COMMENTS FROM EDITORS AND REVIEWERS

Editor:

first of all, I want to apologize for the very long delay in handling this manuscript. It is often very difficult to find suitable reviewers, because less and less colleagues feel the responsibility to take part in the peer review process. Everyone wants to publish in peer-reviewed journals, but nowadays only few are willing to share their expertise and write a review. In many cases we need to approach at least a dozen of colleagues in order to get two reviews. When it is a short purely descriptive manuscript in a field I feel confident, I sometimes rely on one review, but in case of more complex and/or potentially controversial manuscripts I definitely need two reviews of people who are well familiar with the topic. None of the reviewers suggested by the authors took the job. Some never reacted on our request, others had no time, and one colleague promised to write a review and even noticed that it was an interesting paper but never sent us a review despite several reminders. Also a couple of other colleagues approached by me did not want to write a review.

This is a very important manuscript that critically discusses a method that is very widely used. I clearly see the weaknesses of the NRL method and I have always been very sceptical about it, just like the use of so-called "stomata proxies" in the Palaeozoic and Mesozoic. However, in this case I did not feel confident enough to rely on a single review. Eventually, I found two people who have the right expertise and who were willing to do it. I am very pleased that both reviewers recommended publication after minor revision. For their comments I refer to the reviews and the annotated manuscript in the attachment.

Although I do not want to criticize the contents of this manuscript, there are some points that I would like to mention. In many places the phrasing is very ironic or even sarcastic, and words and phrases are printed in italics or in bold. Using different font types in a running text is not the style of the journal; only headers are printed in bold typeface and Latin names in italics. As it is presented here, it is written a very polemic style. This was a style of writing that was used in the 19th century but that is not longer used today. In my opinion this seriously weakens the impact of the paper.

Response E1: We removed all italics and bold font and went over the manuscript, further we gave it to a colleague not involved in the topic at all, to check for personally offensive phrases. Given that CA+PF is promoted solely using methods from the 15th century (intimidation, scorning of critics, restricted access to knowledge, ignorance against hard evidence, rejection of any development, principle of infallibility), a 19th century-style language is more than appropriate. What we write is simply to the point. Polemic would have been if we would have written that the Quan et al. (2012) and Utescher et al. (2014) papers demonstrate that the self-declared climate expert and Psychopomp of NECLIME Torsten Utescher must be mentally insane and institutionalised, that the director of the Senckenberg Gesellschaft and High Chancellor of NECLIME Volker Mosbrugger is bigot and corrupt, promoting an obviously flawed and dubiously applied method (he realised that 17 years ago), and that the NECLIME members who have their data studied by the Members of the Inner Circle and blindly accept the results are a flock of unwitting (or even dumb) sheep. And that all editors and reviewers who have contributed to the uncontrolled publication of CA+PF should be ashamed for the establishment of a fraudulent pseudoscience, which ridicules palaeobotanists throughout the scientific world. Instead, we kept to the facts, which are embarrassing enough.

It is much better to give a straight-forward presentation of the data and the conclusions; they are strong enough and speak for themselves, and that is what they should do. Any remark that the NLR method was "successfully" used during the past fifteen years is redundant and really harms the impact of this paper. From the references it is clear since when the NLR method is used and by whom. I also urge not to criticize persons. Their work should be criticized not the people who wrote it. In this case, being too direct can really affect the intention of this paper and harm your own reputation. Readers should not have a reason to think that that this is just a personal feud between different researcher groups, but it should be a sound scientific discussion. Let the facts speak for themselves. That is the strongest possible criticism that can be given!

Response E2: We removed “successfully”. Nevertheless, the point of this paper is to show that not only the method is problematic in principle (Grimm and Potts, 2016a), but more so how it has been (and apparently still is) applied. The Quan et al. supplements provide a unique glimpse behind the heavy curtains of CA+PF studies and reveal nothing less than either incompetence of leading CA+PF figures or a tradition of fraudulent data manipulation. This is personal failure, nothing else. Only the authors, potentially the reviewers, and ultimately the editors of CA+PF papers are responsible for the foibles in the application and the unconditionally acceptance of the fabled results of CA+PF studies. Problems such as poor taxonomic control, poor NLR tolerance data, dubious filtering of taxa that are demonstrated in the Quan et al. study cannot be blamed on the general inadequacy of the Coexistence Approach.

