



UvA-DARE (Digital Academic Repository)

Plant wax n-alkane biomarkers in the tropical Andes (Ecuador)

Teunissen van Manen, M.L.

Publication date

2020

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):


Teunissen van Manen, M. L. (2020). *Plant wax n-alkane biomarkers in the tropical Andes (Ecuador)*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.



5

“On replicating measurements of the plant wax *n*-alkane molecular proxy”
[in review, 2020].

Milan L Teunissen van Manen¹, Boris Jansen¹,
Samira Absalah¹, William D Gosling¹

¹ *Institute for Biodiversity and Ecosystem Dynamics (IBED),
University of Amsterdam.*



ABSTRACT

n-Alkanes derived from higher terrestrial plants preserved in soils and sedimentary records have been found to provide information on past environments, but the interpretation of these records remains debated. In an effort to improve the interpretation of such *n*-alkane sedimentary records, multiple studies have examined the factors controlling the variability of modern plant wax *n*-alkanes, such as plant species, vegetation composition and environmental factors. However, another source of variability that could confound interpretation is low reproducibility of the *n*-alkane measurement (i.e. processing, extraction, fractionation and measurement). Sample replicate data (which tests the reproducibility of the measurement), available from literature and from our own studies, suggests that absolute concentrations of *n*-alkanes frequently vary between replicate samples; coefficient of variation (CV) ranges between 0-72 %. We examined if data characteristics (such as study authors, concentration, *n*-alkane range and source material measured) could explain the CV in the sample replicates data. We identified that potential sources were related to low reproducibility of the protocol, incomplete homogenization of the sample, error amplification due to standardization, or a combination of these. Although *n*-alkane concentrations varied between replicates, the *n*-alkane chain-length distributions remained consistent between replicates. We conclude that to better understand the relevance of the variability of the *n*-alkane signal in biomarker research it is essential that measurement robustness is explicitly quantified, reported on, and discussed prior to attributing variability to other sources.

1 | INTRODUCTION

Due to the ubiquity of plant wax *n*-alkanes in nature and their persistence in soils and sediments plant wax *n*-alkanes are being developed as a biomarker with many applications in environmental sciences and palaeoecology (Jansen and Wiesenberg, 2017; Kirkels et al., 2013). For example, in environmental sciences, plant wax *n*-alkanes can indicate the source of particulate air pollution (for example, Lyu et al., 2017). In palaeoecology, plant wax *n*-alkanes are being developed as a proxy for determining past vegetation composition and for reconstructing past environmental conditions (for example, Crausbay et al., 2014; Zech et al., 2009, 2012).

Plant wax *n*-alkanes are produced by plants as part of a complex waxy epicuticular layer, that is hypothesized to protect the plant from stressful environmental conditions and herbivory (Eglinton and Hamilton, 1967; Koch and Ensikat, 2008; Shepherd and Griffiths, 2006). In response to environmental stressors, plants have been found to alter the amounts of wax produced, as well as the chemical composition of the wax layer (Eglinton and Hamilton, 1967; Koch and Ensikat, 2008). The plasticity of the epicuticular wax layer is reflected in the *n*-alkane fraction of the wax layer (Feakins et al., 2016; Hoffmann et al., 2013; Teunissen van Manen et al., 2019). However, to what extent *n*-alkane plasticity occurs, and is consistent, across plant taxa and which environmental factors control the amount and composition of the *n*-alkane fraction are active fields of study (e.g. Hoffmann et al., 2013, Feakins et al., 2016, Liu et al., 2018).

The push to understand the *n*-alkane biomarker signal has led to the identification of many factors controlling *n*-alkane concentrations and composition variability, including geography, environmental conditions, seasonality, species traits, vegetation structure and leaf ontogeny (Bush and McInerney, 2015, 2013; Diefendorf and Freimuth, 2017; Feakins et al., 2016; Jansen and Wiesenberg, 2017; Kirkels et al., 2013). However, one potential source of variability seems to have been mostly overlooked, namely, the potential for variability caused during the analysis of the sample (i.e. the extraction, purification and measurement of *n*-alkanes from a sample). There are multiple steps that can help quantify the *n*-alkane measurement error. These include: analysing blank samples alongside normal samples to check for contamination, measuring external standards in repeat, running equipment reliability tests to check for (compound specific) equipment performance, or measuring replicates of a sample (i.e. measuring two or more subsets of a homogenized sample) to check for the overall reproducibility of the measurement.

A survey of the literature studying the variability of *n*-alkane concentrations and composition from 2017 to 2019, shows that out of 42 studies, only three studies report sample replication results (Appendix A). Thirty-three of the studies published did not report any type of reproducibility tests, and six reported equipment testing (such as GC-MS accuracy tests) or repeated measures of external standards (Appendix A). Additionally, the three published

studies that did report sample replicate results each had a different style of reporting the results of sample replication; from very general (stating that sample replicates were below a certain threshold of analytical variability (Lyu et al., 2017), to very explicit (reporting replicate averages, standard deviation and coefficients of variation in table format (Hoffmann et al., 2013; Teunissen van Manen et al., 2019)).

The lack of reporting on reproducibility testing, sample replication in particular, means that: (1) very little is known about how much *n*-alkane measurements vary, (2) there is no literature standard on what comprises acceptable variability (or robust data) or how to report it, and (3) the consequences of measurement variability for interpretation are not discussed. Not understanding *n*-alkane measurement variability impedes the confident employment of *n*-alkanes as a molecular proxy, as interpretations of past *n*-alkane records could be confounded.

Here we present a novel attempt at characterizing *n*-alkane variability, by studying the limited replicate sample data available in literature. Specifically, we present the sample analytical replicate data of all analysis done by the authors (published and under review), and compare our results with the available data of studies that report sample replication. In particular, we take a critical look at the data to: (1) assess the extent of the sample replicates variability in the available data, (2) assess how the observed sample replicate variability affects the results of the available data, and (3) discern likely sources for the observed sample replicate variability, based on the sample replicate data characteristics and the available literature. Finally, we reflect on the implications for the *n*-alkane biomarker field and recommend reproducibility testing practices to further our understanding of *n*-alkane variability. As the literature survey indicates, the available replicate sample data is limited and therefore we recognize that the data cannot conclusively resolve these matters; however, we are able to identify the potential sources of measurement variability that merit further investigation, and encourage further exploration of *n*-alkane measurement variability.

2 | MATERIALS & METHODS

We gathered the data analysed by the authors (Jansen et al., 2006; Teunissen van Manen et al., 2019, 2020) and calculated the replicate sample set mean concentration (CON_w), standard deviation (SD) and coefficient of variation (CV) per replicate set. We gathered additional data of sample replication results that were openly available for comparison (Hoffmann et al., 2013). To simplify terminology we have coded the dataset references as follows (Table 1): Hoff (Hoffmann et al., 2013), Jans (Jansen et al., 2006), TvM1 (Teunissen van Manen et al., 2019), TvM2 (Teunissen van Manen et al., 2020).

Hoff replicated all measurements in triplicate and reported averages and coefficients of variation in their manuscript. We retrieved these data and calculated the SD based on

the given CONw and associated CV values in the table. For the other datasets (Jans, TvM1, TvM2) the measured n -alkane concentration per replicate subsample (Σalk) was available, from which the CONw, SD and CV per replicate set were calculated directly. The n -alkane Σalk , CONw, SD and CV were calculated following:

$$\Sigma alk = \sum_i^n [C_n] \quad (\text{Eq. 1})$$

Where C_n is the concentration of a n -alkane chain, measured between the range $n-i$ from a subsample.

$$CONw = \frac{\sum(\Sigma alk)}{N} \quad (\text{Eq. 2})$$

Where N is the number of subsamples of a replicate sample set and Σalk is the total concentration of n -alkanes in each subsample.

$$SD(\sigma) = \sqrt{\frac{\sum(\Sigma alk - CONw)^2}{N}} \quad (\text{Eq. 3})$$

Where Σalk is the concentration of n -alkanes per subsample, CONw is the mean n -alkane concentration of the replicate sample and N is the number of subsamples of a replicate sample set.

$$CV = \frac{SD}{CONw} \times 100 \quad (\text{Eq. 4})$$

Where SD is the standard deviation from the mean n -alkane concentration of the replicates and CONw is the mean n -alkane concentration of the replicates. Prior to analysis the concentrations of all datasets were recalculated to ng g^{-1} for ease of comparison. The concentrations in the Hoff dataset were standardized to grams of leaf dry weight, instead of grams of dry sample as in other datasets (Hoffmann et al., 2013).

