop a universal model to predict precipitation isotopic composition, but rather to test a comparatively simple and transferable method utilizing easily obtainable atmospheric spatial and climatic information (trajectories) to quantitatively investigate the drivers and their interplay in controlling the isotopic variation of precipitation.

2 Study area

The study area, the Plain of Reeds (Fig. 1), is located in the northern part of the Vietnamese Mekong Delta (VMD), between latitudes 10°42’7"N to 10°48’9"N and longitudes 105°22’45"E to 105°33’54"E. With an area of 697,000 ha, it accounts for 17.7% of the total area of the VMD. About 95% of the Plain of Reeds is used for rice paddy and vegetable cultivating, and shrimp and fish farming (Hung et al., 2014). The average elevation ranges from 1-4 m above sea level.

Located in a tropical monsoon region, the climate of the VMD has a distinct seasonality with two seasons: the rainy season (May to November) resulting from the flow of moisture from the Indian Ocean and Western North-Pacific Ocean accounting for approximately 80-90% of the annual rainfall (Tri, 2012), and the dry season (December to April) controlled by high-pressure systems over the Asian continent (Wang et al., 2001). Precipitation from the Indian monsoon is forced by the convective heat sources over the Bay of Bengal (Wang et al., 2001) and arrives earlier than precipitation from the Western North-Pacific monsoon (Delgado et al., 2012), forced by a convective heat source over the South China Sea – Philippine Sea. The average annual rainfall is 1400-2200 mm, characterized by an uneven distribution, both spatially and temporally (Renaud and Kuenzer, 2012;GSO, 2014).

During the study period, i.e. the period of isotope sampling in rainfall lasting from June 2014 to December 2015, the rainy and dry seasons are defined by the monthly precipitation amounts and the monthly number of days with precipitation for Cao Lanh

(Fig. 2). The dry season is defined as the months with rainfall amount smaller than the overall average (blue line), and a monthly number of days with precipitation smaller than the overall average (red line). All other months are included in the rainy season. The definition used here is particularly developed for the local climatic conditions, the problem to be solved, and the data available. Other definitions could cause some data points to be assigned to the other season. However, those data points will most likely be from the transition period from one season to the other, i.e. other definitions would affect samples that have the least explanatory value for the actual dry and wet seasons.

The data indicates that the rainy season in 2014 lasted from May to November, and in 2015 from June to November. The dry season is thus defined from December 2014 to May 2015 and starts again in December 2015. The study period was very dry with an annual rainfall of 985 mm compared to the long-term average of 1550 mm at the station Cao Lanh. This anomaly needs to be considered in the interpretation of the results.

The annual average temperature is 27°C with the small interannual variability of about 1°C. Variation of temperature throughout the year is small with monthly averages of 25°C to 29°C (Fig. 3). The average annual relative humidity ranges from 82% to 85%, with a seasonal variation of 80% to 88% during the rainy season and 77% to 83% during the dry season (Fig. 3). The mean annual evaporation is 984 mm with a significant difference between the rainy season and the dry season. The monthly evaporation rate ranges from 67 to 80 mm and from 76 to 109 mm in the dry and rainy season, respectively. Daily sunshine duration is about 8.7 to 9.6 hours in the dry season and 5.5 to 5.9 hours in the rainy season (Renaud and Kuenzer, 2012;GSO, 2014).

3 Methodology

An overview of the proposed methodology is given in Fig. 4. For the derivation of local factors relevant for the isotopic composition of precipitation climate data from nearby meteorological stations were collected (section 3.1). At the test location, precipitation samples were analyzed for their isotopic composition (section 3.2 and 3.3). For the construction of local meteoric water lines (LMWL), three regression methods were applied, in order to test the robustness of the LMWL against different regression methods (section 3.4). The regional factors were derived from atmospheric back trajectory modeling (section 3.5). All possible combinations of local and regional predictors were included in multiple linear regressions, and their ability to explain the observed variance of the isotopic composition of precipitation was determined with performance statistics (MLR, section 3.6). Finally, the influence of the different factors on the explained variance of isotopic composition was determined by relative importance analysis (section 3.7).

3.1 Climatic and isotopic data collection

Daily precipitation, air temperature, and relative humidity were obtained from the National Centre for Hydro-Meteorological Forecasting (NCHMF) of Vietnam at two stations (Chau Doc, Cao Lanh, Fig. 1) for the period 2012-2015. Long-term (more than 30 years) climatic data at these stations was provided by SIWRP (2014). Precipitation isotopic data from six selected GNIP stations (IAEA/WMO, 2016) located in the Indochinese Peninsula (Fig. 1) was collected for comparison with the isotopic data sampled in this study in the Plain of Reeds.

3.2 Precipitation sampling at An Long

Precipitation at An Long in the Plain of Reeds (Fig. 1) was sampled on a weekly basis between June 2014 and May 2015 and twice a week between June 2015 and December 2015. The rain collector was a dip-in sampler type as described in the guidelines of the IAEA technical procedure for precipitation sampling (IAEA, 2014). It consists of a 5 L accumulation glass bottle fitted with a vertical 14 cm diameter plastic funnel that reaches almost to the bottom to prevent evaporative losses, and a pressure equilibration plastic tube (2 mm in diameter and 15 m in length) to minimize evaporation. All collected samples were stored in 30 mL plastic sample bottles with tight screw caps to avoid evaporation effects. Between collection and laboratory analysis, the samples were stored in the dark.

3.3 Isotopic laboratory analysis

All stable isotope samples were analyzed at the laboratory of the Alfred-Wegener-Institute (AWI) in Potsdam, Germany. The measurements were performed with a Finnigan MAT Delta-S mass spectrometer using equilibration techniques to determine the ratio of stable oxygen ($^{18}O/^{16}O$) and hydrogen ($^{2}H/^{1}H$) isotopes. Analytical results were reported as $\delta^{2}H$ and $\delta^{18}O$ (‰, relative to Vienna Standard Mean Ocean Water - VSMOW) with internal 1σ errors of better than 0.8‰ and 0.1‰ for $\delta^{2}H$ and $\delta^{18}O$, respectively. The measuring procedure is described in detail in Meyer et al. (2000). The deuterium excess (d-excess) was calculated following Eq. 1 (Dansgaard, 1964):

$$\text{d-excess} = \delta^{2}H - 8 * \delta^{18}O \tag{1}$$

3.4 Development of local meteoric water lines

For the development of local meteoric water lines (LMWL) three methods of linear correlation between $\delta^{18}O$ and $\delta^{2}H$ values were applied, in order to test the robustness of the LMWL against different regression methods:

1) ordinary least squares regression (OLSR),
2) reduced major axis (RMA) regression,
3) precipitation amount weighted least squares regression (PWLSR).

OLSR and RMA give equal weight to all data points regardless of their precipitation amount, while PWLSR minimizes the effect of smaller precipitation amounts (Hughes and Crawford, 2012), which are more likely to have a lower d-excess due to re-evaporation of raindrops below the cloud base (Jacob and Sonntag, 1991), or biases in the sampling method (Froehlich, 2001). OLSR tends to be more useful when investigating the interaction between hydro-climatic processes and stable isotope signatures in precipitation, whereas PWLSR is adequate in studying surface and groundwater hydrology (Hughes and

Crawford, 2012). For a more detailed discussion, the reader is referred to IAEA (1992); Hughes and Crawford (2012); Crawford et al. (2014).

The quality of fit of the three LMWLs resulting from OLSR, RMA, and PWLSR was evaluated based on the coefficient of determination R^2, also referred to as explained variance, the standard error SE and the statistical significance value (p-value). The regression model indicates a good fit to the data when R^2 is close to 1.0, the standard error is small in relation to the magnitude of the data, and the p-value is smaller than 0.0001 (Helsel and Hirsch, 2002).

3.5 Back trajectory modeling

The potential locations of atmospheric moisture sources and the direction of the air mass causing precipitation before reaching An Long station were investigated via back-trajectory analysis. This investigation was performed using the PC Windows-based HYSPLIT (Hybrid Single Particle Lagrangian Integrated Trajectory) model developed by NOAA (National Oceanic and Atmospheric Administration) at the Air Resources Laboratory (ARL) (www.arl.noaa.gov/HYSPLIT_info.php). The model builds on the Lagrangian approach, using a moving frame of reference for the advection and diffusion calculation as the air parcels move from their initial location (Draxler and Rolph, 2003;Stein et al., 2015). The model parameters and inputs are starting time and height of the trajectories, trajectory duration, vertical motion options, type of climatic dataset, and the number of trajectories. The back-trajectory outputs are the hourly locations of the trajectory segment endpoints, the altitude of trajectories, and climatic parameters (e.g. precipitation, temperature, relative humidity) along each trajectory.

The 1°x1° climatic dataset generated by the global data assimilation system (GDAS) was used as input to the HYSPLIT model. This dataset was downloaded from the ARL web server using the HYSPLIT graphical user interface. 10-day backward trajectory analysis was performed every 6 hours between 01-JUN-2014 and 31-DEC-2015 at the sampling site (10.72°N, 105.24°E) for three levels at 1000, 1500, and 2000 m above ground (corresponding to barometric surfaces of approximately 900, 850, and 800 hPa). These barometric surfaces were chosen because the 850-hPa vorticity is highly indicative of the strength of the boundary layer moisture convergence and of rainfall in regions away from the equator (Wang et al., 2001), hence rainfall is expected to mostly originate from these altitudes. Consequently, the combination of 800 hPa and 850 hPa barometric surfaces accounts for the fact that rainfall is expected to mostly originate between 1500 and 2000 m above ground level. Correspondingly, the combination of the barometric surfaces of 800, 850 and 900 hPa means that rainfall is expected to mostly originate between 1000 and 2000 m above ground level. In total, 6948 backward trajectories were computed. The HYSPLIT outputs, i.e. precipitation, temperature, relative humidity, and moving distance of moisture sources, were used to investigate the influence of the different moisture sources on the variation of the isotopic composition ~~in~~of precipitation at An Long. In order to derive figures representative for each trajectory, accumulated precipitation, mean values of temperature and humidity of the hourly HYSPLIT output were calculated along the trajectory and used as predictors in the MLR.

Single backward trajectory computations by the HYSPLIT model can have large uncertainties. The horizontal uncertainty of the trajectory calculations by HYSPLIT has been estimated to be 10–20 % of the travel distance (Draxler and Hess, 1998).

While errors in trajectory calculation computed from analyzed wind fields seem to be typical on the order of 20% of the distance travelled (Stohl, 1998), the statistical analysis of a large number of trajectories arriving at a study site would increase the accuracy of the trajectory analysis (Cabello et al., 2008). Harris et al. (2005) studied trajectory model sensitivity to the input meteorological data (focusing on ERA-40 and NCEP/NCAR reanalysis data) and to the vertical transport method. They pointed out five causes of trajectory uncertainty, expressed as percentage of deviation of the average travel distance: 1) minor differences in the computational methodology: 3–4%; 2) time interpolation: 9–25%; 3) vertical transport method: 18–34%; 4) meteorological input data: 30–40%; and 5) combined two-way differences in the vertical transport method and meteorological input data: 39–47%. However, it would be difficult to prove that in all situations a single meteorological data set or a single method of trajectory modeling was superior to another one (Gebhart et al., 2005;Harris et al., 2005)~~Backward trajectory computations by the HYSPLIT model can have large uncertainties~~. More details about the uncertainties in trajectory modeling were provided by (Stohl, 1998), later by (Fleming et al., 2012)~~.~~ and references therein.

In this study, several quality control measures were applied, as recommended in Stohl (1998), to increase confidence in the HYSPLIT-generated back trajectories and to improve the validity of the air mass history. Firstly, trajectories were computed for three pressure levels (900, 850, and 800 hPa). Similar origins of atmospheric moisture for these pressure levels suggest that resolution errors and atmospheric shearing instabilities are negligible which increases the confidence in the results. Secondly, we use the shortest possible integration time step (i.e. 1 h) and a small value for the parameter TRATIO (0.25), which ~~defines~~is the fraction of a grid cell that a trajectory is permitted to transit in one advection time step~~,~~. Smaller values of TRATIO help to minimize the trajectory computation error~~.~~ using the HYSPLIT model. Thirdly, the statistical analysis of a large number of trajectories (e.g. trajectory cluster analysis) arriving at the study site was applied to confirm the accuracy of the trajectory analysis. The trajectory cluster analysis is conducted by the HYSPLIT model to group trajectories with similar pathways. The cluster analysis merges these trajectories that are near each other and represents those clusters by their mean trajectory. Differences between trajectories within a cluster are minimized while differences between clusters are maximized. Computationally, trajectories are combined to decrease the number of clusters until the total spatial variance (TSV) starts to increase significantly. This occurs when disparate clusters are combined. This number of clusters is then selected as the optimal cluster number for sorting and combining similar trajectories. More information about the HYSPLIT cluster analysis can be found at https://ready.arl.noaa.gov/documents/Tutorial/html/.

3.6 Analysis of factors controlling isotopic variation in precipitation

Multiple linear regression (MLR) was used to assess how the isotopic variation in precipitation is related to regional and local controlling factors. As indicators of regional factors ~~we used~~ the output of the HYSPLIT model was used, consisting of the accumulated precipitation amount along the transport pathways (hereafter P_hysplit), mean temperature (T_hysplit) and mean relative humidity (H_hysplit) along the trajectory, and the distance of moisture sources travelled within the time frame of 10 days (D_hysplit). The local climatic factors are weekly precipitation amount (P_AL) at An Long station, and weekly mean air temperature (T_AL) and weekly mean relative humidity (H_AL) taken from the nearby Cao Lanh station during the sampling period. These seven predictors were related to isotopic values ($\delta^{18}O$, δ^2H, and d-excess) defined as response variables in the MLR. Pearson linear correlation coefficients were computed to show inter-correlations between response and predictor variables~~,~~ and then used to determine the importance of predictors in the MLR.

All possible subset regression models consisting of all possible combinations of predictors (2^7-1 = 127 models) were applied separately for $\delta^{18}O$, δ^2H and d-excess. The coefficient of determination R^2 for the MLR was calculated for each subset regression. The goodness of each MLR model was evaluated based on the Prediction Residual Error Sum of Squares (PRESS) (Eq. 2) and adjusted R^2 (Eq. 3) (Helsel and Hirsch, 2002). The PRESS residuals are defined as $e_{(i)} = y_i - \hat{y}_{(i)}$ where $\hat{y}_{(i)}$ is the regression estimate of y_i based on a regression equation computed leaving out the ith observation. The process is repeated for all n observations:

$$PRESS = \sum_{i=1}^{n} e_{(i)}{}^2 \qquad (2)$$

The selection of best models with PRESS is equivalent to a leave-one-out cross-validation, which tests the regression models for robustness and reduces the chances of model over-fitting, i.e. the chances of finding spurious regression models that provide good results for the given combination of factors and selected time period only.

The adjusted R^2 (R_a^2) is defined as follows

$$R_a^2 = R^2 - (1 - R^2)\frac{p}{(n-p-1)} \qquad (3)$$

Where p is the total number of predictors in the MLR model and n is the number of observations. The statistical significance of all linear regression~~s~~ was evaluated based on the p-value for the F-test as part of a one-way ANOVA analysis. A good MLR model is hereby characterized by:

(i) PRESS close to zero,

(ii) Adjusted R^2 (R^2_a) close to 1.0,

(iii) a p-value smaller than 0.0001.

For each response variable, six pressure layers (800 hPa, 850 hPa, 900 hPa, and mean values of their combinations) and 10 durations of backward trajectories (from 1-day to 10-day backward) were used. The different pressure levels and combinations were chosen in order to tackle the inherent uncertainty regarding the pressure levels from which the rainfall actually stems. Similarly, different durations of the trajectories were chosen in order to avoid fixing the a-priori unknown travel time of precipitation reaching An Long. Overall, this resulted in 7620 MLR models for each response variable $\delta^{18}O$, δ^2H and d-excess (6 pressure levels times 10 trajectory durations times 127 predictor sets). The best MLR model was then identified by the smallest PRESS value (Eq. 2). Furthermore, the goodness of fit of the MLR models was characterized based on the adjusted R^2 values.