We assessed the effect of the reported measurement replicability on the results of the associated dataset in two ways. First, we compared the replicate sample CONw variability to the overall CONw variability of the studied system. For this we chose the Hoff and the soil analysis subset of TvM2 (TvM2-soil) datasets because these were sampled and replicated extensively along a significant environmental gradient. We quantified the variability of the studied system by calculating site and transect variability (CONw, SD and CV), which we then compared to the reported sample replicate variability. The Hoff data was sampled from *Acacia* and *Eucalyptus* (i.e. *Corymbia* and *Eucalyptus*) trees along a latitudinal transect in Australia (12.44°S - 23.70°S, Hoffmann et al., 2013), the TvM2-soil data was sampled along a elevational transect in Ecuador (635 - 3507 m above sea level, Teunissen van Manen et al., 2020). Second, we visually assessed the effect of n -alkane measurement variability on the

n-alkane relative distributions (i.e. the *n*-alkane pattern). Specifically, we selected replicate sample set “14” from the TvM2 dataset because it had the highest CV value in the dataset for which the individual *n*-alkane chain length data was available. For each of the replicates in the sample set, we calculated the relative abundance of each chain length (% of Σ alk). We then calculated the mean relative abundance of each chain length for the replicate sample set, and the standard deviation from the mean. This gives the mean proportion of each chain length, and how much this proportion varies between replicates.

To discern likely sources for the observed sample replicate variability, we noted four key dataset characteristics: (1) study authors, (2) sample source (leaf, necromass, soil or sediment), (3) sample concentration measured, and (4) the *n*-alkane chain length range measured. Due to the differences in the units measured for the sample concentration characteristic ('CONw') we analysed the Hoff dataset separately. For the results to be comparable between datasets we transformed the continuous variable CONw into a categorical variable with three levels ('low', 'middle', 'high') based on the quartiles of the data ('CONw level'). The quartiles were calculated using the 'quantile()' function from the core 'stats' package in R (R Core Team, 2017).

The gathered data was not suitable for direct statistical comparison, due to the many methodological differences between studies and the limited amount of data. For this reason we decided that visual exploration of the potential sources for the reported sample replicate CV was more appropriate.

3 | RESULTS

The CV values of the 63 replicate sets presented varied between 0.1% and 72%, the median CV value was 15% (Table 1, Figure 1). Therefore, more than half (41 replicate sets, 65%) of the replicate set CV values surpassed the lower (10%) acceptability threshold, and furthermore 26 replicate sets (41%) had CV values surpassed higher (20%) threshold (Figure 1).

The CV values of the Hoff and TvM2-soil data varied between 2-72% and 1-18%, respectively (Table 1). The Hoff-*Acacia* data replicate variability falls within the *Acacia* transect variability (transect CV = 88%) (Figure 2a, Appendix B). The Hoff-Eucalyptus data replicate variability falls within both the site (site CV between 3-109%) and the transect variability (transect CV = 95%) (Figure 2a, Appendix B). The TvM2-soil data replicate variability falls within both the site (site CV between 16-116%) and the transect variability (transect CV = 101%) (Figure 2b, Appendix B).

Table 1 | Overview of replicated samples showing the replicated sample (set) name, the number of subsamples in the sample set (N), the average n-alkane concentration of the replicate sample set (CONw), the replicate set standard deviation from the mean n-alkane concentration (SD), the coefficient of variation from the mean n-alkane concentration (CV). Also given are the data characteristics: the CONw level category, the source material, the n-alkane range measured and the study author with respective dataset naming code.

[next page]

Replicate sets	N	CONw*	±SD	CV(%)	CONw category	Source material	n-Alkane range	code
<i>Acacia aneura</i>	3	2818000	1944420	69	high	leaf	10 (C23-C33)	Hoff
<i>Acacia auriculiformis</i>	3	238000	104720	44	high	leaf	10 (C23-C33)	Hoff
<i>Acacia colei</i>	3	740000	229400	31	high	leaf	10 (C23-C33)	Hoff
<i>Acacia colei</i>	3	817000	98040	12	high	leaf	10 (C23-C33)	Hoff
<i>Acacia cowleana</i>	3	1486000	29720	2	high	leaf	10 (C23-C33)	Hoff
<i>Acacia kempiana</i>	3	2752000	165120	6	high	leaf	10 (C23-C33)	Hoff
<i>Acacia kempiana</i>	3	1340000	281400	21	high	leaf	10 (C23-C33)	Hoff
<i>Acacia shirleyi</i>	3	210000	48300	23	high	leaf	10 (C23-C33)	Hoff
<i>Acacia sp.</i>	3	145000	14500	10	middle	leaf	10 (C23-C33)	Hoff
<i>Corymbia aparrerinja</i>	3	265000	84800	32	high	leaf	10 (C23-C33)	Hoff
<i>Corymbia confertiflora</i>	3	10000	3400	34	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia dichromophloia</i>	3	22000	2200	10	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia foelscheana</i>	3	17000	1190	7	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia polycarpa</i>	3	28000	1400	5	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia terminalis</i>	3	110000	60500	55	middle	leaf	10 (C23-C33)	Hoff
<i>Corymbia terminalis</i>	3	30000	4200	14	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia terminalis</i>	3	18000	8280	46	low	leaf	10 (C23-C33)	Hoff
<i>Corymbia terminalis</i>	3	34000	8500	25	low	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus leucophloia</i>	3	117000	39780	34	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus miniata</i>	3	56000	7840	14	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus pruinosa</i>	3	155000	12400	8	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus pruinosa</i>	3	66000	21780	33	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus pruinosa</i>	3	89000	59630	67	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus tectiflora</i>	3	16000	5440	34	low	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus tectiflora</i>	3	38000	25080	66	low	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus tetradonta</i>	3	58000	14500	25	middle	leaf	10 (C23-C33)	Hoff
<i>Eucalyptus tetradonta</i>	3	90000	64800	72	middle	leaf	10 (C23-C33)	Hoff
72 a	3	21543	7347	34	high	soil	22 (C12-C34)	Jans
72 b	3	45640	3838	8	high	soil	22 (C12-C34)	Jans
73 a	3	107847	38673	36	high	soil	22 (C12-C34)	Jans
73 b	3	143393	85338	60	high	soil	22 (C12-C34)	Jans
76 a	3	3396	1225	36	high	soil	22 (C12-C34)	Jans

Table 1 | (continued)

Replicate sets	N	CONw*	±SD	CV(%)	CONw category	Source material	n-Alkane range	code
76 b	3	3286	1239	38	high	soil	22 (C12-C34)	Jans
77 a	3	2953	973	33	high	soil	22 (C12-C34)	Jans
77 b	3	3307	408	12	high	soil	22 (C12-C34)	Jans
<i>Guaera kunthiana</i> 15	3	148	26	18	middle	leaf	10 (C23-C33)	TvM1
<i>Miconia clathrantha</i> 22	2	53	9	16	low	leaf	10 (C23-C33)	TvM1
<i>Miconia theaezans</i> 17	3	291	29	10	middle	leaf	10 (C23-C33)	TvM1
1	3	269	62	23	middle	necromass	12 (C21-C33)	TvM2
7	3	445	39	9	high	necromass	12 (C21-C33)	TvM2
12	3	343	40	12	middle	necromass	12 (C21-C33)	TvM2
14	3	789	244	31	high	necromass	12 (C21-C33)	TvM2
BIOM 006	3	138	7	5	middle	soil	18 (C15-C33)	TvM2
BIOM 007	3	168	31	18	middle	soil	18 (C15-C33)	TvM2
BIOM 010	3	322	6	2	middle	soil	18 (C15-C33)	TvM2
BIOM 018	3	49	5	11	low	soil	18 (C15-C33)	TvM2
BIOM 022	3	90	14	16	low	soil	18 (C15-C33)	TvM2
BIOM 026	3	137	25	18	middle	soil	18 (C15-C33)	TvM2
BIOM 028	3	237	33	14	middle	soil	18 (C15-C33)	TvM2
BIOM 032	3	130	6	4	middle	soil	18 (C15-C33)	TvM2
BIOM 038	2	51	5	10	low	soil	18 (C15-C33)	TvM2
BIOM 040	3	1094	156	14	high	soil	18 (C15-C33)	TvM2
BIOM 046	2	185	2	1	middle	soil	18 (C15-C33)	TvM2
BIOM 049	3	400	17	4	high	soil	18 (C15-C33)	TvM2
BIOM 051	2	112	4	3	middle	soil	18 (C15-C33)	TvM2
P4	2	4	0.4	9	low	sediments	16 (C21-C37)	TvM2
P7	3	6	0.2	3	low	sediments	16 (C21-C37)	TvM2
P23	3	12	3	24	low	sediments	16 (C21-C37)	TvM2
P28	2	8	1	15	low	sediments	16 (C21-C37)	TvM2
P29	2	7	1	13	low	sediments	16 (C21-C37)	TvM2
P39	2	4	0.003	0.1	low	sediments	16 (C21-C37)	TvM2
P43	2	5	0.2	5	low	sediments	16 (C21-C37)	TvM2
P40	3	13	1.1	9	low	sediments	16 (C21-C37)	TvM2

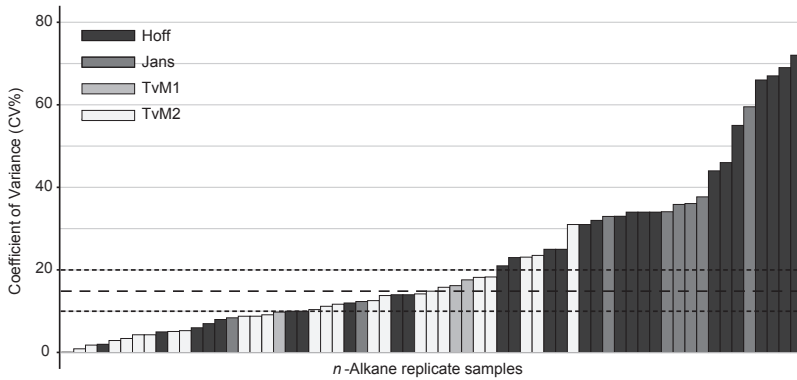


Figure 1 | Replicate sample set coefficient of variation (CV) values, ordered from low to high, per author. The short dashed lines represent the 10% and 20% rule of thumb values considered acceptable variance in literature (Lyu et al., 2017; Teunissen van Manen et al., 2019). The long dashed lines indicate the median CV value of all replicates.

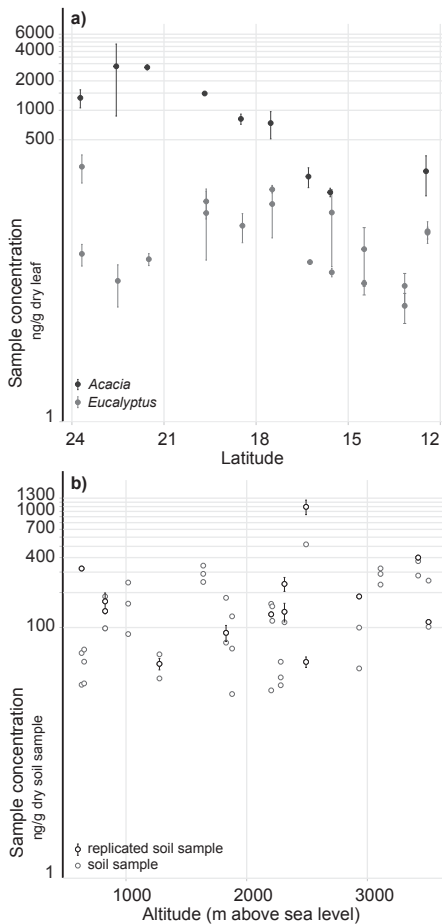


Figure 2 | Effect of measurement variability on the CONw results along the Hoff (a) and TvM2-soil (b) sampling transects. Points are the CONw of the (replicate) sample, lines show measurement variability (SD) of the replicated sample. The Hoff data (a) shows the Acacia (closed black points) and Eucalyptus (closed grey points) \log^{10} n-alkane ng g⁻¹ of dry leaf, along the latitudinal transect. The TvM2-soil data shows the \log^{10} n-alkane ng g⁻¹ of dry soil sample of replicated samples (open black points) and single measurement samples (open grey points), along the altitudinal transect.

The CONw of the TvM2-necromass replicate sample set “14” has an associated CV value of 18% (Table 1). Looking in detail at the absolute concentrations fingerprint, the chain length CV is between 16-35% (median CV 26%) (Figure 3a, Appendix C). In the relative abundances fingerprint, the CV is between 3-28% (median CV 13%) (Figure 3b, Appendix C).

Overall there was no obvious systematic variance in the CV values attributable to the data characteristics recorded, but there were some visible differences between study authors, source material, CONw category, and *n*-alkane chain length range measured (Figure 4). It should be noted that in our study, all categories overlapped extensively so that no true distinction can be made between them (Table 1).

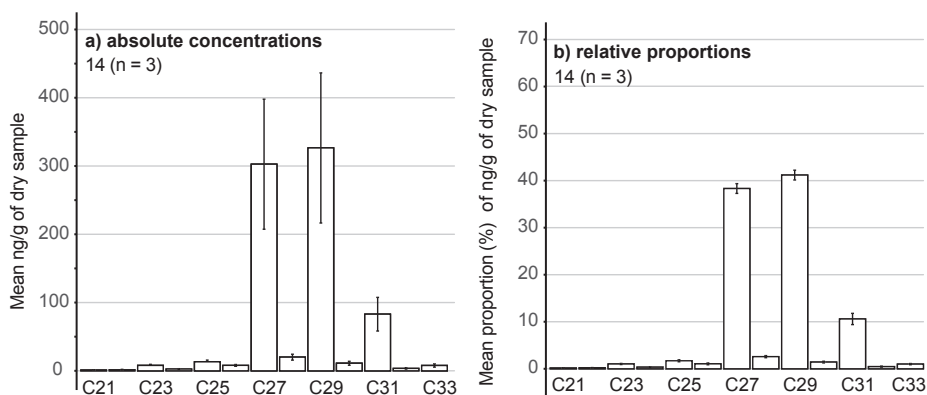


Figure 3 | The absolute (a, ng g⁻¹) and relative (b, %) fingerprints of the replicate sample set 14 (TvM2-necromass). Lines indicate the standard deviation from the mean *n*-alkane concentration (a) or mean *n*-alkane proportion (b).

There were similarities in the distribution of the replicate CV values between authors (Figure 4a). Specifically, replicates from the Hoff study had the widest range of CV values (Figure 4a). Jans replicates had the highest median CV value, whereas the TvM1 and TvM2 replicates had lower median CV compared to the other datasets (Figure 4a).

A visual trend was observed between the CV values and source material (Figure 4b), where generally the median CV of each category decreased as the source material became increasingly degraded (from leaf to sediment). However, it should be noted that the sample size for the necromass and sediment categories was very small (four and eight, respectively) when compared with the leaf and soil categories (30 and 21, respectively), which could incorrectly exaggerate the decreasing trend in CV with sample degradation (Figure 4b).

We observed that the range and median of the replicate set CV values in the high CONw category was wider and higher than in the low and middle categories, for the datasets measured in ngg⁻¹ of dry sample (Figure 4c). This was not reflected in the CV values of CONw expressed in ngg⁻¹ of dry leaf weight, where the range and median of replicate set CV values was similar across CONw categories (Figure 4c).

The shortest range of *n*-alkanes measured (10, C₂₃-C₃₃) had the widest range of replicate set CV values (Figure 4d). The widest range of *n*-alkanes measured (22, C₁₂-C₃₄) had the highest median replicate set CV value. Other *n*-alkane ranges did not differ from each other strongly in replicate set CV range or median (Figure 4d).

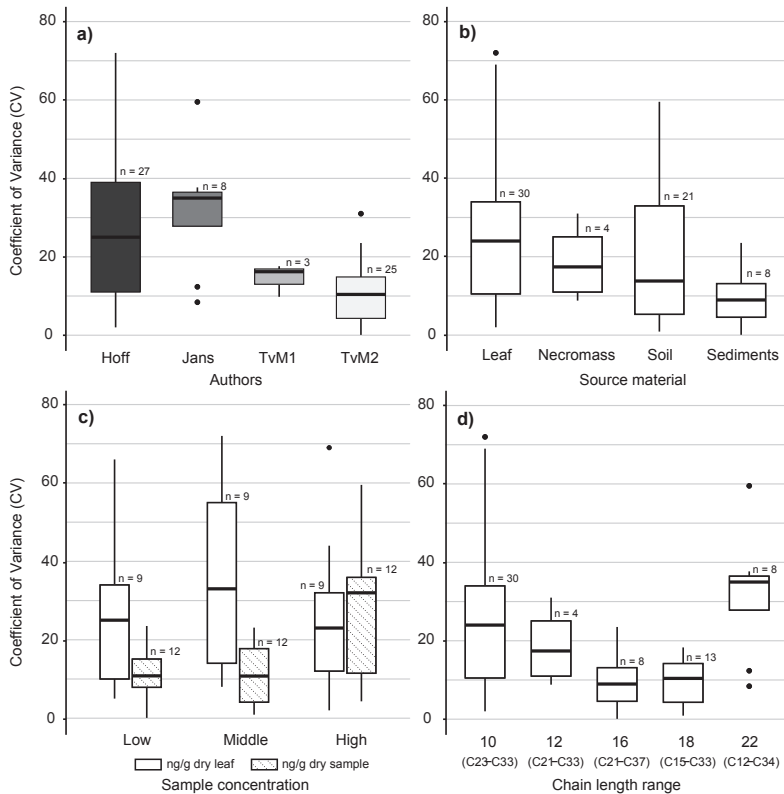


Figure 4 | The distribution of sample replicate set mean concentration (CONw) CV values split across data characteristics: study authors (a), sample source (b), sample set CONw category (c) and, the range of *n*-alkanes measured (d). Boxplots indicate median and interquartile ranges of the CV data. The CONw categories were determined by the quartile ranges of the CONw data, per unit (ng/g dry leaf or ng/g dry sample). Specifically, “low” is the lower quartile, “middle” is the interquartile range and “high” is the upper quartile.