3.7 Relative importance analysis

Relative importance analysis determines the proportion of the variance explained by the individual predictors in the regression. However, this is difficult when predictors are correlated, since multicollinearity can lead to a high sensitivity of regression coefficients caused by small changes in the model. This means that the importance can strongly shift from one predictor to another well correlated one if the data set is changed even only slightly. The leave-one-out cross-validation may be particularly vulnerable to this effect. Therefore two methods were applied, namely relative weight analysis (Johnson, 2000) ~~and~~which has been developed to quantify the power of predictors when they are correlated, and the relative partial sum of squares (Gardner and Trabalka, 1985)~~, which have both been developed to quantify the power of predictors when they are correlated.~~. For a review of approaches to estimate predictor importance, readers are referred to Tonidandel and LeBreton (2011); Kraha et al. (2012).

Relative weight analysis approximates the relative importance of a set of predictors by creating a set of variables that are highly related to the original set of variables but are uncorrelated with each other. The response variable is then regressed on the uncorrelated set of predictors to approximate the relative weight of the original set of predictors, defined as the relative contribution of each predictor to R^2. This method is computationally efficient even for a large number of predictors and produces very similar results compared to more complex methods. Details are given in Johnson (2000); Tonidandel et al. (2009).

In the relative partial sum of squares (RPSS) method (Gardner and Trabalka, 1985), the total sum of squares of the response variable is partitioned based on multiple linear regression between all predictors. Briefly, the RPSS represents the percentage of the total sum of squares attributable to each of the predictors. To calculate RPSS for predictor V_i, the difference between the regression sum of squares (RSS) for the full model and the regression sum of squares for the model with V_i missing (RSS-i) is divided by the total sum of squares (TSS) (Rose et al., 1991), and expressed as a percentage using Eq. 4.

$$RPSS = 100 * (RSS - RSS_{-i})/TSS \quad (4)$$

The relative importance derived by the methods above quantifies the proportion of the variance explained by the individual regression factors, and thus identifies the dominant controls on the isotopic composition of rainfall.

4 Results and discussion

4.1 Variability of moisture sources

Because there is no daily precipitation data recorded at An Long, we used daily precipitation data at Cao Lanh instead. This is the closest national meteorological station, located approximately 37 km Southeast of An Long. Backtracking trajectories in Fig. 5 are plotted for the days when rainfall was recorded at Cao Lanh. This is based on assumption that days with precipitation at Cao Lanh and An Long coincide.

Figure 5 shows back-calculated trajectories of atmospheric moisture prior to rainy days at An Long for the sampling period from June 2014 to December 2015. Left and right panels show the results of 850 hPa trajectories for 2014 and 2015, and the upper, middle, and lower panels show the results for the early (June – September) and late (October – November) rainy season and dry season (December – May), respectively. Figure 6 shows the spatial distribution of vapour trajectories (cluster means) for precipitation days at An Long for 3 barometric surfaces (800, 850, 900 hPa) between June 2014 and December 2015, and the change in total spatial variance (TVS) for different cluster numbers. The TSV was used to identify the optimum number of clusters. The similarity of back-calculated trajectories (Fig. 5) and trajectory cluster analysis (Fig. 6) at three barometric surfaces (900, 850, and 800 hPa~~) (not shown~~) illustrates that the trajectories and thus the source regions do not differ between different atmospheric layers. This indicates a barotropic atmosphere, with the consequence that it is unlikely that the selection of the pressure layer for the HYSPLIT trajectories modifies the results of the MLR significantly.

Figure 5 and Figure 6 ~~demonstrates~~ demonstrate that the dry-season precipitation (from December to May) in the Plain of Reeds mainly originates from the moisture sources of the Asian continental air masses and the oceanic air masses carried by the equatorial easterlies, whereas during the rainy season (from June to November) air masses travel a longer distance over the tropical Indian Ocean (from June to September) and the South Pacific Ocean (October to November).

These findings for An Long agree with the general characterization of monsoonal circulation and precipitation over the Southeast Asia region, with moisture from the Indian Ocean dominating during the initial stage of monsoon evolution, and the Pacific Ocean dominating in the later stages. This indicates that the HYSPLIT model provides valid trajectories to be used in the MLR.

The mean $\delta^{18}O$ values for the 5 clusters are plotted in Figure 6 (in brown). The mean cluster values are similar for the three pressure levels. Also, the mean values of the two clusters from the Indian Ocean, as well as the two clusters from the Pacific, are similar. For a fingerprinting one also has to consider the variation of the values within the clusters, which partly overlap. This means that the $\delta^{18}O$ values of precipitation in the Mekong Delta cannot be used to uniquely identify the origin of the trajectory. However, they provide a coarse indication of their origin.

4.2 Isotopic composition of precipitation

4.2.1 Meteoric water lines

The linear-regression analyses of 74 pairs of $\delta^{18}O$ and δ^2H values at An Long yield LMWLs for the Plain of Reeds as follows:

1) Ordinary least squares regression (OLSR):

$\delta^2H = (7.56 \pm 0.11) * \delta^{18}O + (7.26 \pm 0.67)$

(SE = 2.26; $r^2 = 0.99$; $p < 0.0001$; n = 74),

2) Reduced major axis regression (RMA):

$\delta^2H = (7.61 \pm 0.11) * \delta^{18}O + (7.58 \pm 0.68)$

(SE = 2.27; $r^2 = 0.99$; $p < 0.0001$; n = 74),

3) Precipitation amount weighted least squares regression (PWLSR):

$\delta^2H = (7.61 \pm 0.11) * \delta^{18}O + (7.87 \pm 0.73)$

($SE = 2.29$; $r^2 = 0.99$; $p < 0.0001$; $n = 74$).

The numbers in brackets indicate the estimates of slope and intercept plus/minus the standard deviation, indicating the parameter uncertainty.

The close fit of all considered regressions indicates a very good linear relationship between $\delta^{18}O$ and δ^2H in the study area that is independent of the applied regression method. On a global scale, a good linear relationship between $\delta^{18}O$ and δ^2H is usually observed at sites where secondary fractionation processes, e.g. sub-cloud evaporation, are insignificant (Crawford et al., 2014). The LMWL for An Long is slightly different from the global meteoric water line (GMWL; defined by $\delta^2H = 8*\delta^{18}O + 10$ using OLSR, (Craig, 1961) and the LMWLs derived for six selected GNIP stations (IAEA/WMO, 2016) located in the Indochinese Peninsula (Fig. ~~6~~7). The small difference in slope between these LMWLs compared to that of GMWL, and the distribution of isotope values along the GMWL indicate that evaporative isotopic enrichment during rainfall is not significant. However, the less positive intercepts of LMWLs (<10‰) (Fig. ~~6~~7) may reflect smaller kinetic effects during evaporation (Ingraham, 1998) over the Mekong Delta compared to the worldwide average.

4.2.2 Seasonal variation and spatial homogeneity

The 74 precipitation samples at An Long showed that $\delta^{18}O$ ranges between -12.6‰ and -1.0‰, with an arithmetic mean value and standard deviation of -5.8‰ ± 2.5‰, and δ^2H ranges between -89.3‰ and 0.9‰, with an arithmetic mean value and standard deviation of -36.2‰ ± 18.7‰. Generally, less negative isotopic values are observed in the dry-season precipitation samples. The most negative values occur in the second half of the rainy season (September and October), whereas the least negative values are observed ~~at~~in the late dry season in April and May (Fig. ~~6~~7 and Fig. ~~7~~8). This shows that the isotopic composition of precipitation at An Long station exhibits marked seasonal variations, which in turn indicates different dominant moisture sources and/or processes in the different seasons. A comparison of the seasonal variation of $\delta^{18}O$ with the short-term (2014-2015) and long-term (1968-2015) monthly averages of Bangkok (Fig. ~~7~~8) reveals very similar seasonality, both in terms of timing and magnitude. The differences between $\delta^{18}O$ for An Long and Bangkok are likely caused by the exceptional low rainfall in the study period compared to the long-term monthly values, particularly during May and July. Considering additionally the similarity of general factors controlling stable isotopic composition of precipitation between the two stations, i.e. annual rainfall amount, air temperature, altitude and latitude (Dansgaard, 1964;Ingraham, 1998), it can be concluded that the isotopic variations of An Long and Bangkok follow the same dynamics and controls, both on an annual and seasonal scale, and can represent or complement each other.

In order to test the representativeness of the An Long data for a wider area, the variability of the monthly mean $\delta^{18}O$ data of An Long was compared to the available GNIP data of the Indochinese Peninsula (Table 1). The Levene test (Levene, 1960) for equality of variances was used to compare the data of the different stations. As shown in Fig. ~~8~~9, the test results in four distinct groups of data series with similar variances: the Northern part of the Indochinese Peninsula (Hanoi and Luang Prabang, Fig. ~~8b~~9b), the Southern part of the Indochinese Peninsula (Bangkok and An Long, Fig. ~~8c~~9c), the islands in the Gulf of Thailand (Ko Samui and Ko Sichang, Fig. ~~8d~~9d), and finally Kuala Lumpur showing only little seasonal variability. The Northern and Southern parts of the Indochinese Peninsula show generally a similar seasonal behavior with a distinct higher depletion during the rainy season, ~~still~~but in the Northern part the highest depletion is one month earlier (August)~~ in the Northern part~~ than in the Southern part, and the magnitude of the depletion is larger. The seasonal $\delta^{18}O$ variability in precipitation on the islands is much lower than on the stations located on the continent. This is likely due to the maritime setting and could indicate a continental effect. In addition, the short-term time series of Bangkok and An Long (i.e. 2014-2015) show similar variances, resulting in a highly significant Levene test statistic of 0.98. The variation of the short-term time series of Bangkok and An Long is also very similar to the long-term time series, again shown by a highly significant Levene test statistic of 0.90 (Fig. 9c). This indicates that the isotopic variation of the An Long time series is almost identical to the one from Bangkok. In summary, the analyzed GNIP data suggests that the data and results from this study are likely to be representative ~~for~~of the Southern continental part of the Indochinese Peninsula.

4.3 Factors controlling isotopic composition of precipitation

Prior to the MLR, the correlation of the predictors was analyzed (Table 2). The absolute values of the correlation coefficients between local (P_AL, T_AL, H_AL) and regional (P_hysplit, T_hysplit, H_hysplit, D_hysplit) climatic parameters are relatively small and mostly not significant ($|r| < 0.4$, Table 2b). However, the correlation coefficients between regional predictors are in most cases high and significant (Table 2c). Highest correlations are found between temperature and humidity for local factors, and between the regional humidity and precipitation for regional factors. Interestingly, the correlation between P_AL and H_AL is quite low. This indicates that the local precipitation is mainly controlled by large-scale circulation. The correlation between the predictors underlines the necessity to consider multicollinearity when investigating how the predictors control the response variables ($\delta^{18}O$ and $\delta^{2}H$).

4.3.1 Local factors and isotopic composition in precipitation

Typically, in tropical regions subject to a monsoon climate the correlation between $\delta^{18}O$ and $\delta^{2}H$ values of precipitation and air temperature is virtually nonexistent, whereas a strong relation between $\delta^{18}O$ and amount of precipitation has been observed (Rozanski et al., 1992;Araguás-Araguás et al., 1998). Our data show that the correlation of local precipitation amount (P_AL) and local temperature (T_AL) with isotopic values ($\delta^{18}O$ and $\delta^{2}H$) are both low ($|r| < 0.45$, Table 2a). This suggests that $\delta^{18}O$ and $\delta^{2}H$ variation is neither dominated by local precipitation amount nor by local temperature during the sampling period. This lack of a significant correlation ($|r| < 0.5$) between $\delta^{18}O$ and local rainfall amount was also observed in other regions affected by the Asian monsoon climate such as Bangkok, Hong Kong, New Delhi (Ishizaki et al., 2012), and Cherrapunji, India

(Breitenbach et al., 2010). This again supports the statement that $\delta^{18}O$ may not be an adequate proxy for local climatic conditions (e.g. temperature or rainfall amount) in the Asian monsoon region (Aggarwal et al., 2004;Vuille et al., 2005).

Secondary fractionation processes such as sub-cloud evaporation or secondary evaporation from open water bodies tend to decrease d-excess in the residual rainwater (Stewart, 1975) and enrich it in the heavy isotopes (Guan et al., 2013). The negative correlation of humidity (H_AL) with $\delta^{18}O$ and $\delta^{2}H$ (r = -0.53, Table 2a) combined with a positive correlation with d-excess (r = 0.2, Table 2a), indicates that some secondary fractionation processes (Risi et al., 2008b;Crawford et al., 2017) may take place during some months at An Long. To examine in which month secondary fractionation processes are likely significant, amount-weighted mean and arithmetic mean, for both $\delta^{18}O$ and d-excess are compared. The rationale is that if secondary fractionation processes are important (with the assumption that the moisture sources of different events within the month are the same), the arithmetic mean should have a $\delta^{18}O$ value that is more enriched in heavy isotopes, and a much smaller d-excess than the weighted mean (Guan et al., 2013). Figure 10 shows that secondary fractionation processes may take place during the dry season, in December 2014, and in April and May 2015, because in these months a) less negative $\delta^{18}O$ values and lower d-excess values compared to the overall arithmetic mean are observed, while at the same time the monthly arithmetic means are higher for $\delta^{18}O$, and lower for d-excess compared to the monthly weighted means.

To further corroborate this finding, linear regression ~~were~~was performed for different seasons to derive seasonal LMWL's and relations between local humidity and $\delta^{18}O$ and d-excess. Table 3 suggests that secondary fractionation processes are likely to take place in the dry season between December 2014 and May 2015. This is depicted by a slope of lower than 8 (slope = 6.9) for the dry season, the slightly negative correlation between $\delta^{18}O$ and local relative humidity, and the markedly positive correlation between humidity and d-excess. This is a distinctly different behavior compared to the rainy season as a whole, but also for the first (early monsoon) and second (late monsoon) parts of the rainy season. In summary, these findings indicate that secondary fractionation processes influence the isotopic composition of precipitation primarily in the dry season, which is characterized by lower humidity and higher temperature in the Plain of Reeds. While this conclusion is plausible due to the climatic conditions and low rainfall amounts, one has to consider the low number of rainfall samples during the dry season, which associates some uncertainty to the regressions and thus the interpretation.

4.3.2 Regional factors and isotopic composition in precipitation

In comparison to other regional and local parameters, the precipitation amount along the transport pathways of moisture sources (P_hysplit) shows the strongest correlation with $\delta^{18}O$ and $\delta^{2}H$ as depicted by a correlation coefficient of -0.76 (Table 2a). Thus, P_hysplit is likely the dominant factor controlling the isotopic composition of precipitation. Other predictors show weaker correlations with $|r| < 0.55$. This, however, does not exclude that these predictors do have some predictive power for the isotopic composition of precipitation in An Long when used in combination with other predictors. Although $\delta^{18}O$ and $\delta^{2}H$ are rather well correlated with some climatic parameters, d-excess (which is a function of both) is not well correlated. This is because of the relative difference of the variation of $\delta^{18}O$ and d-excess, which is expressed by a low correlation coefficient between two these variables (r = -0.44). The weak correlation between d-excess and all climatic parameters ($|r|<0.36$) indicates that the selected predictors (i.e. selected climatic parameters) are not sufficient to explain the processes responsible for the variability of the d-excess. On a global scale, drivers controlling d-excess variation are likely sea surface temperature or near-surface relative humidity of moisture sources (Pfahl and Wernli, 2008;Uemura et al., 2008;Pfahl and Sodemann, 2014), which are not considered in this study. In tropical areas, a major contribution to the seasonal variation in d-excess can be convective processes, e.g. re-evaporation and rain–vapor interactions (Risi et al., 2008a;Risi et al., 2010a), or the influence of large-scale processes, e.g. conditions at the vapor source, convection and recycling of moisture along trajectories (Landais et al., 2010). A complete investigation of factors controlling d-excess in precipitation is thus not possible by the presented study design and selected predictors. However, some conclusions about the factors controlling the d-excess can be obtained with the presented method, see below.

4.4 MLR and relative importance analysis

The results of the MLR indicate that $\delta^{18}O$ signal in precipitation at An Long is best explained by moisture sources of 5-day backward trajectories (Fig. ~~10~~11). The MLR of these trajectories produces the lowest PRESS and highest R^2 values, indicating that about 80% of the variability of precipitation $\delta^{18}O$ (Fig. ~~10~~11) and δ^2H (not shown) at An Long can be explained by the best MLR model. The explained variance differs only slightly between the different pressure levels used. But still, the best performance in terms of the lowest PRESS value was obtained by the mean backward trajectories of the 800 hPa and 850h Pa levels.

Contrary to $\delta^{18}O$ and δ^2H, the MLR fails to explain the variation of d-excess over the whole study period to a large extent, with a maximal R^2 of 0.3 (Fig. ~~10~~11). This indicates that the climatic parameters used in our MLR models have only little impact on the annual d-excess variation, which corroborates the findings of the linear correlation analysis in section 4.3.2.

In the next step, the importance of the MLR predictors was analyzed. Figure 12 shows the results applying Johnson's relative weight analysis for the best performing MLR models, i.e. using the mean of the 800 hPa and 850 hPa 5-day backward trajectories. In general, the predictive power of the MLR models increases with increasing number of predictors. Both importance methods, i.e. relative weight analysis and the relative partial sum of squares, yield very similar results (not shown). The results indicate that regional factors are always more important than local factors if the R^2 value is above 0.5. The local factors dominate only in MLR models with low performance, or when no regional factors are used as predictors. This is also highlighted by the sum ratio line (black line in Fig. ~~11~~12), defined as the fraction of R^2 explained by regional factors normalized to the overall R^2. In the best MLR model (124th model) with the lowest PRESS value and an R^2 of 0.80, which is equivalent to an explained variance of 80%, the regional factors explain 56% of the absolute $\delta^{18}O$ variance (which is equivalent to 70% relative to $R^2 = 0.80$), while local factors explain only 24% (30% relative to $R^2 = 0.80$). This result agrees with the two-factor analysis of Ishizaki et al. (2012) who stated that distillation during transport from source regions is the dominant contributor to inter-annual variability of $\delta^{18}O$ precipitation in Bangkok, Bombay, and Hong Kong, accounting for 70%, 60% and 70%

relative to the overall explained variance, while the amount of local precipitation contributed the remaining 27%, 33%, and 25% of the explained variance, respectively.

In all models where precipitation amount along transport pathways from moisture source regions (P_hysplit) is included, this factor explains the highest proportion of R^2, which is always at least double and up to triple of the explained variance of other factors (Fig. ~~11~~12). In turn, the absence of P_hysplit as a predictor in the MLR model considerably decreases the R^2, indicating that P_hysplit is the most dominant factor. In the best MLR model (124^{th} model) the most important predictor is P_hysplit, explaining 47% of the total $\delta^{18}O$ variance (partial $R^2 = 0.47$, Fig. ~~11~~12). The second dominant factor is T_AL, accounting for 21% of the explained the total variance. The remaining factors account for less than 13% of the $\delta^{18}O$ variance. This result indicates that the regional amount effect is a dominant process in controlling isotopic variation, whereas the local amount effect is not important in the VMD. Similar findings are reported for other regions in Asia (e.g. Rozanski et al., 1992;Araguás-Araguás et al., 1998). The local ~~humidity H~~temperature T_AL, however, can be regarded as a modulating factor for the isotopic composition on top of P_hysplit.

In a next step, the predictor importance analysis is performed for different seasons, in order to analyze if seasonal differences in the dominating factors for the isotopic composition exist, as the correlation analysis of local factors and isotopic composition suggests (section 4.3.1). The samples were split into dry season and rainy season subsets, for which the MLR was applied individually. The definition of the seasons follows the analysis in section 2, i.e. the dry season lasts from December to May. However, due to the low number of samples during this period, the dry season samples were taken from mid-November to mid-June in order to increase the sample number, thus enabling a more robust MLR fitting. This selection can be justified: Because the delineation of the dry and wet season above is based on monthly data, the "sharp" distinction between the rainy and dry season is forced by the temporal resolution ~~in~~of the data used. In reality, the transition between rainy and dry season is rather gradual, thus the delineation between the rainy and dry season should rather be regarded as fuzzy. Using data from the last two weeks of November and the first two weeks of June can be seen as one way to consider this.

Furthermore, the rainy season was subdivided according to the different moisture source regions shown in section 4.1: the Indian Ocean, dominating during the initial and high stage of the Indian monsoon from June to September/mid-October, and the South China Sea – Philippine Sea and the North-West Pacific Ocean from October to May during the late rainy and dry seasons, with some contribution from continental Asia (Fig. 5). In order to test if the factors have different importance caused by different source regions during the rainy season, the MLR models and relative importance analysis were applied for these two time periods in addition to the dry season. The number of samples for the different subsets was 42, 18 and 14 for the early rainy season, late rainy season and dry season, respectively.

Figure 13 shows the results of the MLR and importance analysis for the three seasonal subsets for $\delta^{18}O$. The sorting of the models is the same as in Fig. ~~11~~12. On a first glance, the results for the rainy season subsets (Fig. 1~~2~~3a and Fig. 1~~2~~3b) are quite similar to each other and to the overall data set. The best performing model in terms of the lowest PRESS value is in all cases the model 124. However, in terms of R^2, the performance of the early rainy season is somewhat lower compared to the overall data set, while for the late rainy season it is significantly better, with $R^2 = 0.96$. This increase in explained variance is caused by an increased contribution of the regional factors. In the late rainy season, the regional factors alone contribute 76% to the overall R^2 of 0.96 of the best PRESS model, which equals 79% of the explained variance (Table 4).

This is a much larger contribution compared to the partial R^2 values of 56% and 51% for the whole data set and the early rainy season, respectively. The increase stems from a larger importance of the other regional factors H_hysplit and/or T_hysplit. While their contribution to the whole data set and the early rainy period is rather low and P_hysplit dominates the contribution of the regional factors, it is raised to about 30% in the late rainy season, either individually or in combination. For the best PRESS model marked with the cyan dot in Fig. 1~~2~~3b, T_hysplit contributes 27% to the overall R^2 of 0.96. It indicates that temperature and humidity play a larger role in the isotopic fractionation along the trajectories of water stemming from the North-West Pacific/South China Sea and continental Asia compared to water stemming from the Indian Ocean during the boreal summer months. The large regional and thus climatic heterogeneity of water sources during the late rainy season offers a plausible explanation for this result. The source regions during this period are located in oceans and continental regions in higher latitudes outside the tropics, where large climatic differences may occur during the transport along the trajectories. Therefore, fractionation processes caused not only by the rainfall amount, but also by evaporation and condensation are likely to have a larger effect on the final isotopic composition of rainfall reaching An Long during this period, as compared to the low climatic variability of the tropical Indian Ocean region, where the rainfall during the early rainy seasons originates.

A completely different picture reveals the MLR fitting and importance analysis for the dry seasons shown in Fig. 1~~2~~3c. While the overall performance in terms of R^2 is comparable to the early rainy season, the importance of the local and regional factors is very different from the other seasons. For the dry season, the local factors dominate. In the best performing MLR model with the lowest PRESS value (cyan dot in Fig. 1~~2~~3c), T_AL contributes 78% of the explained variance. Similar results are obtained for almost all of the MLR models. For the models with $R^2 > 0.5$, T_AL is the most important factor, followed by P_AL and H_AL with similar importance. The regional factors generally do not contribute more than 22% of the explained variance, if $R^2 > 0.6$. This finding corroborates the assumed higher importance of secondary fractionation processes during the dry season in the VMD, as already hypothesized in section 4.3.1. However, in combination with other predictors, T_AL seems to be a better predictor of the secondary fractionation processes compared to H_AL, which was used in 4.3.1. As T_AL and H_AL are closely correlated (Table 2), the findings of section 4.3.1 and the MLR of the dry season presented in this section agree well.

~~The results for~~ The MLR modeling of δ^2H shows very similar results to $\delta^{18}O$ leading to the same conclusions (Fig. S1 and Fig. S2 ~~-~~in the supplement). The ~~seasonal~~MLR modeling of seasonal d-excess also shows an improved ~~the MLR~~ fit for the late rainy and dry seasons (Fig. S3 and Fig. S4 ~~-~~in the supplement), while for the early rainy season the results are not as satisfying as for the whole dataset. In contrast to $\delta^{18}O$ and δ^2H~~,~~ regional factors explain the bulk of the d-excess variance also for the dry season. Among the regional factors~~,~~ P_hysplit has the lowest importance for d-excess, while the others factors T_hysplit, H_hysplit, and D_hysplit explain about 65% of the best R^2 of 0.66. This is also a distinctively different result compared to

$\delta^{18}O$ and $\delta^{2}H$, where P_hysplit always dominated the regional factor contribution. The remaining explained variance ~~is~~stems mainly ~~explained by~~from the local precipitation P_AL, with some contribution of T_AL. This finding is in line with the rationale outlined in section 4.3.1, that evaporation along the transport pathway decreases the d-excess (Stewart, 1975). This effect is much more variable during the late rainy and dry season due to the transport pathways from higher latitudes, as compared to the rather uniform climatic conditions along the transport pathways during the rainy season, as already argued in the previous paragraph for the late rainy season results of $\delta^{18}O$. This means in summary that the MLR and relative importance analysis of d-excess for the late rainy and dry season corroborate the hypothesis that secondary fractionation processes caused by evaporation are relevant during the dry season, respectively for rainfall stemming from the Pacific region and continental Asia. However, for $\delta^{18}O$ and $\delta^{18}H$ local factors describing evaporation are more important, while for d-excess regional factors and thus evaporation processes along the transport pathways dominate.

Overall, applying all possible subset regression MLR models can much better explain the isotopic variation in rainfall compared to approaches considering only one predictor, i.e. a simple correlation analysis. Moreover, the associated relative importance analysis enables the identification of the dominant factors, thus offering interpretation aids for the identification of the processes responsible for the isotopic signature of ~~the~~ local rainfall. The presented analysis illustrates~~,~~ that ~~investigating~~the investigation of dominant factors controlling isotopic composition in precipitation with simple correlation analyses may lead to wrong conclusions, particularly when predictors are ~~cross-~~correlated. Additionally, MLR is able to consider the combination of different local and regional factors, thus enabling a better identification and interpretation of the manifold processes controlling the isotopic composition of rainfall.

5 Conclusions

This study analyzes the influence of local and regional meteorological factors on the isotopic composition of rainfall, expressed as $\delta^{18}O$, $\delta^{2}H$, and d-excess, in the Vietnamese Mekong Delta (VMD). For this purpose rainfall samples were taken on a weekly to a bi-weekly basis for a period of 1.5 years at An Long in the North-Eastern part of the VMD and analyzed ~~in the laboratory.~~for stable water isotopes. The regional factors potentially influencing ~~factors~~isotopic composition were derived by back-tracing of water particles up to 10 days from the target location using the HYSPLIT model, while the local factors were derived from local climate records. The influence of the different factors on the isotopic condition was quantified by multiple linear regression~~s~~ (MLR) of all factor combinations and relative importance analysis.

The MLR showed that up to 80% of the total variation of $\delta^{18}O$ can be explained by linear combinations of the selected factors. Similar results are obtained for $\delta^{2}H$. Contrary to this, only about 30% of the total variation of the d-excess can be explained by the selected factors, if the whole data series is used. General considerations regarding the controls of d-excess in tropical areas suggest that additional factors, like sea surface temperatures of the source region, need to be taken into account for an improved modeling of d-excess variation by MLR ~~over all~~overall seasons and source regions.

The study showed that local climatic factors, specifically~~ local~~ rainfall amount and temperature, play a minor role in controlling the isotopic composition of the rainfall at An Long. However, there is evidence that sub-cloud evaporation has a small effect during the dry season. Regional factors, on the contrary, dominate the isotopic composition of rainfall at An Long. 70% of the explained variance, i.e. a partial R^2 of up to 0.56, can be attributed to regional factors, among which precipitation amount along the transport pathway can ~~explains~~explain most of the ~~observed ~~variance~~ best~~. The remaining 30% of the explained variance is attributed to local factors, among which the temperature plays the most important role. These findings indicate that local secondary fractionation processes like sub-cloud evaporation modulate the isotopic composition, which is otherwise dominated by the rainout along the transport pathway of the precipitation.

Furthermore, the analysis of ~~different ~~transport durations implies that the moisture-producing precipitation ~~at~~reaching An Long travels about 4-6 days from its source, as the best regression results are obtained for these durations. For longer travel durations the explained variability of the regression decreases, suggesting that the moisture is recycled, i.e. precipitated and evaporated again, when the travel time exceeds 6 days.

If the data set is divided into seasonal subsets ~~based on~~defined by precipitation amount and water source regions, the MLR and importance analysis enables a better identification of factors and thus processes controlling the isotopic composition ~~for~~in the different seasons. For the late rainy and dry seasons (i.e. October to May~~)~~), the importance of ~~other~~regional (late rainy season) and local (dry season) factors ~~than P_hysplit ~~increases ~~and adds~~compared to P_hysplit, raising the explained variance, particularly for the late rainy season. The source regions and the associated transport pathways as well as local processes are more important for these periods, indicating that secondary fractionation processes by evaporation, either along the pathway (for d-excess) or locally (for $\delta^{18}O$ and $\delta^{2}H$), are more important than the amount effect, which is dominant during the Indian monsoon period. This is reasonable, because moisture transported to the Mekong Delta from the Pacific region and continental Asia passes through different climatic regimes, compared to the more uniform climatic conditions along the pathway from the Indian Ocean during the Indian summer monsoon.

In summary, it can be concluded that the proposed approach, consisting of simultaneous testing of all possible factors by MLR combined with relative importance analysis, is able to detect the relevant factors controlling the isotopic composition of rainfall as well as their individual contributions. If applied to seasonal data subsets, the predictions can be improved and the seasonal differences in controlling factors and processes can be identified. The validity of the approach is confirmed by similar, but mainly qualitative results obtained in other studies. The comparable results provide a strong indication that the method is able to identify the dominant factors responsible for the isotopic composition of rainfall without a priori knowledge or assumptions. In contrast to previous studies, the presented approach and results provide, however, a quantitative assessment of the impact of different factors, and thus information about the dominant processes of isotopic fractionation. It can support the interpretation of processes responsible for observed patterns of isotopic composition. The rather simple approach can, of course, not provide detailed information about atmospheric dynamics, but it provides a relatively simple and easy to apply approach supplementing or preceding more complex studies of isotopic composition with circulation models. Due to the simplicity, any scientist can easily apply this method in order to investigate factors controlling isotopic composition in precipitation at any given study area around the world without the requirement of setting up and in-depth knowledge about running a complex numerical atmospheric circulation model. Furthermore, the approach is easily reproducible and contains a rigorous quantitative analysis of the interplay of different driving factors. Moreover, the analysis can easily be extended to other factors and processes of importance in order to capture particularly the d-excess better, e.g. the sea surface temperatures at the source regions.

The similarity of isotopic signatures and LMWLs of stations all over Southeast Asia, as well as similar general climatic conditions, allows the conclusion~~,~~ that the findings are representative of a larger area. Particularly the similarity of the LMWLs, the variability of the monthly isotopic composition of rainfall, and climatic conditions of the VMD and Bangkok suggests that the results are representative for the whole Mekong Delta, and possibly for large areas of the southern tip of the continental Indochinese Peninsula.

The results have direct implications for the interpretation of ~~paleo-records~~paleorecords of stable water isotopes in terms of past climate conditions for ~~the Asian monsoon region~~Southeast Asia. Because this study shows that the factors controlling the isotopic signature of precipitation are changing between and even within seasons and that regional factors have large impacts on the local isotopic composition of rainfall. This needs to be considered in the reconstruction of past climates based on isotopic records: $\delta^{18}O$ and $\delta^{2}H$ values are likely to be representative for the rainfall during the dry season. However, as regional factors dominate during most of the rainy season (receiving the bulk of the total annual rainfall), reconstructions of the past climate have to be carefully interpreted. ~~Moreover, the analysis can easily be extended to other factors and processes of importance in order to capture the d-excess better, e.g. the sea surface temperatures at the source regions.~~ The proposed approach ~~could easily include additional factors, and could thus~~might open a pathway for ~~a better~~an improved reconstruction of ~~paleo climate~~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.

Acknowledgements~~.~~

We acknowledge the German Academic Exchange Service (DAAD) for the scholarship that enabled the first author to pursue a Ph.D. degree at the University of Potsdam, Germany. We acknowledge the International Atomic Energy Agency (IAEA) for their on-going maintenance of the Global Network of Isotopes in precipitation. The authors gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model and/or READY website (http://www.ready.noaa.gov) used in this publication. We would also like to thank the staff at An Long ~~station~~ for isotope sample collection of precipitation. Our thanks extend to the staff who analyzed part of the isotope samples at the laboratory of the Alfred-Wegener-Institute (AWI) in Potsdam, Germany. We thank anonymous reviewers for their constructive comments to improve the manuscript.