4 | DISCUSSION

4.1 | Replicate sample variability

The reported CV values from the four datasets (Hoff, Jans, TvM1, TvM2) varied widely with a median CV of 15% (Figure 1). There is no consensus in the literature of what comprises an acceptable standard of variance within sample replication, e.g. one study accepted <10% (Teunissen van Manen et al., 2019) and another study accepted <20% (Lyu et al., 2017).

Furthermore, a study that reported sample replication only in figure format, qualitatively assessed the sample replicate variability as “small” (Zhao et al., 2018); however, our estimates of the individual *n*-alkane chain length CV values based on reading data from the figure (mean values and whiskers represented in figure 3 in Zhao et al., 2018) range from 8-75%. Aside from reflecting how flexible the notion of ‘acceptable variability’ is in literature, it also illustrates how hard it is to assess sample replicate variability visually. In the compiled data 65% of the replicate sample sets reported CV values above 10%, and 41% of the replicate sets reported CV values above 20% (Figure 1, Table 1). This demonstrates that the concentration of the sample replicates frequently vary above the rules of thumb of acceptable variability suggested in the literature (Lyu et al., 2017; Teunissen van Manen et al., 2019).

4.2 | Variance in *n*-alkane datasets

To contextualize the observed replicate sample variability we next zoom in on the largest and most extensively replicated two datasets (Hoff and TvM2-soil). Specifically, using these data we: (1) explore the effect of the reported *n*-alkane sample replicate variability on the interpretation of results, and (2) give an indication of what might be considered acceptable variability.

4.2.1 | Effect of variability on results

None of the Hoff and TvM2-soil replicate sample sets varied more than their respective site and transect range, despite some high replicate sample set variability (up to 72%) in the Hoff dataset (Table 1, Appendix 2). In the Hoff and TvM2-soil datasets, the high variability in the measurement did not exceed the studied system variability, suggesting that the measurement was robust enough for the purpose of exploring the relationship between CONw and the environment. In the case of the Hoff dataset, it should be noted that if the replicate set CV values of the *Acacia* data had been consistently high across replicate sets, the measurement would not have been sufficiently accurate to infer a significant relationship between *Acacia* *n*-alkane concentrations and the transect.

Although the CV values of the absolute concentrations reported here suggest low reproducibility of the *n*-alkane measurement, the *n*-alkane relative abundances (i.e. the *n*-alkane patterns) are much more robust (Figure 3, Appendix C). This is illustrated by the TvM2-necromass replicate sample set “14” (Figure 3), where mean *n*-alkane chain length specific concentrations vary widely (between 16-35%), but the relative abundances of each chain length vary less (between 3-24%).

4.2.2 | Defining acceptable variability within *n*-alkane datasets

Overall our results suggest the effect of measurement inconsistencies, and therefore what might be ‘acceptable’ variability, is dependent on the overall variability of the system

studied. Therefore, arguing what type of measurement precision is required for a particular study is more important than comparing with a rule of thumb that has been developed for a different study context. That is not to say that having a rule of thumb for acceptable variability cannot be useful, the subjective choice for a 10% CV seems appropriate (considering it is on the lower end of the median CV values recorded, Figure 1), but the rule of thumb should be validated in the context of any specific study.

4.3 | Potential sources of *n*-alkane measurement variability.

We will next discuss potential sources of the variance in *n*-alkane analysis, and in measurements of absolute concentrations in particular, based on the data characteristics and the literature available.

4.3.1 | Variance between authors

We find that there are differences in the measurement reproducibility between authors, where the Hoff and Jans have higher median CV values than the TvM1 and TvM2 replicate sample sets (Figure 4a). There are many differences between the studies, suggesting that overall, the differences in their approach, sampling and/or protocol led to different measurement reproducibility. One major difference between the studies is that Hoff standardized the *n*-alkane concentration to leaf weight, which could amplify the measurement error of the sample. This is supported by the fact that Hoffmann et al. (2013) also present their measurements standardized to leaf area, which have considerably lower CV values (between 2-21%).

4.3.2 | Variance between sample types

Of the replicate sample data available, the leaf and soil samples had a wide range of CV values (2-72% and 1-60%, respectively), suggesting that successfully replicating the measurement of these samples is hard to do (Figure 4b). This could be due to incomplete sample homogenization. Plant material is known to be highly heterogeneous at the individual and leaf(let) level (Ardenghi et al., 2017; Eglinton and Hamilton, 1967). Similarly, a study on lipid yields extracted from soils found that the extract concentrations varied with grain size and organic matter in the soils (Quénéa et al., 2012), suggesting that sample homogenization can also play an important role in the success of replicating a soil sample. Although the sample sizes in our study are limited, the decreasing median CV from leaf, to necromass, to soils and to sediments suggests that the level of degradation of the *n*-alkane sample (or alternatively the increasing spatiotemporal averaging effect of the sample) has some effect on the reproducibility of the measurement (Figure 2d).

The co-extraction of chlorophyll could be another reason for inconsistent measurement reproducibility that is particular to leaf samples. Literature indicates that the commonly used MeOH/DCM solvent mixture's relative apolarity (MeOH) causes more chlorophyll to extract which can subsequently lead to clogging of the silica gel chromatography columns

(Ardenghi et al., 2017). However, the addition of an internal standard prior to extraction or fractionation should compensate for these irregularities introduced by the sample (as long as the internal standard compound is similar to the *n*-alkanes measured).

4.3.3 | Variance between *n*-alkane concentrations

We hypothesise that the measuring of very large, or very small, amounts of *n*-alkanes could push the systems to the limits of reliable detection and expected the range and median CV for these categories to be wider and higher than the middle category. We see that for the datasets presented in ng/g of dry sample, measurement reliability decreases (i.e. CV range and median increase) as the concentration increases (Figure 4c). However, we do not see this reproduced in the Hoff data (ng/g dry leaf) (Figure 4c). This suggests that measuring very small or very large amounts is likely not intrinsically harder to do consistently.

4.3.4 | Variance between range of measurements

Some literature has indicated that the range of the *n*-alkane measured can affect reproducibility (Aebig et al., 2017; Ardenghi et al., 2017). For example, shorter (C12-C20) and longer (C34-C40) *n*-alkane chain lengths have been found to be measured under detection limit regularly (Ardenghi et al., 2017). Inconsistently detecting these chain lengths can add variance in replicate measurements. Additionally, low molecular weight *n*-alkanes (C19 and below) have been shown to volatilize more easily than middle to high molecular weight *n*-alkanes (C21 and above) (Aebig et al., 2017). The irregularities introduced by measuring particularly short and particularly long *n*-alkanes (either due to detection limit or volatilization problems) could partially explain why the Jans samples, which measured a relatively wide range of *n*-alkanes chain lengths (C12-C34), exhibited a relatively higher median CV value of 35% compared to the other studies (Figure 4d). In contrast, the TVM2-soil sample replicates, which were measured at a slightly narrower range of *n*-alkanes (C15-C33) (Table 1), exhibit a much lower median CV of 10%. However, it seems unlikely that the large difference in reproducibility (CV 35% vs. 10%) is due to a relatively small shift (C13 - C34 vs. C15-C33) in *n*-alkanes measured alone (Figure 4d).

Methodological differences, other than those discussed in the sections above, between the four studies likely influence *n*-alkane measurement reproducibility, which the available data cannot quantify. For example, irregularities of the equipment and GC-MS measurement error. Jansen et al. (2006) found that *n*-alkane extraction amounts using ASE were strongly affected by pressure, it is possible that insufficient system reliability could introduce variance at this step. Similarly, the GC-MS has a measurement error which contributes to variance of the replicate measurement. However, internal standards can compensate for extraction irregularities and equipment measurement error can be tested (for example by repeatedly measuring known *n*-alkane mixtures, see Schemmel et al., 2017; Tao et al., 2017).