References

Aggarwal, P. K., Fröhlich, K., Kulkarni, K. M., and Gourcy, L. L.: Stable isotope evidence for moisture sources in the Asian summer monsoon under present and past climate regimes, Geophysical Research Letters, 31, 2004.

Angert, A., LEE, J. E., and Yakir, D.: Seasonal variations in the isotopic composition of near-surface water vapour in the eastern Mediterranean, Tellus B, 60, 674-684, 2008.

Araguás-Araguás, L., Froehlich, K., and Rozanski, K.: Stable isotope composition of precipitation over southeast Asia, Journal of Geophysical Research: Atmospheres, 103, 28721-28742, 1998.

Baker, A. J., Sodemann, H., Baldini, J. U., Breitenbach, S. F., Johnson, K. R., Hunen, J., and Zhang, P.: Seasonality of westerly moisture transport in the East Asian summer monsoon and its implications for interpreting precipitation δ18O, Journal of Geophysical Research: Atmospheres, 120, 5850-5862, 2015.

Bowen, G. J.: Spatial analysis of the intra-annual variation of precipitation isotope ratios and its climatological corollaries, Journal of Geophysical Research: Atmospheres, 113, 2008.

Breitenbach, S. F., Adkins, J. F., Meyer, H., Marwan, N., Kumar, K. K., and Haug, G. H.: Strong influence of water vapor source dynamics on stable isotopes in precipitation observed in Southern Meghalaya, NE India, Earth and Planetary Science Letters, 292, 212-220, 2010.

Cabello, M., Orza, J., Galiano, V., and Ruiz, G.: Influence of meteorological input data on backtrajectory cluster analysis? a seven-year study for southeastern Spain, Advances in Science and Research, 2, 65-70, 2008.

Chakraborty, S., Sinha, N., Chattopadhyay, R., Sengupta, S., Mohan, P., and Datye, A.: Atmospheric controls on the precipitation isotopes over the Andaman Islands, Bay of Bengal, Scientific reports, 6, 2016.

Clark, I. D., and Fritz, P.: Environmental isotopes in hydrogeology, CRC press, 1997.

Conroy, J. L., Cobb, K. M., and Noone, D.: Comparison of precipitation isotope variability across the tropical Pacific in observations and SWING2 model simulations, Journal of Geophysical Research: Atmospheres, 118, 5867-5892, 2013.

Craig, H.: Isotopic variations in meteoric waters, Science, 133, 1702-1703, 1961.

Crawford, J., Hughes, C. E., and Lykoudis, S.: Alternative least squares methods for determining the meteoric water line, demonstrated using GNIP data, Journal of Hydrology, 519, 2331-2340, 2014.

Crawford, J., Hollins, S. E., Meredith, K. T., and Hughes, C. E.: Precipitation stable isotope variability and subcloud evaporation processes in a semi-arid region, Hydrological Processes, 31, 20-34, 10.1002/hyp.10885, 2017.

Curio, J., and Scherer, D.: Seasonality and spatial variability of dynamic precipitation controls on the Tibetan Plateau, Earth System Dynamics, 7, 767, 2016.

Dansgaard, W.: Stable isotopes in precipitation, Tellus, 16, 436-468, 1964.

Dayem, K. E., Molnar, P., Battisti, D. S., and Roe, G. H.: Lessons learned from oxygen isotopes in modern precipitation applied to interpretation of speleothem records of paleoclimate from eastern Asia, Earth and Planetary Science Letters, 295, 219-230, 2010.

Delgado, J. M., Merz, B., and Apel, H.: Monsoon variability and the Mekong flood regime, in: The Mekong Delta System, Springer, 233-244, 2012.

Draxler, R. R., and Hess, G.: An overview of the HYSPLIT_4 modelling system for trajectories, Australian meteorological magazine, 47, 295-308, 1998.

Draxler, R. R., and Rolph, G.: HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) model access via NOAA ARL READY website (http://ready.arl.noaa.gov/HYSPLIT.php). NOAA Air Resources Laboratory, Silver Spring, in, Md, 2003.

Fleming, Z. L., Monks, P. S., and Manning, A. J.: Untangling the influence of air-mass history in interpreting observed atmospheric composition, Atmospheric Research, 104, 1-39, 2012.

Froehlich, K.: Deuterium excess in precipitation and its climatological significance, International Atomic Energy Agency, C&S Papers Series 13, 54-66, 2001.

Galewsky, J., Steen-Larsen, H. C., Field, R. D., Worden, J., Risi, C., and Schneider, M.: Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle, Reviews of Geophysics, 54, 809-865, 2016.

Gardner, R. H., and Trabalka, J. R.: Methods of uncertainty analysis for a global carbon dioxide model, Oak Ridge National Lab., TN (USA), 1985.

Gat, J. R.: Oxygen and hydrogen isotopes in the hydrologic cycle, Annual Review of Earth and Planetary Sciences, 24, 225-262, 1996.

Gebhart, K. A., Schichtel, B. A., and Barna, M. G.: Directional biases in back trajectories caused by model and input data, Journal of the Air & Waste Management Association, 55, 1649-1662, 2005.

Genereux, D. P., and Hooper, R. P.: Oxygen and hydrogen isotopes in rainfall-runoff studies, Isotope tracers in catchment hydrology, 319-346, 1998.

Gonfiantini, R., Fröhlich, K., Araguas-Araguas, L., and Rozanski, K.: Isotopes in groundwater hydrology, Isotope tracers in catchment hydrology, 203-246, 1998.

GSO, V.: Statistical Yearbook of Vietnam 2013, Hanoi: General Statistics Office (GSO) Vietnam, 2014.

Guan, H., Zhang, X., Skrzypek, G., Sun, Z., and Xu, X.: Deuterium excess variations of rainfall events in a coastal area of South Australia and its relationship with synoptic weather systems and atmospheric moisture sources, Journal of Geophysical Research: Atmospheres, 118, 1123-1138, 2013.

Harris, J. M., Draxler, R. R., and Oltmans, S. J.: Trajectory model sensitivity to differences in input data and vertical transport method, Journal of Geophysical Research: Atmospheres, 110, 2005.

Helsel, D. R., and Hirsch, R. M.: Statistical methods in water resources, US Geological survey Reston, VA, 2002.

Helsen, M., Van de Wal, R., Van den Broeke, M., Masson-Delmotte, V., Meijer, H., Scheele, M., and Werner, M.: Modeling the isotopic composition of Antarctic snow using backward trajectories: Simulation of snow pit records, Journal of Geophysical Research: Atmospheres, 111, 2006.

Hughes, C. E., and Crawford, J.: A new precipitation weighted method for determining the meteoric water line for hydrological applications demonstrated using Australian and global GNIP data, Journal of Hydrology, 464, 344-351, 2012.

Hung, N. N., Delgado, J. M., Güntner, A., Merz, B., Bárdossy, A., and Apel, H.: Sedimentation in the floodplains of the Mekong Delta, Vietnam. Part I: suspended sediment dynamics, Hydrological Processes, 28, 3132-3144, 2014.

IAEA: Statistical treatment of data on environmental isotopes in precipitation. , IAEA, 781, 1992.

IAEA: IAEA/GNIP precipitation sampling guide, International Atomic Energy Agency, 2014.

IAEA/WMO: Global Network of Isotopes in Precipitation. The GNIP Database. Accessible at: http://www.iaea.org/water, 2016.

Ichiyanagi, K., and Yamanaka, M. D.: Interannual variation of stable isotopes in precipitation at Bangkok in response to El Niño Southern Oscillation, Hydrological processes, 19, 3413-3423, 2005.

Ingraham, N. L.: Isotopic variations in precipitation, Isotope Tracers in Catchment Hydrology, 87-118, 1998.

Ishizaki, Y., Yoshimura, K., Kanae, S., Kimoto, M., Kurita, N., and Oki, T.: Interannual variability of H218O in precipitation over the Asian monsoon region, Journal of Geophysical Research: Atmospheres, 117, 2012.

Jacob, H., and Sonntag, C.: An 8-year record of the seasonal variation of 2H and 18O in atmospheric water vapour and precipitation at Heidelberg, Germany, Tellus B, 43, 291-300, 1991.

Johnson, J. W.: A heuristic method for estimating the relative weight of predictor variables in multiple regression, Multivariate behavioral research, 35, 1-19, 2000.

Joussaume, S., Sadourny, R., and Jouzel, J.: A general circulation model of water isotope cycles in the atmosphere, Nature, 311, 24-29, 1984.

Kendall, C., and Caldwell, E. A.: Fundamentals of isotope geochemistry, Isotope tracers in catchment hydrology, 51-86, 1998.

Kraha, A., Turner, H., Nimon, K., Zientek, L., and Henson, R.: Tools to Support Interpreting Multiple Regression in the Face of Multicollinearity, Frontiers in Psychology, 3, 44, 2012.

Kurita, N., Ichiyanagi, K., Matsumoto, J., Yamanaka, M. D., and Ohata, T.: The relationship between the isotopic content of precipitation and the precipitation amount in tropical regions, Journal of Geochemical Exploration, 102, 113-122, 2009.

Kurita, N., Fujiyoshi, Y., Nakayama, T., Matsumi, Y., and Kitagawa, H.: East Asian Monsoon controls on the inter-annual variability in precipitation isotope ratio in Japan, Climate of the Past, 11, 339-353, 2015.

Lachniet, M. S.: Sea surface temperature control on the stable isotopic composition of rainfall in Panama, Geophysical Research Letters, 36, 2009.

Landais, A., Risi, C., Bony, S., Vimeux, F., Descroix, L., Falourd, S., and Bouygues, A.: Combined measurements of 17 O excess and d-excess in African monsoon precipitation: Implications for evaluating convective parameterizations, Earth and Planetary Science Letters, 298, 104-112, 2010.

Lee, J. E., Fung, I., DePaolo, D. J., and Henning, C. C.: Analysis of the global distribution of water isotopes using the NCAR atmospheric general circulation model, Journal of Geophysical Research: Atmospheres, 112, 2007.

Lee, J. E., Risi, C., Fung, I., Worden, J., Scheepmaker, R. A., Lintner, B., and Frankenberg, C.: Asian monsoon hydrometeorology from TES and SCIAMACHY water vapor isotope measurements and LMDZ simulations: Implications for speleothem climate record interpretation, Journal of Geophysical Research: Atmospheres, 117, 2012.

Lee, K.-S., Grundstein, A. J., Wenner, D. B., Choi, M.-S., Woo, N.-C., and Lee, D.-H.: Climatic controls on the stable isotopic composition of precipitation in Northeast Asia, Climate Research, 23, 137-148, 2003.

LeGrande, A. N., and Schmidt, G. A.: Sources of Holocene variability of oxygen isotopes in paleoclimate archives, Clim. Past, 5, 441-455, 10.5194/cp-5-441-2009, 2009.

Levene, H.: Robust tests for equality of variances, Contributions to probability and statistics, 1, 278-292, 1960.

Liu, Z., Wen, X., Brady, E., Otto-Bliesner, B., Yu, G., Lu, H., Cheng, H., Wang, Y., Zheng, W., and Ding, Y.: Chinese cave records and the East Asia summer monsoon, Quaternary Science Reviews, 83, 115-128, 2014.

Mathieu, R., Pollard, D., Cole, J. E., White, J. W., Webb, R. S., and Thompson, S. L.: Simulation of stable water isotope variations by the GENESIS GCM for modern conditions, Journal of Geophysical Research: Atmospheres, 107, 2002.

Merlivat, L., and Jouzel, J.: Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation, Journal of Geophysical Research: Oceans, 84, 5029-5033, 1979.

Meyer, H., Schönicke, L., Wand, U., Hubberten, H.-W., and Friedrichsen, H.: Isotope studies of hydrogen and oxygen in ground ice-experiences with the equilibration technique, Isotopes in Environmental and Health Studies, 36, 133-149, 2000.

Noone, D., and Simmonds, I.: Associations between δ 18O of Water and Climate Parameters in a Simulation of Atmospheric Circulation for 1979-95, Journal of Climate, 15, 3150-3169, 2002.

Noone, D., and Sturm, C.: Comprehensive dynamical models of global and regional water isotope distributions, in: Isoscapes, Springer, 195-219, 2010.

Okazaki, A., and Yoshimura, K.: Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction, Climate of the Past, 13, 379, 2017.

Pausata, F. S., Battisti, D. S., Nisancioglu, K. H., and Bitz, C. M.: Chinese stalagmite [delta] 18O controlled by changes in the Indian monsoon during a simulated Heinrich event, Nature Geoscience, 4, 474-480, 2011.

Peng, T.-R., Wang, C.-H., Huang, C.-C., Fei, L.-Y., Chen, C.-T. A., and Hwong, J.-L.: Stable isotopic characteristic of Taiwan's precipitation: A case study of western Pacific monsoon region, Earth and Planetary Science Letters, 289, 357-366, 2010.

Pfahl, S., and Wernli, H.: Air parcel trajectory analysis of stable isotopes in water vapor in the eastern Mediterranean, Journal of Geophysical Research: Atmospheres, 113, 2008.

Pfahl, S., and Sodemann, H.: What controls deuterium excess in global precipitation?, Climate of the Past, 10, 771-781, 2014.

Renaud, F. G., and Kuenzer, C.: The Mekong Delta system: Interdisciplinary analyses of a river delta, Springer Science & Business Media, 2012.

Risi, C., Bony, S., and Vimeux, F.: Influence of convective processes on the isotopic composition (δ18O and δD) of precipitation and water vapor in the tropics: 2. Physical interpretation of the amount effect, Journal of Geophysical Research: Atmospheres, 113, 2008a.

Risi, C., Bony, S., Vimeux, F., Descroix, L., Ibrahim, B., Lebreton, E., Mamadou, I., and Sultan, B.: What controls the isotopic composition of the African monsoon precipitation? Insights from event-based precipitation collected during the 2006 AMMA field campaign, Geophysical Research Letters, 35, 2008b.

Risi, C., Bony, S., Vimeux, F., Chong, M., and Descroix, L.: Evolution of the stable water isotopic composition of the rain sampled along Sahelian squall lines, Quarterly Journal of the Royal Meteorological Society, 136, 227-242, 2010a.

Risi, C., Bony, S., Vimeux, F., and Jouzel, J.: Water-stable isotopes in the LMDZ4 general circulation model: Model evaluation for present-day and past climates and applications to climatic interpretations of tropical isotopic records, Journal of Geophysical Research: Atmospheres, 115, 2010b.

Risi, C., Noone, D., Worden, J., Frankenberg, C., Stiller, G., Kiefer, M., Funke, B., Walker, K., Bernath, P., and Schneider, M.: Process-evaluation of tropospheric humidity simulated by general circulation models using water vapor isotopic observations: 2. Using isotopic diagnostics to understand the mid and upper tropospheric moist bias in the tropics and subtropics, Journal of Geophysical Research: Atmospheres, 117, 2012.

Rose, K. A., Smith, E. P., Gardner, R. H., Brenkert, A. L., and Bartell, S. M.: Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty, Journal of Forecasting, 10, 117-133, 1991.

Rozanski, K., Araguas-Araguas, L., and Gonfiantini, R.: Relation between long-term trends of oxygen-18 isotope composition of precipitation and climate, Science, 258, 981-985, 1992.

Sánchez-Murillo, R., Birkel, C., Welsh, K., Esquivel-Hernández, G., Corrales-Salazar, J., Boll, J., Brooks, E., Roupsard, O., Sáenz-Rosales, O., and Katchan, I.: Key drivers controlling stable isotope variations in daily precipitation of Costa Rica: Caribbean Sea versus Eastern Pacific Ocean moisture sources, Quaternary Science Reviews, 131, 250-261, 2016.

SIWRP: Flood Planning in Vietnamese Mekong Delta until 2020, 2014.

Sodemann, H., Schwierz, C., and Wernli, H.: Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation influence, Journal of Geophysical Research: Atmospheres, 113, 2008.

Stein, A., Draxler, R., Rolph, G., Stunder, B., Cohen, M., and Ngan, F.: NOAA's HYSPLIT atmospheric transport and dispersion modeling system, Bulletin of the American Meteorological Society, 96, 2059-2077, 2015.

Stewart, M. K.: Stable isotope fractionation due to evaporation and isotopic exchange of falling waterdrops: Applications to atmospheric processes and evaporation of lakes, Journal of Geophysical Research, 80, 1133-1146, 1975.

Stohl, A.: Computation, accuracy and applications of trajectories—a review and bibliography, Atmospheric Environment, 32, 947-966, 1998.

Sturm, C., Zhang, Q., and Noone, D.: An introduction to stable water isotopes in climate models: benefits of forward proxy modelling for paleoclimatology, Climate of the Past, 6, 115-129, 2010.

Tan, M.: Circulation effect: climatic significance of the short term variability of the oxygen isotopes in stalagmites from monsoonal China—dialogue between paleoclimate records and modern climate research, Quaternary Sciences, 29, 851-862, 2009.

Tan, M.: Circulation effect: response of precipitation δ18O to the ENSO cycle in monsoon regions of China, Climate Dynamics, 42, 1067-1077, 2014.

Tindall, J., Valdes, P., and Sime, L. C.: Stable water isotopes in HadCM3: Isotopic signature of El Niño–Southern Oscillation and the tropical amount effect, Journal of Geophysical Research: Atmospheres, 114, 2009.

Tonidandel, S., LeBreton, J. M., and Johnson, J. W.: Determining the statistical significance of relative weights, Psychological methods, 14, 387, 2009.

Tonidandel, S., and LeBreton, J. M.: Relative importance analysis: A useful supplement to regression analysis, Journal of Business and Psychology, 26, 1-9, 2011.

Tri, V. K.: Hydrology and hydraulic infrastructure systems in the Mekong Delta, Vietnam, in: The Mekong Delta System, Springer, 49-81, 2012.

Uemura, R., Matsui, Y., Yoshimura, K., Motoyama, H., and Yoshida, N.: Evidence of deuterium excess in water vapor as an indicator of ocean surface conditions, Journal of Geophysical Research: Atmospheres, 113, 2008.

Vuille, M., Bradley, R., Werner, M., Healy, R., and Keimig, F.: Modeling δ18O in precipitation over the tropical Americas: 1. Interannual variability and climatic controls, Journal of Geophysical Research: Atmospheres, 108, 2003.

Vuille, M., Werner, M., Bradley, R., and Keimig, F.: Stable isotopes in precipitation in the Asian monsoon region, Journal of Geophysical Research: Atmospheres, 110, 2005.

Wang, B., Wu, R., and Lau, K.: Interannual variability of the Asian summer monsoon: contrasts between the Indian and the Western North Pacific-East Asian Monsoons*, Journal of climate, 14, 4073-4090, 2001.

Yang, H., Johnson, K., Griffiths, M., and Yoshimura, K.: Interannual controls on oxygen isotope variability in Asian monsoon precipitation and implications for paleoclimate reconstructions, Journal of Geophysical Research: Atmospheres, 121, 8410-8428, 2016.

Yoshimura, K., Oki, T., Ohte, N., and Kanae, S.: A quantitative analysis of short-term 18O variability with a Rayleigh-type isotope circulation model, Journal of Geophysical Research: Atmospheres, 108, 2003.

Yoshimura, K., Kanamitsu, M., Noone, D., and Oki, T.: Historical isotope simulation using reanalysis atmospheric data, Journal of Geophysical Research: Atmospheres, 113, 2008.

Yoshimura, K., Miyoshi, T., and Kanamitsu, M.: Observation system simulation experiments using water vapor isotope information, Journal of Geophysical Research: Atmospheres, 119, 7842-7862, 2014.

Table 1. Isotopic composition of precipitation at An Long and six selected GNIP stations throughout the Indochinese Peninsula.

Station Name	Location	Country	AL	P (mm)	T (ºC)	Period	$\delta^{18}O$		$\delta^{2}H$		d-excess	LMWL (OLSR)		LMWL (RMA)		LMWL (PWLSR)	
							WM	M	WM	M	WM	S	Int.	S	Int.	S	Int.
An Long	105.24º E - 10.72º N	Vietnam	2	985* (1550)	27.4* (27.2)	06.2014 - 12.2015	-6.4 ±1.5	-5.8 ±2.5	-40.9 ±11.5	-36.2 ±18.7	10.4 ±1.8	7.56 ±0.11	7.26 ±0.67	7.61 ±0.11	7.58 ±0.67	7.61 ±0.11	7.87 ±0.73
Hanoi **	105.84º E - 21.02º N	Vietnam	10	1659 ±257	24.8 ±0.5	2004-2007	-8.8 ±0.7	-5.9 ±0.5	-56.9 ±4.2	-33.8 ±3.6	13.5 ±1.5	7.91 ±0.10	12.45 ±1.25	7.99 ±0.18	12.90 ±1.22	7.77 ±0.21	10.92 ±1.91
Bangkok **	100.50º E - 13.73º N	Thailand	2	1558 ±314	28.5 ±0.6	1968-2015	-6.5 ±1.0	-5.2 ±1.0	-42.6 ±7.6	-33.2 ±7.2	9.4 ±1.6	7.35 ±0.04	5.36 ±0.47	7.53 ±0.08	6.29 ±0.47	7.68 ±0.07	7.25 ±0.49
Ko Samui **	100.03º E - 09.28º N	Thailand	7	1265 ±611	27.9 ±0.2	1979-1983	-5.8 ±1.4	-4.8 ±0.9	-28.8 ±7.1	-24.1 ±4.4	10.8 ±0.0	7.18 ±0.10	6.89 ±1.20	7.30 ±0.25	7.41 ±1.16	7.45 ±0.25	7.89 ±1.26
Ko Sichang **	100.80º E - 13.17º N	Thailand	26	877 ±320	27.9 ±0.6	1983-1995	-6.2 ±0.6	-6.2 ±1.1	-39.3 ±5.1	-39.7 ±8.8	10.2 ±0.6	7.62 ±0.06	7.61 ±1.15	7.72 ±0.18	8.16 ±1.12	7.77 ±0.23	8.65 ±1.44
Luang Prabang **	102.13º E - 19.88º N	Lao PDR	305	1228 ±178	25.7 ±0.7	1961-1967	-7.8 ±1.2	-6.7 ±0.3	-54.2 ±7.6	-45.9 ±0.9	8.4 ±1.9	7.90 ±0.13	7.97 ±2.00	8.01 ±0.27	8.70 ±1.93	7.80 ±0.28	7.52 ±2.29
Kuala Lumpur **	101.68º E - 03.13º N	Malaysia	26	1801 ±787	-	1993-2012	-7.3 ±0.8	-7.0 ±0.7	-46.6 ±7.7	-45.1 ±6.8	11.8 ±4.1	7.63 ±0.07	8.10 ±1.93	8.26 ±0.26	12.53 ±1.92	7.73 ±0.29	8.95 ±2.24

Note:

* Measured at An Long in 2015; numbers in parentheses show mean values of long-term measurements at Cao Lanh.

** Data is from https://nucleus.iaea.org/wiser/gnip.php (IAEA/WMO, 2016)

P: annual precipitation (mm/year); T: average temperature (ºC); AL: altitude (meter above sea level); WM: weighted mean value; M: mean value; S: slope; Int.: intercept

Table 2. Pairwise correlation coefficients between regional factors (P_hysplit, T_hysplit, H_hysplit, D_hysplit) and local factors (P_AL, T_AL, H_AL) and stable isotopic values ($\delta^{18}O$, $\delta^{2}H$, and d-excess). Bold and italic numbers denote significance at the 0.01 and 0.05 level (2-tailed), respectively. The meteorological data are aggregated to weekly values corresponding to the precipitation sampling at An Long.

(a)	P_hysplit	H_hysplit	T_hysplit	D_hysplit	P_AL	H_AL	T_AL	
$\delta^{18}O$	**-0.74**	**-0.45**	**-0.38**	*0.24*	**-0.34**	**-0.53**	**0.45**	Isotopic values vs. Regional and Local factors
$\delta^{2}H$	**-0.76**	**-0.47**	**-0.39**	0.20	-0.32	**-0.53**	**0.45**	
d-excess	0.18	0.04	0.07	**-0.36**	*0.27*	0.20	-0.15	

(b)	P_hysplit	H_hysplit	T_hysplit	D_hysplit	
P_AL	0.13	0.23	0.04	0.03	Regional factors vs. Local factors
H_AL	**0.38**	0.17	0.21	0.10	
T_AL	-0.21	0.05	0.17	**-0.33**	

(c)	P_hysplit	H_hysplit	T_hysplit	D_hysplit	
P_hysplit	1				Regional factors vs. Regional factors
H_hysplit	**0.77**	1			
T_hysplit	**0.59**	**0.67**	1		
D_hysplit	-0.10	-0.17	**-0.49**	1	

(d)	P_AL	H_AL	T_AL	
P_AL	1			Local factors vs. Local factors
H_AL	0.20	1		
T_AL	-0.14	**-0.78**	1	

Table 3. Results of the linear regression analysis between local relative humidity (H_AL) and isotopic values at An Long. Regressions that are statistically significant at the 0.05 level are marked in bold.

	Linear regression line	*r*	*R^2*	*p-value*	*n*	Period
$\delta^{18}O$ - δ^2H	$\delta^2H = 7.56*\delta^{18}O+7.26$	**0.99**	**0.99**	**0.000**	74	full year
	$\delta^2H = 7.62*\delta^{18}O+7.74$	**0.99**	**0.99**	**0.000**	67	rainy season (Jun-Nov)
	$\delta^2H = 7.58*\delta^{18}O+7.21$	**0.99**	**0.98**	**0.000**	42	early monsoon (Jun-Sep)
	$\delta^2H = 7.68*\delta^{18}O+8.6$	**0.99**	**0.99**	**0.000**	25	late monsoon (Oct-Nov)
	$\delta^2H = 6.9*\delta^{18}O+3.98$	**0.98**	**0.96**	**0.000**	7	dry season (Dec-May)
$\delta^{18}O$ - Humidity	$\delta^{18}O$ = -0.51*H_AL+36.05	**-0.53**	**0.28**	**0.000**	74	full year
	$\delta^{18}O$ = -0.46*H_AL+32.09	**-0.47**	**0.22**	**0.000**	67	rainy season (Jun-Nov)
	$\delta^{18}O$ = -0.33*H_AL+21.84	**-0.42**	**0.17**	**0.006**	42	early monsoon (Jun-Sep)
	$\delta^{18}O$ = -0.83*H_AL+63.12	**-0.61**	**0.37**	**0.001**	25	late monsoon (Oct-Nov)
	$\delta^{18}O$ = -0.56*H_AL+41.34	**-0.88**	**0.77**	**0.010**	7	dry season (Dec-May)
d-excess - Humidity	d-excess = 0.2*H_AL-6.36	0.20	0.04	0.090	74	full year
	d-excess = 0.13*H_AL-0.46	0.13	0.02	0.301	67	rainy season (Jun-Nov)
	d-excess = 0.18*H_AL-5.35	0.21	0.04	0.211	42	early monsoon (Jun-Sep)
	d-excess = -0.08*H_AL+17.44	-0.07	0.01	0.734	25	late monsoon (Oct-Nov)
	d-excess = 0.34*H_AL-19.42	0.31	0.10	0.455	7	dry season (Dec-May)

Table 4. Explained variance (partial R^2) of regional and local factors of the best MLR model according to the PRESS value. The first value indicates the absolute partial R^2, the second value the relative contribution to the overall explained variance.

	Whole period	Early rainy season	Late rainy season	Dry season
Regional factors	0.56 | 70%	0.51 | 68%	0.76 | 79%	0.14 | 22%
Local factors	0.24 | 30%	0.24 | 32%	0.20 | 21%	0.51 | 78%
Total	0.80 | 100%	0.75 | 100%	0.96 | 100%	0.65 | 100%

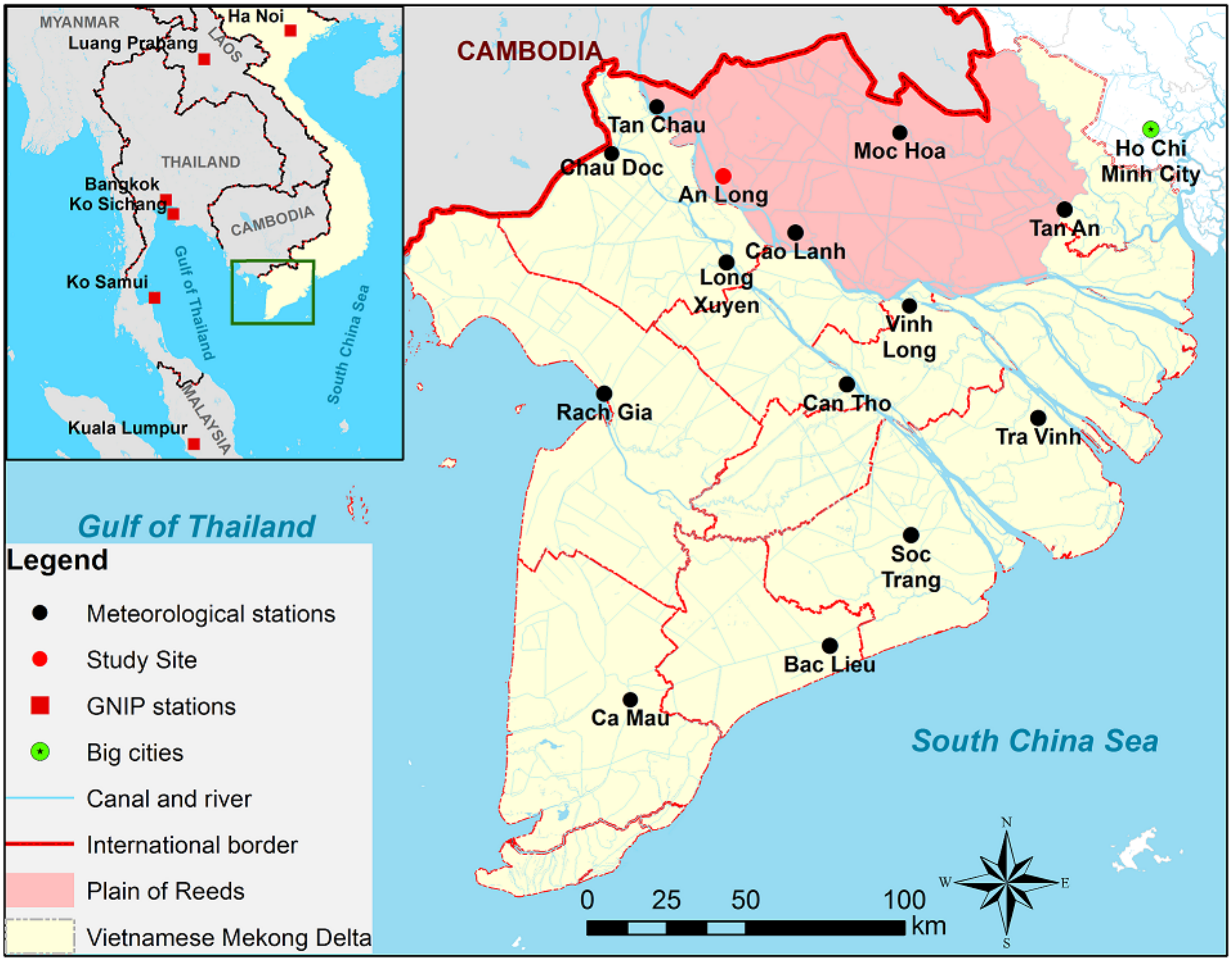


Figure 1: Sampling and monitoring sites in the study area

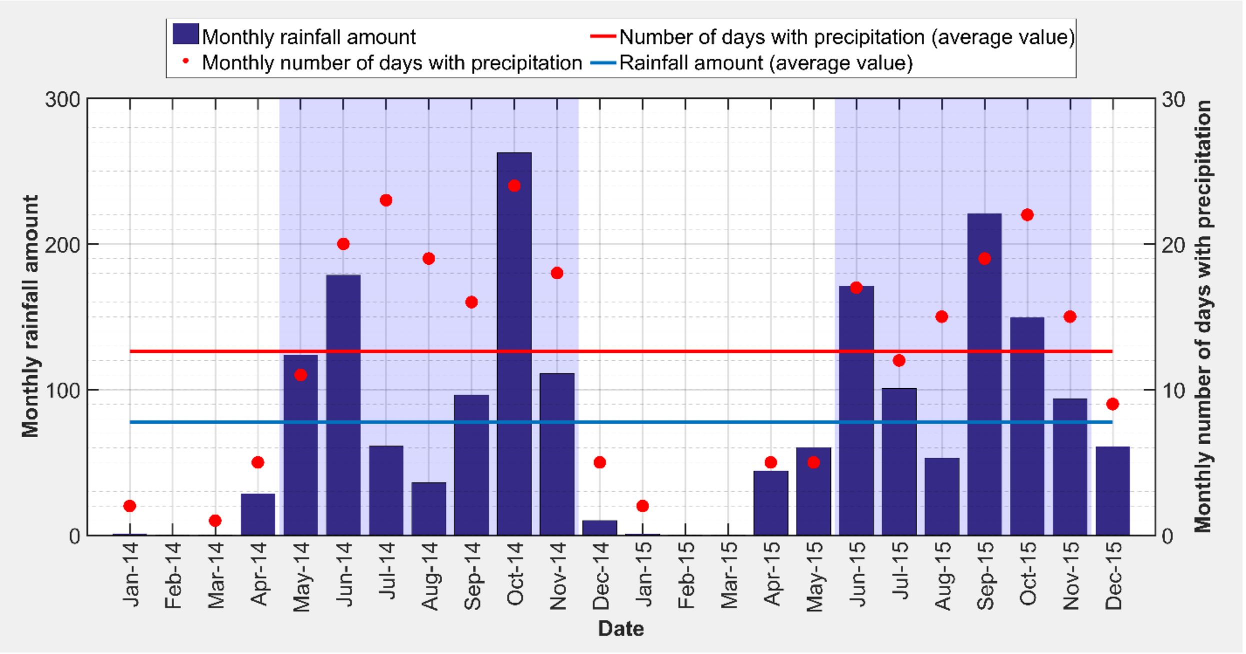


Figure 2: Monthly precipitation (mm) and a monthly number of days with precipitation for Cao Lanh station. Light blue background indicates rainy season.