4.4 | Implications for the *n*-alkane molecular proxy

We found there is very little data on *n*-alkane replicate measurement variability in the published academic literature. Although the concentration and pattern results of the data presented here were not altered by their respective reported measurement inconsistencies, the robustness of the interpretation of these data can only be assessed in the study specific context of these variance data. Not assessing the replicability of the measurement can impede unifying the *n*-alkane data produced and, in worst case scenarios, could confound interpretations of the *n*-alkane data.

In order for the *n*-alkane molecular proxy research to produce comparable and robust results, it needs a standardized methodology and replication practices. Studies are arising on method optimization, for example studying the effect of temperature, pressure, solvents and extraction methods (Aebig et al., 2017; Ardenghi et al., 2017; Jansen et al., 2006; Quéneá et al., 2012; Wiesenberg et al., 2004). Despite the efforts to improve and standardize methods, testing the reproducibility of the *n*-alkane measurement remains overlooked. We therefore suggest the *n*-alkane molecular proxy research field adopt a best practice of replicating (part of) the studied samples and report the mean concentrations, standard deviation and coefficients of variation in table format (datasheet), so the robustness of the measurement can be explicitly discussed in context of the system variability and compare between studies.

5 | CONCLUSION

Based on our own measurements and the available data in literature, we find that replicated *n*-alkane measurements vary to the extent that low analytical reproducibility is to be expected rather than the exception. Although the low analytical reproducibility did not alter the interpretation of the *n*-alkanes extracted from leaves and soils along the environmental gradients studied (Hoff and TvM2-soil datasets), they do show that acceptable variability is dependent on the study context and on whether the focus of the study is on *n*-alkane absolute concentrations (amounts) or *n*-alkane patterns (composition). We therefore suggest that *n*-alkane measurement reproducibility should always be discussed clearly and explicitly in *n*-alkane molecular proxy studies.

No common characteristic of the data examined was found to explain the observed variability between replicates; however, potential sources could be attributed to low reproducibility of (some steps in) the protocol, incomplete sample homogenization, amplification of the error due to standardization choices, or a combination of both. The apparent low reproducibility of *n*-alkane measurements calls into question our understanding of the factors that contribute to the variability of the *n*-alkane signal, especially the variability in *n*-alkane concentrations. In order to further our understanding of (non-analytical) *n*-alkane variability,

it is essential that the *n*-alkane molecular proxy research field adopts replication practices and invests in developing robust, reproducible protocols. We encourage *n*-alkane biomarker studies to include sample replicates, and to explicitly quantify, report, and discuss the observed sample replicate variability prior to attributing variability to other sources.

6 | DATA AVAILABILITY

All data replicate samples data owned by the authors is available at:
<https://doi.org/10.21942/uva.11537487>

7 | FUNDING SOURCES

This study was funded by the Institute for Biodiversity and Ecosystem Dynamics (IBED) and the University of Amsterdam (UvA).

8 | GLOSSARY

Σ alk: the concentration of the replicate subsample. See also sample replication and replicate sample set.

CONw: mean concentration of the replicate sample set. See also replicate sample set.

CV: coefficient of variation from the replicate samples. In this study it is the standard deviation from the replicate samples CONw, expressed as a percentage of the CONw. See also CONw.

Sample replication: measuring multiple subsets of a homogenized sample, with the aim to test the consistency of the measurement. This requires the sample to be subsampled multiple times.

Replicate sample: A sample that was measured multiple times. See also sample replication, replicate subsample and replicate sample set.

Replicate subsample: One of the subsamples from the homogenized replicate sample.

Replicate sample set: The collection of subsamples taken from the same replicate sample. See also sample replication and replicate sample.

9 | REFERENCES

- Aebig, C.H.F., Curtin, L., Hageman, K.J., Moy, C.M., Beltran, C., 2017. Quantification of low molecular weight *n*-alkanes in lake sediment cores for paleoclimate studies. *Org. Geochem.* 107, 46–53. <https://doi.org/10.1016/j.orggeochem.2017.02.009>
- Ardenghi, N., Mulch, A., Pross, J., Maria Niedermeyer, E., 2017. Leaf wax *n*-alkane extraction: An optimised procedure. *Org. Geochem.* 113, 283–292. <https://doi.org/10.1016/j.orggeochem.2017.08.012>
- Bush, R.T., McInerney, F.A., 2015. Influence of temperature and C4 abundance on *n*-alkane chain length distributions across the central USA. *Org. Geochem.* 79, 65–73. <https://doi.org/10.1016/j.orggeochem.2014.12.003>
- Bush, R.T., McInerney, F.A., 2013. Leaf wax *n*-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Acta* 117, 161–179. <https://doi.org/10.1016/j.gca.2013.04.016>
- Crausbay, S., Genderjahn, S., Hotchkiss, S., Sachse, D., Kahmen, A., Arndt, S.K., 2014. Vegetation dynamics at the upper reaches of a tropical montane forest are driven by disturbance over the past 7300 years. *Arctic, Antarct. Alp. Res.* 46, 787–799. <https://doi.org/10.1657/1938-4246-46.4.787>
- Diefendorf, A.F., Freimuth, E.J., 2017. Extracting the most from terrestrial plant-derived *n*-alkyl lipids and their carbon isotopes from the sedimentary record: A review. *Org. Geochem.* 103, 1–21. <https://doi.org/10.1016/j.orggeochem.2016.10.016>
- Eglinton, G., Hamilton, R.J., 1967. Leaf Epicuticular Waxes. *Science*. 156, 1322–1335. <https://doi.org/10.1126/science.156.3780.1322>
- Feakins, S.J., Peters, T., Wu, M.S., Shenkin, A., Salinas, N., Girardin, C.A.J., Bentley, L.P., Blonder, B., Enquist, B.J., Martin, R.E., Asner, G.P., Malhi, Y., 2016. Production of leaf wax *n*-alkanes across a tropical forest elevation transect. *Org. Geochem.* 100, 89–100. <https://doi.org/10.1016/j.orggeochem.2016.07.004>
- Hoffmann, B., Kahmen, A., Cernusak, L.A., Arndt, S.K., Sachse, D., 2013. Abundance and distribution of leaf wax *n*-alkanes in leaves of acacia and eucalyptus trees along a strong humidity gradient in Northern Australia. *Org. Geochem.* 62, 62–67. <https://doi.org/10.1016/j.orggeochem.2013.07.003>
- Jansen, B., Nierop, K.G.J., Kotte, M.C., de Voogt, P., Verstraten, J.M., 2006. The applicability of accelerated solvent extraction (ASE) to extract lipid biomarkers from soils. *Appl. Geochemistry* 21, 1006–1015. <https://doi.org/10.1016/j.apgeochem.2006.02.021>
- Jansen, B., Wiesenberg, G.L.B., 2017. Opportunities and limitations related to the application of plant-derived lipid molecular proxies in soil science. *SOIL* 3, 211–234. <https://doi.org/10.5194/soil-3-211-2017>
- Kirkels, F.M.S.A., Jansen, B., Kalbitz, K., 2013. Consistency of plant-specific *n*-alkane patterns in plaggen ecosystems: A review. *Holocene* 23, 1355–1368. <https://doi.org/10.1177/0959683613486943>
- Koch, K., Ensikat, H.J., 2008. The hydrophobic coatings of plant surfaces: Epicuticular wax crystals and their morphologies, crystallinity and molecular self-assembly. *Micron* 39, 759–772. <https://doi.org/10.1016/j.micron.2007.11.010>
- Liu, J., An, Z., Liu, H., 2018. Leaf wax *n*-alkane distributions across plant types in the central Chinese Loess Plateau. *Org. Geochem.* 125, 260–269. <https://doi.org/10.1016/j.orggeochem.2018.09.006>