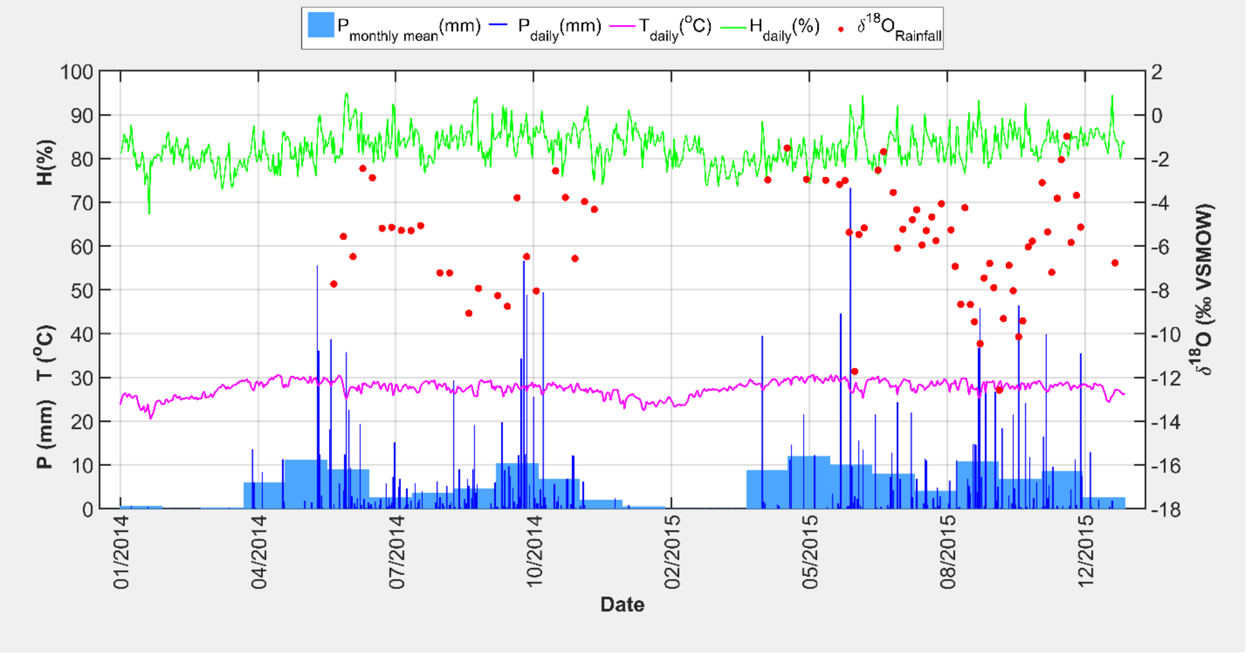


Figure 3: Climate data from the Cao Lanh meteorological station for the study period. Daily temperature (T) is given together with monthly and daily precipitation (P) and daily relative humidity (H). Weekly and bi-weekly $\delta^{18}O$ (‰ VSMOW) values of rainwater are presented as red circles.

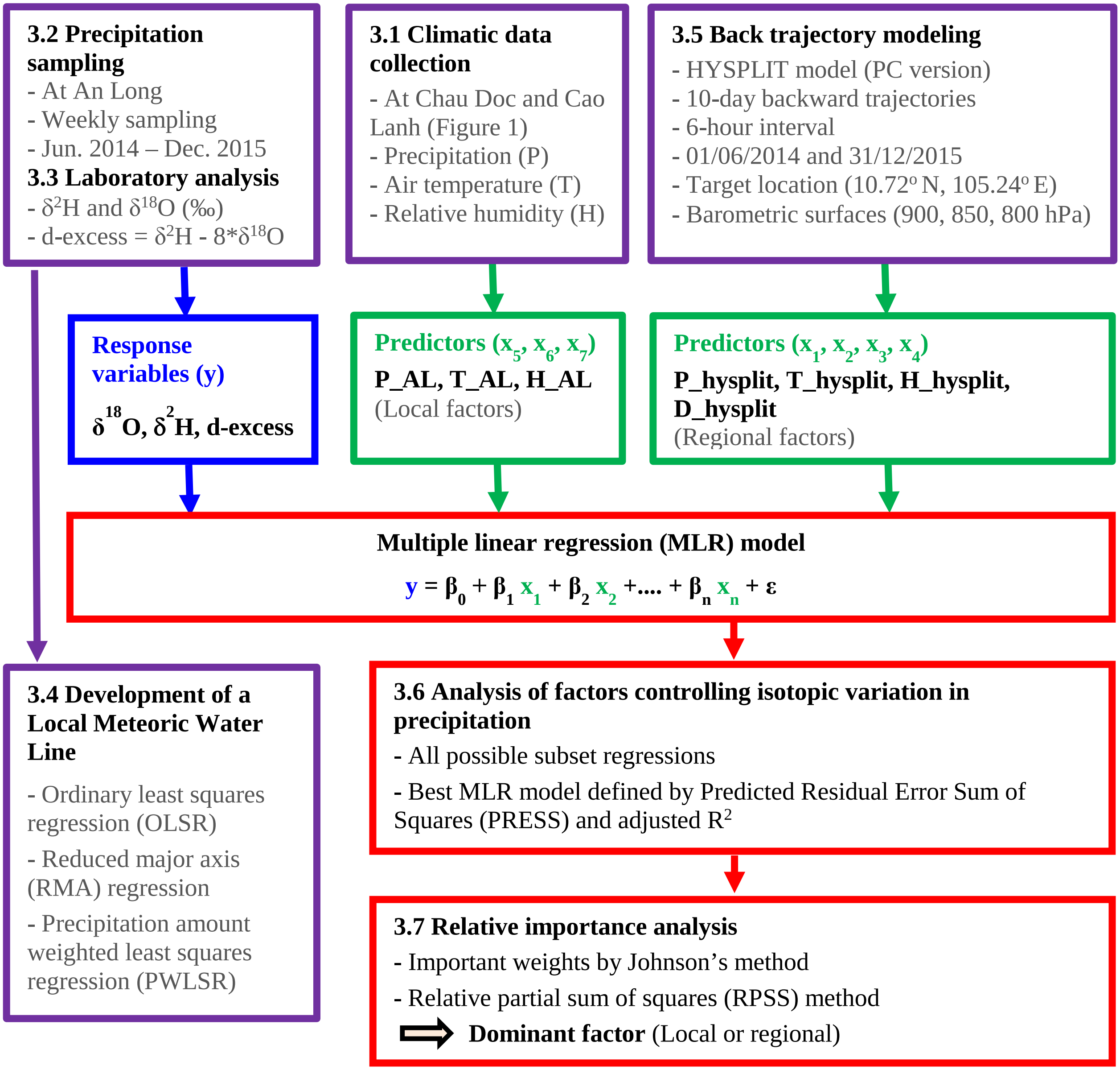


Figure 4: Methodology used in the study. Local precipitation (P_AL), air temperature (T_AL), and relative humidity (H_AL) at An Long. Precipitation amount (P_hysplit), mean temperature (T_hysplit) and relative humidity (H_hysplit) along the transport pathways, and moving distance from moisture sources (D_hysplit).

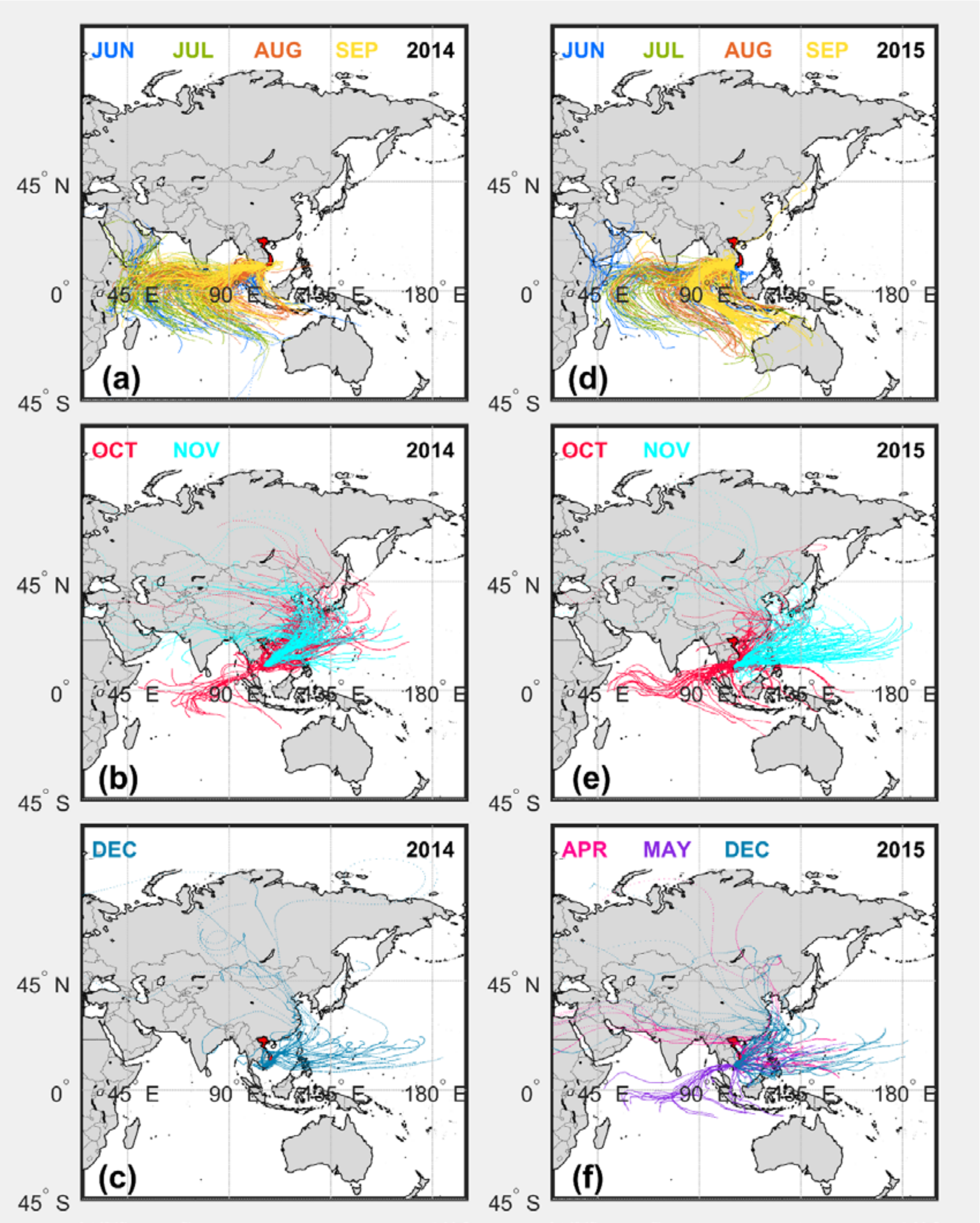
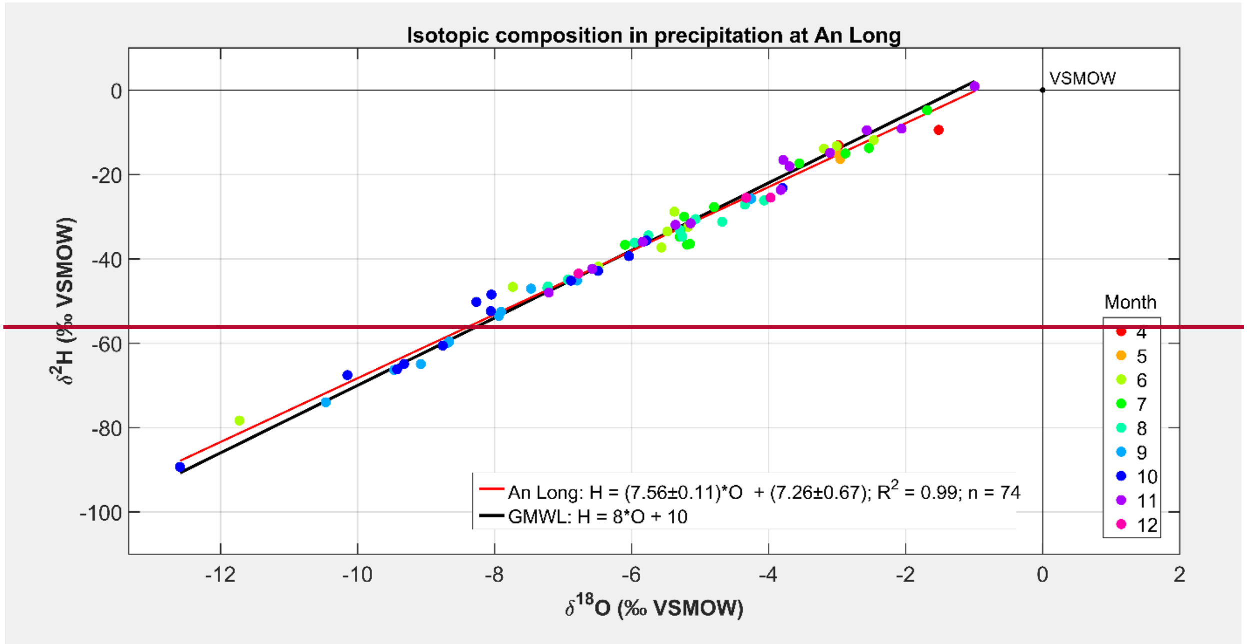


Figure 5: Back-trajectories indicating potential moisture sources of precipitation (plotted only for days with precipitation) at An Long station for the barometric surfaces at 850 hPa between June 2014 and December 2015. Left panels show the results for 2014, right panels for 2015; top row (a, d) early rainy season (June – September), middle row (b, e) late rainy season (October – November), bottom row (c, f) dry season (December – May). In January, February and March 2015 no rainfall was recorded.



Isotopic composition in precipitation at An Long
VSMOW
Month
4
5
6
7
8
9
10
11
12
An Long: H = (7.56±0.11)*O + (7.26±0.67); R2 = 0.99; n = 74
GMWL: H = 8*O + 10
δ2H (‰ VSMOW)
δ18O (‰ VSMOW)
0
-20
-40
-60
-80
-100
-12
-10
-8
-6
-4
-2
0
2

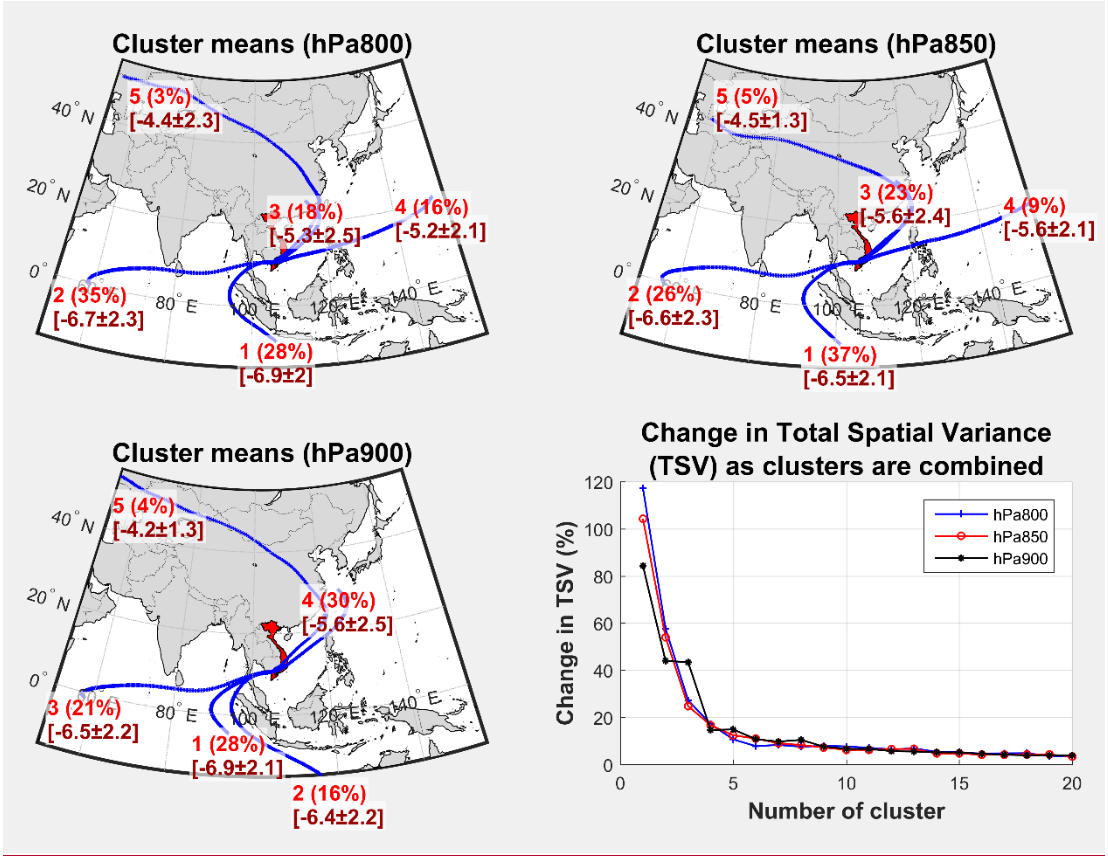


Figure 6: Spatial distribution of vapor trajectories (cluster means) for precipitation days at An Long for 3 barometric surfaces (800, 850, 900 hPa) between June 2014 and December 2015, and change in total spatial variance (TVS) for different cluster numbers. The TSV was used to identify the optimum number of clusters. Red texts indicate the cluster number (1-5) and the percent of all trajectories assigned to each of the five clusters. Brown texts indicate the mean $\delta^{18}O$ values for each cluster plus/minus the standard deviation of each cluster.

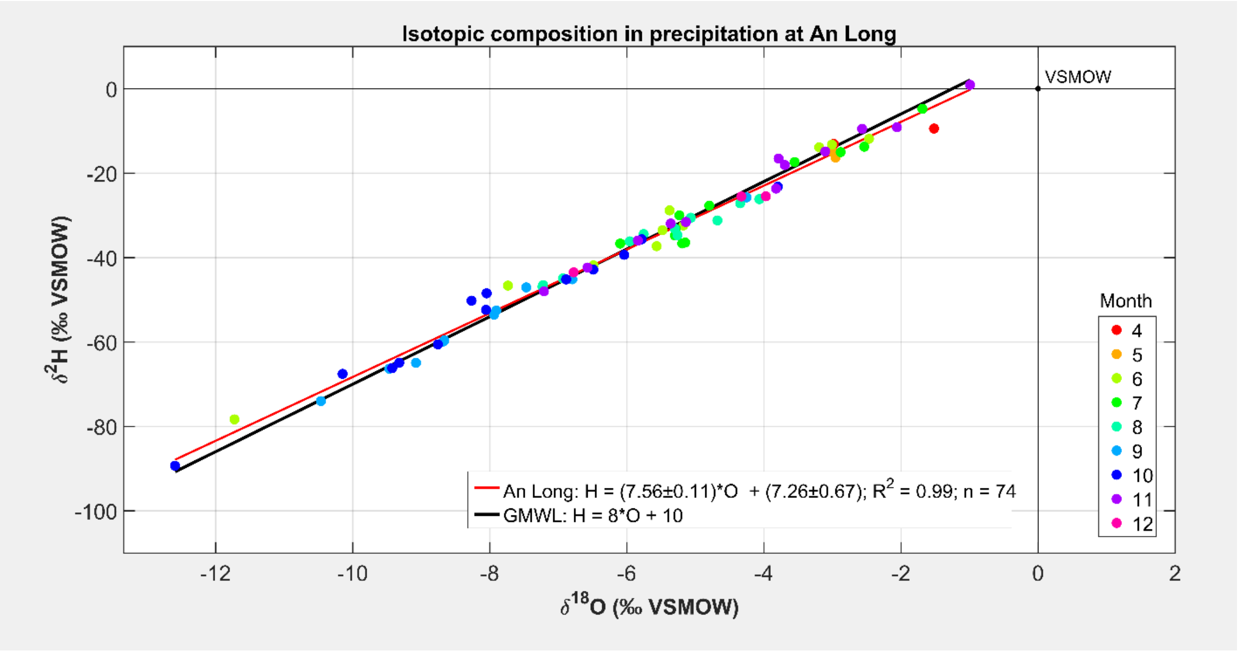
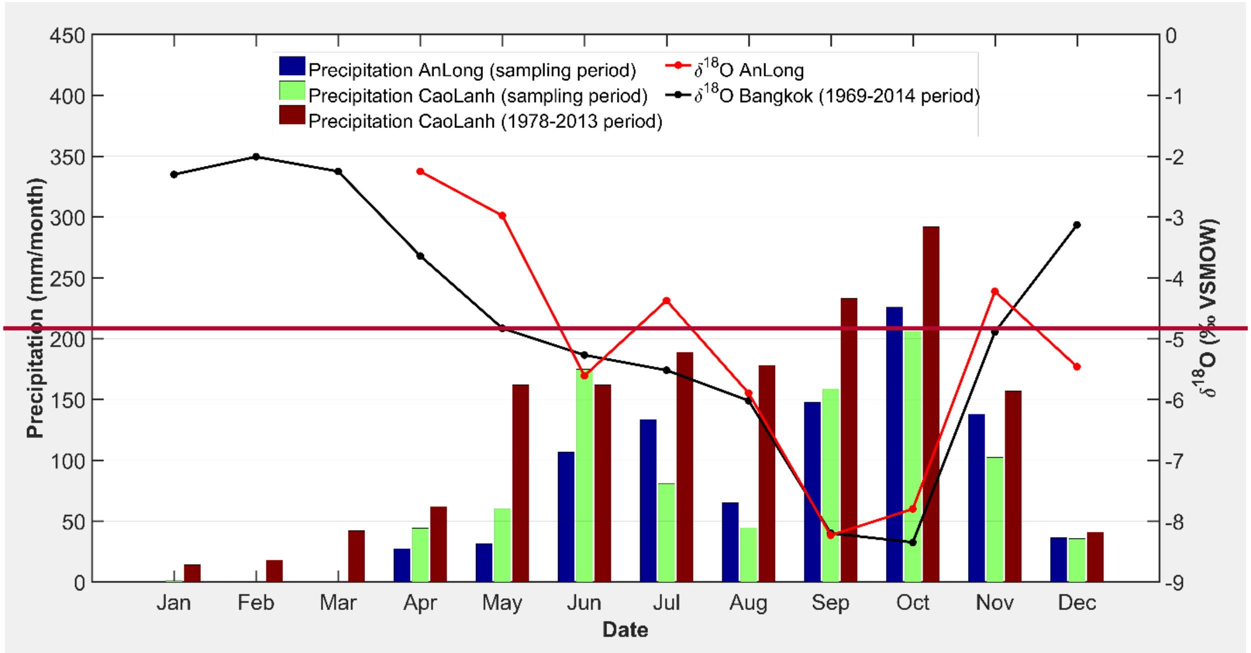


Figure 7: The LMWL of An Long in comparison to the GMWL.



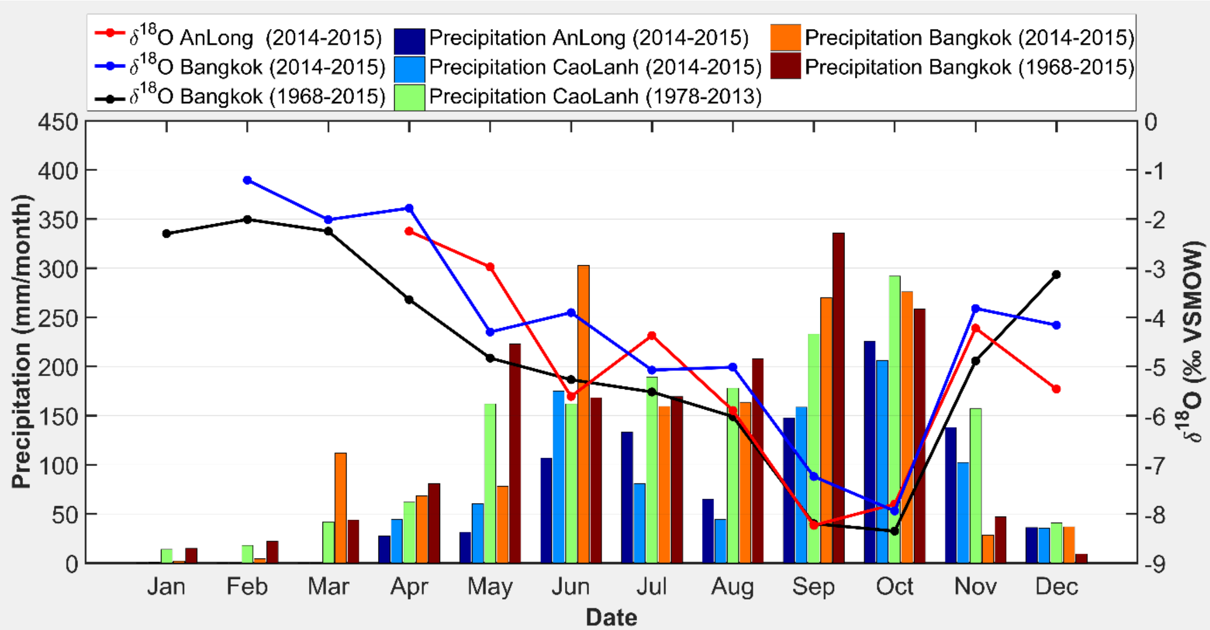
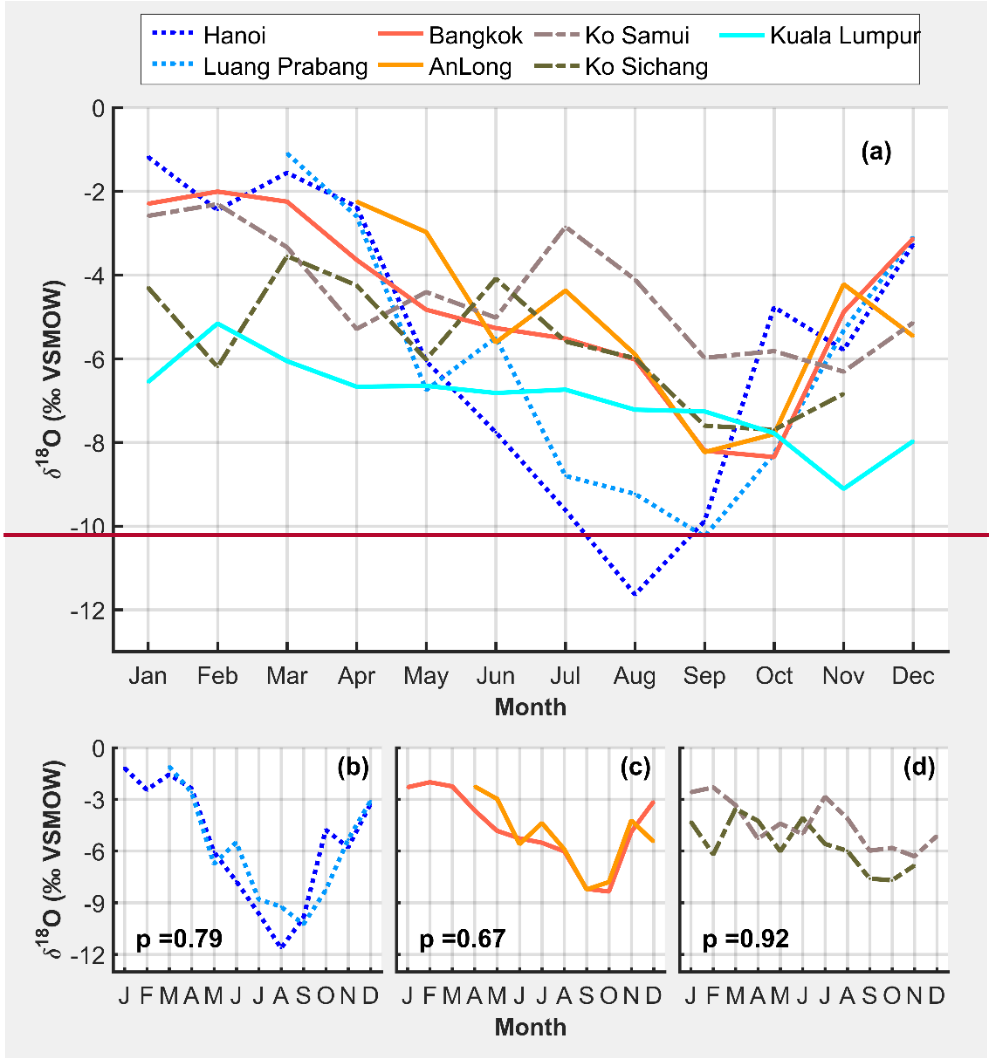


Figure 8~~: Annual~~: Seasonal variation of the average monthly precipitation for An Long and Cao Lanh and $\delta^{18}O$ values of precipitation ~~δ18O~~ for An Long (for the period of observation (red)) and Bangkok~~.~~ (both for the period of observation (blue) and the long-term mean (black)).