- Lyu, Y., Xu, T., Yang, X., Chen, J., Cheng, T., Li, X., 2017. Seasonal contributions to size-resolved *n*-alkanes (C8–C40) in the Shanghai atmosphere from regional anthropogenic activities and terrestrial plant waxes. *Sci. Total Environ.* 579, 1918–1928. <https://doi.org/10.1016/j.scitotenv.2016.11.201>
- Quénéa, K., Mathieu, J., Derenne, S., 2012. Soil lipids from accelerated solvent extraction: Influence of temperature and solvent on extract composition. *Org. Geochem.* 44, 45–52. <https://doi.org/10.1016/j.orggeochem.2011.11.009>
- R Core Team, 2017. R: A language and environment for statistical computing. <https://www.r-project.org/>.
- Schemmel, F., Niedermeyer, E.M., Koutsodendris, A., Pross, J., Fiebig, J., Mulch, A., 2017. Paleohydrological changes in the Eastern Mediterranean region during the early to mid-Holocene recorded in plant wax *n*-alkane distributions and $\delta^{13}\text{C}$ TOC – New data from Tenaghi Philippon, NE Greece. *Org. Geochem.* 110, 100–109. <https://doi.org/10.1016/j.orggeochem.2017.04.006>
- Shepherd, T., Griffiths, D.W., 2006. The effects of stress on plant cuticular waxes. *New Phytol.* 171, 469–499. <https://doi.org/10.1111/j.1469-8137.2006.01826.x>
- Tao, S., Yin, X., Jiao, L., Zhao, S., Chen, L., 2017. Temporal Variability of Source-Specific Solvent-Extractable Organic Compounds in Coastal Aerosols over Xiamen, China. *Atmosphere (Basel)*. 8, 33. <https://doi.org/10.3390/atmos8020033>
- Teunissen van Manen, M.L., Jansen, B., Cuesta, F., León-yáñez, S., William, D., Teunissen van Manen, M.L., Jansen, B., Cuesta, F., León-yáñez, S., Gosling, W.D., 2020. From leaf to soil: *n*-alkane signal preservation, despite degradation along an environmental gradient in the tropical Andes. *Biogeosciences Discuss.* 1–31. <https://doi.org/10.5194/bg-2019-456>
- Teunissen van Manen, M.L., Jansen, B., Cuesta, F., León-Yáñez, S., Gosling, W.D., 2019. Leaf wax *n*-alkane patterns of six tropical montane tree species show species-specific environmental response. *Ecol. Evol.* 9, 9120–9128. <https://doi.org/10.1002/ece3.5458>
- Wiesenberg, G.L.B.B., Schwark, L., Schmidt, M.W.I.I., 2004. Improved automated extraction and separation procedure for soil lipid analyses. *Eur. J. Soil Sci.* 55, 349–356. <https://doi.org/10.1111/j.1351-0754.2004.00601.x>
- Zech, M., Rass, S., Buggle, B., Löscher, M., Zöller, L., 2012. Reconstruction of the late Quaternary paleoenvironments of the Nussloch loess paleosol sequence, Germany, using *n*-alkane biomarkers. *Quat. Res. (United States)* 78, 226–235. <https://doi.org/10.1016/j.yqres.2012.05.006>
- Zech, M., Zech, R., Morrás, H., Moretti, L., Glaser, B., Zech, W., 2009. Late Quaternary environmental changes in Misiones, subtropical NE Argentina, deduced from multi-proxy geochemical analyses in a palaeosol-sediment sequence. *Quat. Int.* 196, 121–136. <https://doi.org/10.1016/j.quaint.2008.06.006>
- Zhao, B., Zhang, Y., Huang, X., Qiu, R., Zhang, Z., Meyers, P.A., 2018. Comparison of *n*-alkane molecular, carbon and hydrogen isotope compositions of different types of plants in the Dajiuhe peatland, central China. *Org. Geochem.* 124, 1–11. <https://doi.org/10.1016/j.orggeochem.2018.07.008>

10 | APPENDICES

Appendix A

In order to survey reported *n*-alkane measurement testing, and sample replication in particular, we searched Web of Science and Science Direct websites. The search commands used were:

- Web of Science: TS=(("Leaf wax" OR "plant wax") AND (*n*-alkane OR *n*-alkanes) AND (biomarkers OR concentration))
- Science Direct website: KEY (("Leaf wax" OR "plant wax") AND (*n*-alkane OR *n*-alkanes) OR (biomarkers OR concentration))

Both searches were done on July 10th 2019. Search results were filtered to cover only the last three years (2017-2019, included). The lists were compared and combined to a total of 74 unique records. We then checked each of the records to assess whether the studies: (1) presented original *n*-alkane data (not a review paper), and (2) reported *n*-alkane data (concentrations, distributions, CPI, ACL, and/or *n*-alkane chain ratios). This resulted in a narrowing of the list to 42 records, which we then checked to see if they reported on any sampling or measurement replication in the materials and methods, and results sections (Table A.1). Studies that had no mention of measurement reliability testing in the methods or results sections (text or figures) were labeled “no”, studies that mentioned some form of measurement reliability testing, but not sample replication, were labeled “external”. Studies that mentioned sample replication were labeled “yes”.

Table A.1 | List of studies included in survey, noted by their short reference and reported measurement replicability.

Replication	Reference
no	Aichner <i>et al.</i> , 2018
no	Andrae <i>et al.</i> , 2019
no	Ardenghi <i>et al.</i> , 2017
external	Badejo <i>et al.</i> , 2017
no	Baker, Routh and Roychoudhury, 2018
no	Balascio <i>et al.</i> , 2018
no	Bliedtner <i>et al.</i> , 2018
no	Collins <i>et al.</i> , 2017
no	Deng and Jia, 2018
external	Denis <i>et al.</i> , 2017
no	Dunne <i>et al.</i> , 2016
no	Freimuth <i>et al.</i> , 2019
no	Griepentrog <i>et al.</i> , 2019

Table A.1 | (continued)

Replication	Reference
no	Häggi <i>et al.</i> , 2019
no	Howard <i>et al.</i> , 2018
no	Huang <i>et al.</i> , 2018
no	G. Li <i>et al.</i> , 2018
no	Y. Li <i>et al.</i> , 2018
no	Liu <i>et al.</i> , 2017
no	Liu and An, 2018
yes	Lyu <i>et al.</i> , 2017
no	Makou <i>et al.</i> , 2018
no	Nelson <i>et al.</i> , 2017
no	Nelson <i>et al.</i> , 2018
no	Norström <i>et al.</i> , 2017
no	Pu <i>et al.</i> , 2018
no	Li <i>et al.</i> , 2017
external	Ruan <i>et al.</i> , 2018
external	Schemmel <i>et al.</i> , 2017
no	Schreuder <i>et al.</i> , 2018
no	Simoneit <i>et al.</i> , 2017
no	Suh and Diefendorf, 2018
external	Tao <i>et al.</i> , 2017
yes	Teunissen van Manen <i>et al.</i> , 2019
no	Trigui <i>et al.</i> , 2019
no	van den Bos <i>et al.</i> , 2018
no	Wang, Eley, <i>et al.</i> , 2017
no	Wang, Hren, <i>et al.</i> , 2017
external	Wang, Axia, <i>et al.</i> , 2018
no	Wang, Xu, <i>et al.</i> , 2018
no	Wu, West and Feakins, 2019
yes	Zhao <i>et al.</i> , 2018
42	total
33	no
6	external
3	yes

REFERENCES

- Aichner, B., Ott, F., Slowiński, M., Norýskiewicz, A.M., Brauer, A., Sachse, D., 2018. Leaf wax *n*-alkane distributions record ecological changes during the Younger Dryas at Trzechowskie paleolake (northern Poland) without temporal delay. *Clim. Past* 14, 1607–1624. <https://doi.org/10.5194/cp-14-1607-2018>
- Andrae, J.W., McInerney, F.A., Tibby, J., Henderson, A.C.G., Hall, P.A., Marshall, J.C., McGregor, G.B., Barr, C., Greenway, M., 2019. Variation in leaf wax *n*-alkane characteristics with climate in the broad-leaved paperbark (*Melaleuca quinquenervia*). *Org. Geochem.* 130, 33–42. <https://doi.org/10.1016/j.orggeochem.2019.02.004>
- Ardenghi, N., Mulch, A., Pross, J., Maria Niedermeyer, E., 2017. Leaf wax *n*-alkane extraction: An optimised procedure. *Org. Geochem.* 113, 283–292. <https://doi.org/10.1016/j.orggeochem.2017.08.012>
- Badejo, A.O., Hyun, S., Kim, W., Ju, S.J., Song, B., 2017. Sources and distribution of allochthonous organic matter in surface sediment from the Seomjin River to the southern inner shelf of Korea. *Ocean Sci. J.* 52, 489–499. <https://doi.org/10.1007/s12601-017-0052-1>
- Baker, A., Routh, J., Roychoudhury, A.N., 2018. *n*-Alkan-2-one biomarkers as a proxy for palaeoclimate reconstruction in the Mfabeni fen, South Africa. *Org. Geochem.* 120, 75–85. <https://doi.org/10.1016/j.orggeochem.2018.03.001>
- Balascio, N.L., D'Andrea, W.J., Anderson, R.S., Wickler, S., 2018. Influence of vegetation type on *n*-alkane composition and hydrogen isotope values from a high latitude ombrotrophic bog. *Org. Geochem.* 121, 48–57. <https://doi.org/10.1016/j.orggeochem.2018.03.008>
- Bliedner, M., Schäfer, I.K., Zech, R., Von Suchodoletz, H., 2018. Leaf wax *n*-alkanes in modern plants and topsoils from eastern Georgia (Caucasus) - Implications for reconstructing regional paleovegetation. *Biogeosciences* 15, 3927–3936. <https://doi.org/10.5194/bg-15-3927-2018>
- Collins, J.A., Carr, A.S., Schefuß, E., Boom, A., Sealy, J., 2017. Investigation of organic matter and biomarkers from Diepkloof Rock Shelter, South Africa: Insights into Middle Stone Age site usage and palaeoclimate. *J. Archaeol. Sci.* 85, 51–65. <https://doi.org/10.1016/j.jas.2017.06.011>
- Deng, L., Jia, G., 2018. High-relief topography of the Nima basin in central Tibetan Plateau during the mid-Cenozoic time. *Chem. Geol.* 493, 199–209. <https://doi.org/10.1016/j.chemgeo.2018.05.041>
- Denis, E.H., Pedentchouk, N., Schouten, S., Pagani, M., Freeman, K.H., 2017. Fire and ecosystem change in the Arctic across the Paleocene–Eocene Thermal Maximum. *Earth Planet. Sci. Lett.* 467, 149–156. <https://doi.org/10.1016/j.epsl.2017.03.021>
- Dunne, J., Mercuri, A.M., Evershed, R.P., Bruni, S., Di Lernia, S., 2016. Earliest direct evidence of plant processing in prehistoric Saharan pottery. *Nat. Plants* 3, 1–6. <https://doi.org/10.1038/nplants.2016.194>
- Freimuth, E.J., Diefendorf, A.F., Lowell, T. V., Wiles, G.C., 2019. Sedimentary *n*-alkanes and *n*-alkanoic acids in a temperate bog are biased toward woody plants. *Org. Geochem.* 128, 94–107. <https://doi.org/10.1016/j.orggeochem.2019.01.006>
- Griepentrog, M., De Wispelaere, L., Bauters, M., Bodé, S., Hemp, A., Verschuren, D., Boeckx, P., 2019. Influence of plant growth form, habitat and season on leaf-wax *n*-alkane hydrogen-isotopic signatures in equatorial East Africa. *Geochim. Cosmochim. Acta* 263, 122–139. <https://doi.org/10.1016/j.gca.2019.08.004>
- Hägg, C., Eglinton, T.I., Zech, W., Sosin, P., Zech, R., 2019. A 250 ka leaf-wax δD record from a loess section in Darai Kalon, Southern Tajikistan. *Quat. Sci. Rev.* 208, 118–128. <https://doi.org/10.1016/j.quascirev.2019.01.019>

- Howard, S., McInerney, F.A., Caddy-Retalic, S., Hall, P.A., Andrae, J.W., 2018. Modelling leaf wax *n*-alkane inputs to soils along a latitudinal transect across Australia. *Org. Geochem.* 121, 126–137. <https://doi.org/10.1016/j.orggeochem.2018.03.013>
- Huang, X., Pancost, R.D., Xue, J., Gu, Y., Evershed, R.P., Xie, S., 2018. Response of carbon cycle to drier conditions in the mid-Holocene in central China. *Nat. Commun.* 9, 1–9. <https://doi.org/10.1038/s41467-018-03804-w>
- Li, G., Li, L., Tarozo, R., Longo, W.M., Wang, K.J., Dong, H., Huang, Y., 2018. Microbial production of long-chain *n*-alkanes: Implication for interpreting sedimentary leaf wax signals. *Org. Geochem.* 115, 24–31. <https://doi.org/10.1016/j.orggeochem.2017.10.005>
- Li, R., Fan, J., Xue, J., Meyers, P.A., 2017. Effects of early diagenesis on molecular distributions and carbon isotopic compositions of leaf wax long chain biomarker *n*-alkanes: Comparison of two one-year-long burial experiments. *Org. Geochem.* 104, 8–18. <https://doi.org/10.1016/j.orggeochem.2016.11.006>
- Li, Y., Yang, S., Xiao, J., Jiang, W., Yang, X., 2018. Hydrogen isotope ratios of leaf wax *n*-alkanes in loess and floodplain deposits in northern China since the Last Glacial Maximum and their paleoclimatic significance. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 509, 91–97. <https://doi.org/10.1016/j.palaeo.2017.08.009>
- Liu, J., An, Z., 2018. A hierarchical framework for disentangling different controls on leaf wax δD *n*-alkane values in terrestrial higher plants. *Quat. Sci. Rev.* 201, 409–417. <https://doi.org/10.1016/j.quascirev.2018.10.026>
- Liu, X., Feakins, S.J., Dong, X., Xue, Q., Marek, T., Leskovar, D.I., Neely, C.B., Ibrahim, A.M.H., 2017. Experimental study of leaf wax *n*-alkane response in winter wheat cultivars to drought conditions. *Org. Geochem.* 113, 210–223. <https://doi.org/10.1016/j.orggeochem.2017.07.020>
- Lyu, Y., Xu, T., Yang, X., Chen, J., Cheng, T., Li, X., 2017. Seasonal contributions to size-resolved *n*-alkanes (C8–C40) in the Shanghai atmosphere from regional anthropogenic activities and terrestrial plant waxes. *Sci. Total Environ.* 579, 1918–1928. <https://doi.org/10.1016/j.scitotenv.2016.11.201>
- Makou, M., Eglinton, T., McIntyre, C., Montluçon, D., Antheaume, I., Grossi, V., 2018. Plant Wax *n*-Alkane and *n*-Alkanoic Acid Signatures Overprinted by Microbial Contributions and Old Carbon in Meromictic Lake Sediments. *Geophys. Res. Lett.* 45, 1049–1057. <https://doi.org/10.1002/2017GL076211>
- Nelson, D.B., Knohl, A., Sachse, D., Schefuß, E., Kahmen, A., 2017. Sources and abundances of leaf waxes in aerosols in central Europe. *Geochim. Cosmochim. Acta* 198, 299–314. <https://doi.org/10.1016/j.gca.2016.11.018>
- Nelson, D.B., Ladd, S.N., Schubert, C.J., Kahmen, A., 2018. Rapid atmospheric transport and large-scale deposition of recently synthesized plant waxes. *Geochim. Cosmochim. Acta* 222, 599–617. <https://doi.org/10.1016/j.gca.2017.11.018>
- Norström, E., Katrantsiotis, C., Smittenberg, R.H., Kouli, K., 2017. Chemotaxonomy in some Mediterranean plants and implications for fossil biomarker records. *Geochim. Cosmochim. Acta* 219, 96–110. <https://doi.org/10.1016/j.gca.2017.09.029>
- Pu, Y., Cao, J.C., Jia, J.H., Shao, X.H., Han, Y., 2018. Unusual hydrocarbon waxes detected in *Salix oritrepha* leaf from Nianbaoyeze Mountains, eastern Qinghai-Tibetan Plateau. *J. Mt. Sci.* 15, 2445–2452. <https://doi.org/10.1007/s11629-017-4646-y>
- Ruan, X., Ye, H., Lu, Q., Yang, Z., 2018. Historical record, source, and toxicity assessment of sedimentary organic matter using molecular composition of hydrocarbons in an urban lake, Wuhan, China. *J. Soils Sediments* 18, 2081–2092. <https://doi.org/10.1007/s11368-018-1956-9>
- Schemmel, F., Niedermeyer, E.M., Koutsodendris, A., Pross, J., Fiebig, J., Mulch, A., 2017. Paleohydrological changes in the Eastern Mediterranean region during the early to mid-Holocene recorded in plant wax