Hanoi
Luang Prabang
Bangkok
AnLong
Ko Samui
Ko Sichang
Kuala Lumpur
(a)
δ18O (‰ VSMOW)
0
-2
-4
-6
-8
-10
-12
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
(b)
(c)
(d)
p =0.79
p =0.67
p =0.92
-3
-6
-9
J F M A M J J A S O N D

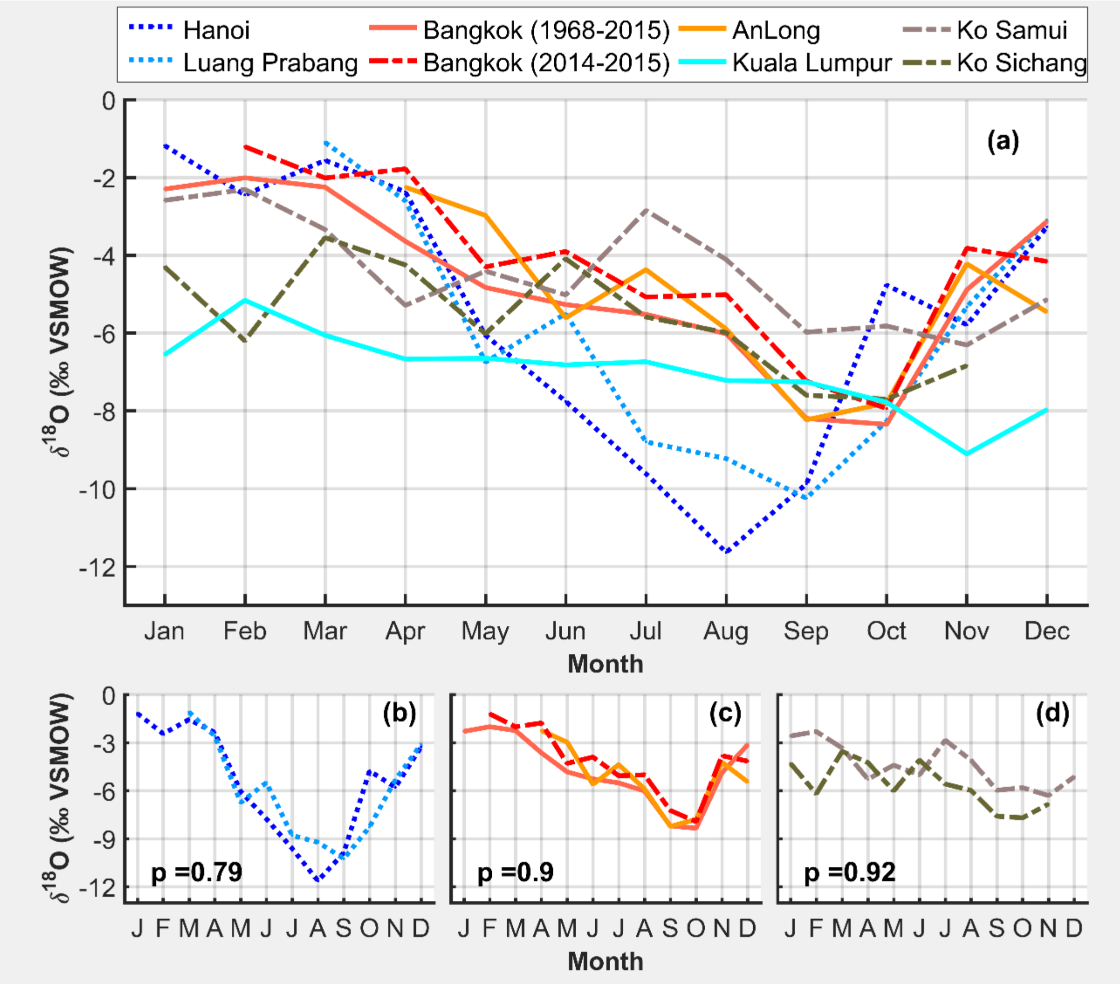


Figure 9: Seasonal monthly mean δ^{18}O values for An Long and GNIP data from the Indochinese Peninsula. The data is grouped according to similar variability tested with the Levene test. The p-values given in (b) to (d) are the test statistics. High values indicate similar variance. The time series of Bangkok is plotted for short-term (2014-2015) and long-term (1968-2015) periods.

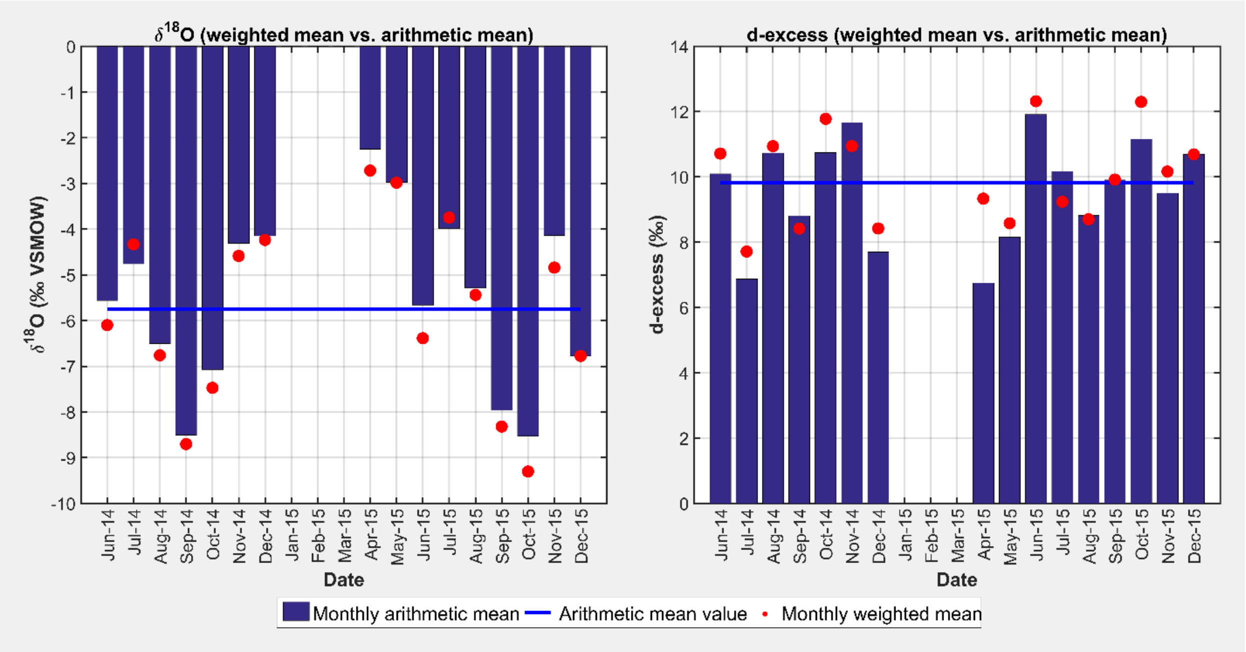


Figure 10: Arithmetic mean and amount-weighted mean monthly $\delta^{18}O$ (left) and d-excess (right) at An Long for the sampling period June 2014 to December 2015.

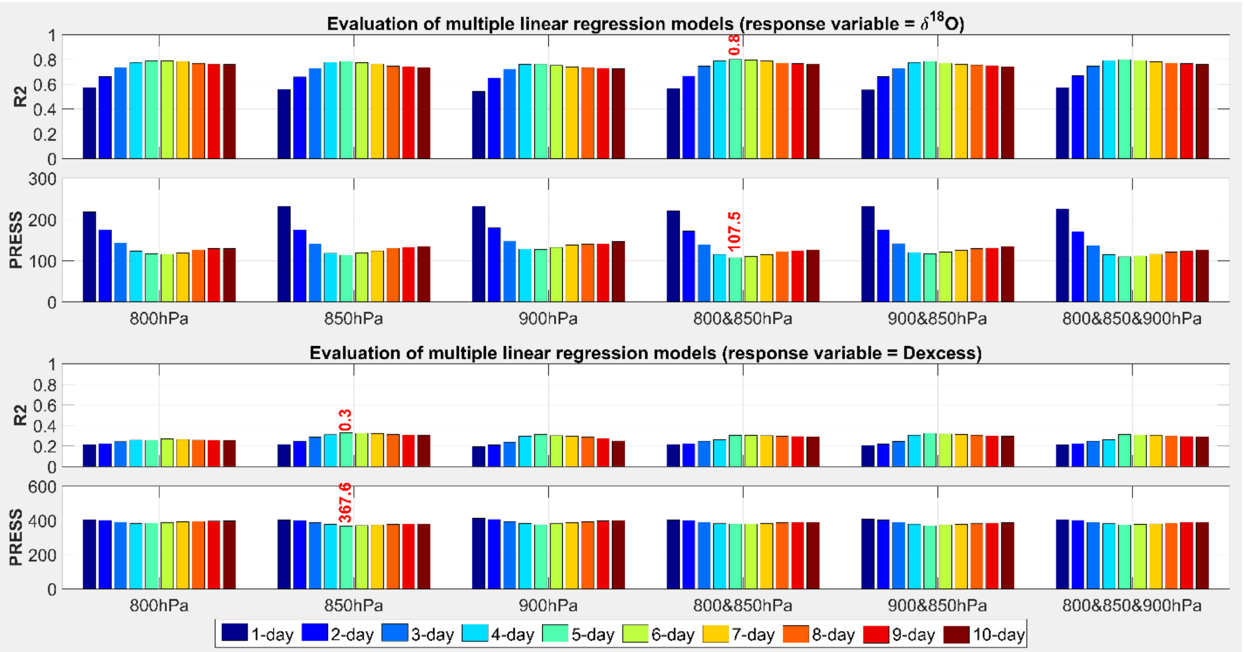


Figure 11: Evaluation of multiple linear regression (MLR) models applied for $\delta^{18}O$ and d-excess as response variables for different pressure levels used for three HYSPLIT backward trajectories and their combinations (mean values of the different levels). The best MLR model is highlighted with red text.

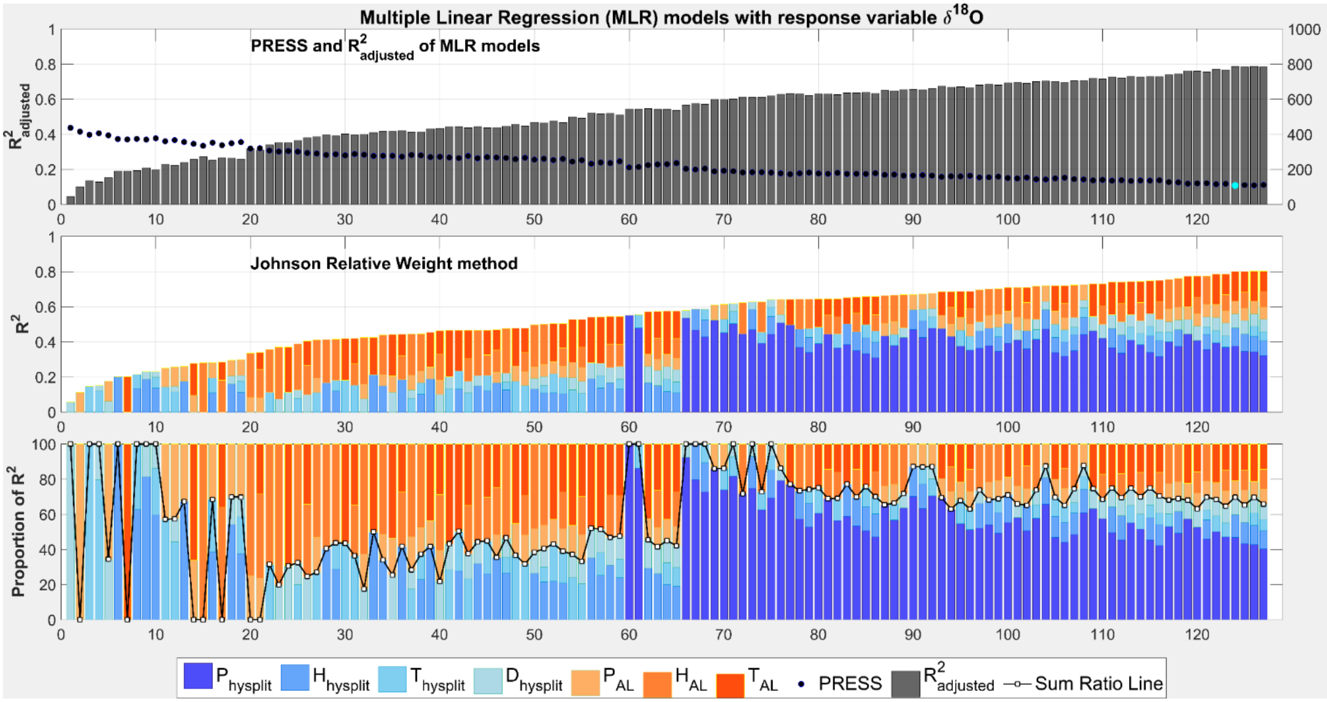


Figure 12. MLR with response variable $\delta^{18}O$ and relative importance analysis applied for all possible subsets. The 127 MLR models are sorted according to their R^2 values in ascendant order. Colors represent the relative contribution (in %) of the predictors. The sum ratio line separates the contribution of local (in red and orange) and regional (in blue) factors. PRESS and adjusted R^2 values indicate the quality of the MLR model. The best MLR model depicted by the lowest PRESS (model 124, highlighted by the cyan dot) explains 80% of the $\delta^{18}O$ variation (R^2 = 0.8).

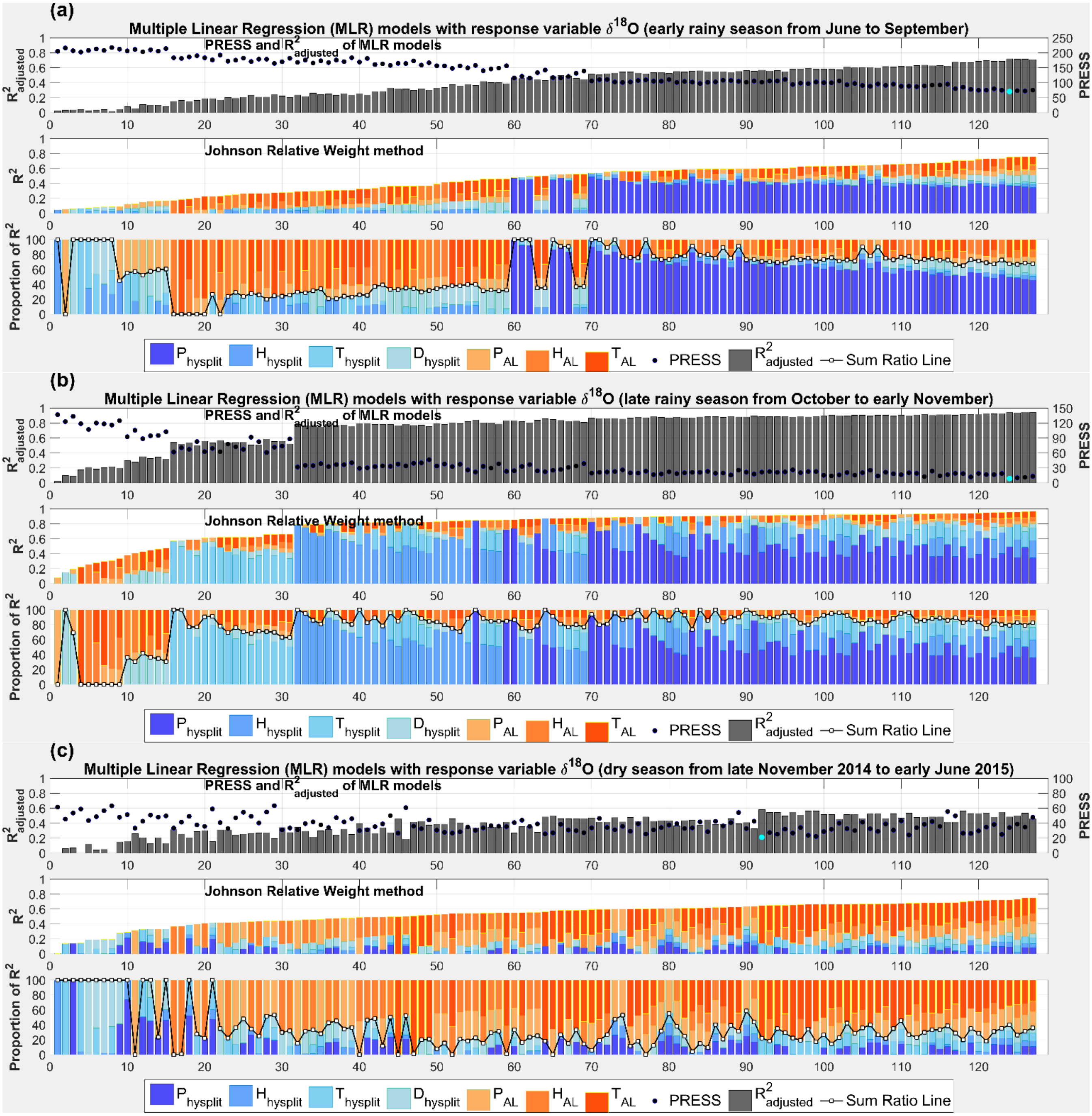


Figure 13. MLR with response variable $\delta^{18}O$ and relative importance analysis applied for all possible subsets (127 MLR models) for different seasons: a) early monsoon from June to September, b) late monsoon from October to mid-November, and c) the dry season from mid-November to mid-June.