- n*-alkane distributions and $\delta^{13}\text{C}$ TOC – New data from Tenaghi Philippon, NE Greece. *Org. Geochem.* 110, 100–109. <https://doi.org/10.1016/j.orggeochem.2017.04.006>
- Schreuder, L.T., Stuut, J.B.W., Korte, L.F., Sinnighe Damsté, J.S., Schouten, S., 2018. Aeolian transport and deposition of plant wax *n*-alkanes across the tropical North Atlantic Ocean. *Org. Geochem.* 115, 113–123. <https://doi.org/10.1016/j.orggeochem.2017.10.010>
- Simoneit, B.R.T., Pisani, O., Ekpo, B.O., Fubara, E.P., Nna, P.J., Ekpa, O.D., 2017. Lipid Biomarker Analysis of Suspended Particulate Matter from the Great Kwa River, SE Nigeria: Origins and Environmental Implications of Biogenic and Anthropogenic Organic Compounds. *Aquat. Geochemistry* 23, 89–108. <https://doi.org/10.1007/s10498-017-9311-0>
- Suh, Y.J., Diefendorf, A.F., 2018. Seasonal and canopy height variation in *n*-alkanes and their carbon isotopes in a temperate forest. *Org. Geochem.* 116, 23–34. <https://doi.org/10.1016/j.orggeochem.2017.10.015>
- Tao, S., Yin, X., Jiao, L., Zhao, S., Chen, L., 2017. Temporal Variability of Source-Specific Solvent-Extractable Organic Compounds in Coastal Aerosols over Xiamen, China. *Atmosphere (Basel)*. 8, 33. <https://doi.org/10.3390/atmos8020033>
- Teunissen van Manen, M.L., Jansen, B., Cuesta, F., León-Yáñez, S., Gosling, W.D., 2019. Leaf wax *n*-alkane patterns of six tropical montane tree species show species-specific environmental response. *Ecol. Evol.* 9, 9120–9128. <https://doi.org/10.1002/ece3.5458>
- Trigui, Y., Wolf, D., Sahakyan, L., Hovakimyan, H., Sahakyan, K., Zech, R., Fuchs, M., Wolpert, T., Zech, M., Faust, D., 2019. First Calibration and Application of Leaf Wax *n*-Alkane Biomarkers in Loess-Paleosol Sequences and Modern Plants and Soils in Armenia. *Geosciences* 9, 263. <https://doi.org/10.3390/geosciences9060263>
- van den Bos, V., Engels, S., Bohncke, S.J.P., Cerli, C., Jansen, B., Kalbitz, K., Peterse, F., Renssen, H., Sachse, D., 2018. Late Holocene changes in vegetation and atmospheric circulation at Lake Uddelermeer (The Netherlands) reconstructed using lipid biomarkers and compound-specific δD analysis. *J. Quat. Sci.* 33, 100–111. <https://doi.org/10.1002/jqs.3006>
- Wang, C., Eley, Y., Oakes, A., Hren, M., 2017a. Hydrogen isotope and molecular alteration of *n*-alkanes during heating in open and closed systems. *Org. Geochem.* 112, 47–58. <https://doi.org/10.1016/j.orggeochem.2017.07.006>
- Wang, C., Hren, M.T., Hoke, G.D., Liu-Zeng, J., Garziona, C.N., 2017b. Soil *n*-alkane δD and glycerol dialkyl glycerol tetraether (GDGT) distributions along an altitudinal transect from southwest China: Evaluating organic molecular proxies for paleoclimate and paleoelevation. *Org. Geochem.* 107, 21–32. <https://doi.org/10.1016/j.orggeochem.2017.01.006>
- Wang, J., Axia, E., Xu, Y., Wang, G., Zhou, L., Jia, Y., Chen, Z., Li, J., 2018a. Temperature effect on abundance and distribution of leaf wax *n*-alkanes across a temperature gradient along the 400 mm isohyet in China. *Org. Geochem.* 120, 31–41. <https://doi.org/10.1016/j.orggeochem.2018.03.009>
- Wang, J., Xu, Y., Zhou, L., Shi, M., Axia, E., Jia, Y., Chen, Z., Li, J., Wang, G., 2018b. Disentangling temperature effects on leaf wax *n*-alkane traits and carbon isotopic composition from phylogeny and precipitation. *Org. Geochem.* 126, 13–22. <https://doi.org/10.1016/j.orggeochem.2018.10.008>
- Wu, M.S., West, A.J., Feakins, S.J., 2019. Tropical soil profiles reveal the fate of plant wax biomarkers during soil storage. *Org. Geochem.* 128, 1–15. <https://doi.org/10.1016/j.orggeochem.2018.12.011>
- Zhao, B., Zhang, Y., Huang, X., Qiu, R., Zhang, Z., Meyers, P.A., 2018. Comparison of *n*-alkane molecular, carbon and hydrogen isotope compositions of different types of plants in the Dajiuhu peatland, central China. *Org. Geochem.* 124, 1–11. <https://doi.org/10.1016/j.orggeochem.2018.07.008>

Appendix B

Table B.1 | Site and transect variability, calculated for the Hoff and TvM2-soil datasets. Per site (site name) and transect the sample source, mean concentration (CONw), standard deviation (SD), coefficient of variation (CV) and number of samples per site or transect (N). CONw is given in ng g⁻¹ dry leaf weight in the Hoff dataset and in ng g⁻¹ dry soil sample weight in the TvM2-soil dataset.

[next page]

Site/transect	Sample	CONw	SD	CV	N	Dataset
Alice	<i>Acacia</i>	1340	-	-	1	Hoff
Barrow	<i>Acacia</i>	2752	-	-	1	Hoff
Daly	<i>Acacia</i>	210	-	-	1	Hoff
Darwin	<i>Acacia</i>	238	-	-	1	Hoff
Elliott	<i>Acacia</i>	740	-	-	1	Hoff
Helen	<i>Acacia</i>	817	-	-	1	Hoff
Larrimah	<i>Acacia</i>	145	-	-	1	Hoff
Tennant	<i>Acacia</i>	1486	-	-	1	Hoff
Territory	<i>Acacia</i>	2818	-	-	1	Hoff
Darwin	<i>Eucalyptus*</i>	57	1	3	2	Hoff
Tennant	<i>Eucalyptus*</i>	103	20	19	2	Hoff
Elliott	<i>Eucalyptus*</i>	133	32	24	2	Hoff
Adelaide	<i>Eucalyptus*</i>	13	4	33	2	Hoff
Katherine	<i>Eucalyptus*</i>	28	15	54	2	Hoff
Larrimah	<i>Eucalyptus*</i>	56	48	86	2	Hoff
Alice	<i>Eucalyptus*</i>	150	163	109	2	Hoff
Barrow	<i>Eucalyptus*</i>	30	-	-	1	Hoff
Daly	<i>Eucalyptus*</i>	28	-	-	1	Hoff
Helen	<i>Eucalyptus*</i>	66	-	-	1	Hoff
Territory	<i>Eucalyptus*</i>	18	-	-	1	Hoff
Transect	<i>Acacia</i>	1172	1031	88	9	Hoff
Transect	<i>Eucalyptus*</i>	68	64	95	18	Hoff
VERD_03	soil	282	45	16	3	TvM2
RIBR_01	soil	293	47	16	3	TvM2
VERD_01	soil	351	63	18	3	TvM2
CEDR_03	soil	133	27	20	2	TvM2
MIND_01	soil	48	11	23	3	TvM2
BECL_02	soil	40	10	24	3	TvM2
MALO_01	soil	147	38	26	4	TvM2
MAPI_01	soil	49	16	32	3	TvM2
BECL_01	soil	162	67	41	3	TvM2

Table B.1 | (continued)

Site/transect	Sample	CONw	SD	CV	N	Dataset
MALO_02	soil	164	78	48	3	TvM2
INTI_02	soil	115	57	50	3	TvM2
YANA_01	soil	156	85	55	3	TvM2
VERD_02	soil	110	71	65	3	TvM2
BECL_03	soil	106	69	65	3	TvM2
INTI_01	soil	72	49	68	3	TvM2
CEDR_01	soil	555	523	94	3	TvM2
MAPI_02	soil	138	160	116	3	TvM2
Transect	soil	172	174	101	51	TvM2

**following Hoffmann et al (2013), includes Corymbia and Eucalyptus species*

Appendix C

Table C.1 | The absolute (**a**, ng g⁻¹ of dry necromass sample) and relative (**b**, %) fingerprints of the replicate sample set 14 (TvM2-necromass), visualized in Figure 3. Per unit and per chain length, the mean, standard deviation (SD) and coefficient of variation (CV, %) are given. Values in bold are above 20% rule of thumb considered acceptable measurement variability.

Alkane	a) Absolute (ng g ⁻¹)			b) Relative (% of ng g ⁻¹)		
	Mean	SD	CV	Mean	SD	CV
C21	1	0.2	20	0.2	0.04	24
C22	2	0.6	35	0.2	0.06	28
C23	8	1.6	20	1.0	0.11	11
C24	3	0.5	21	0.3	0.08	22
C25	13	2.5	19	1.7	0.21	12
C26	8	1.2	16	1.0	0.19	18
C27	303	95.2	31	38.3	1.04	3
C28	20	4.3	22	2.6	0.22	9
C29	326	109.9	34	41.2	1.04	3
C30	11	2.9	26	1.4	0.18	13
C31	83	24.6	30	10.6	1.19	11
C32	4	1.0	28	0.5	0.10	22
C33	8	2.4	30	1.0	0.13	13