The fact that there has been no CA+PF study on North American (where best-possible bioclimatic data are long available) and Japanese floras demonstrate that the method only still exists because of political (i.e. personal), and not scientific reasons. The way how our 2012 paper is cited in Utescher et al. (2014) and their unwillingness to use the open review opportunity to defend their views (see also Grimm, 2015, and the authors' response to this interactive comment), prove further the syndicate’s medieval attitude. Our motivation is not to start a discussion (see also **Response R18)** on a fundamentally flawed (Grimm and Potts, 2016a), uninformative (Klotz, 1999; Grimm, 2015), and outdated (Anonymous Referee #2, 2016) method, but solely to give people a heavy club in the hand to batter down the pseudo-ivory tower of CA. Obviously the palaeobotanical community is afraid of openly criticising the CA, so we “sacrificed” ourselves. If the king is naked, somebody has to just say it.

Once again, I really like to see this important manuscript in print, but please keep the remarks given above and the comments of the reviewers in mind during the final revision. IO am looking forward to the final version.

Response E3: We hope the revised version is more acceptable. The throughout fine reviewer comments have been largely incorporated. An according appreciation has been included in the Acknowledgements.

Reviewer #1:

I found this is a complex manuscript as it needs to steer a fine-line between a critique of three things: a method, the method and its central database, and a particularly sloppy instance of applying that method and database.

So-saying, I have little problem with the manuscript overall - it's an excellent example of how science should be. The critique is justified, the topic is worthy of publication in the journal and is sure to get plenty of attention.

Response R1. Thanks.

I have raised a few minor points and some broader points that should be addressed.

Introduction

Line 55- "Here we examine whether the Coexistence Approach is scientifically sound in terms of its methodological approach; in a separate study we consider the theoretical underpinnings of the Coexistence Approach and other methods based on the mutual climate range coupled with nearest living relative associations"

I assume this can now be referenced to Grimm and Potts (2015). Suggest re-order to something like:

"In a previous study (Grimm and Potts, 2015) we considered the theoretical underpinnings of the Coexistence Approach and other methods based on the mutual climate range coupled with nearest living relative associations. In the present study we examine whether the Coexistence Approach is scientifically sound in terms of its methodological approach."

Response R2. Updated. We did not write "*previous study*" because that study is a follow-up and was conceived based on what we show here (hence, submitted nine months later).

Subjective bias

line 367 MATs … "more likely due to different observation periods"

This is doing my head in, but I don't have access to the data to test it. The claim is that a difference of 20 years would change an average MAT by 5C? The MATs of those 20 years would have to be fantastically different from the other 30 to do that. Mmmm?

Response R3. Good point. Hijman et al. note that they included also records with unknown observation periods, following the "*that in most cases these records will represent the 1950-2000 time period, and that insufficient capture of spatial variation is likely to be a larger source of error than in high resolution surfaces than than effects climatic change during the past 50 years*"; their primary sample included all records with more than 10 years (http://www.worldclim.org/methods); i.e. we have a potential difference of observation periods of max. ± 20 years. Location mismatch may be another reason (e.g. altitude difference between the grid cell centre and that of the climate station within the grid cell). We conferred with R. Hijmans on the topic. Neither he nor we have access to the station data used for Palaeoflora, but reckon that their station coverage is hardly larger (or more representative) than the trimming data used for the WorldClim climate surfaces (Fig. R1, R2). He confirmed location mismatch as a likely explanation. We rewrote the first part of section 3.2 accordingly and added a sentence to the end stating that with respect to the subjectivity of hand-picking 4-6 extreme climate stations and the data basis (1200 stations vs. >25,000 precipitation and temperature records), best-possible tolerance data can only be obtained using the extrapolated climate surface.

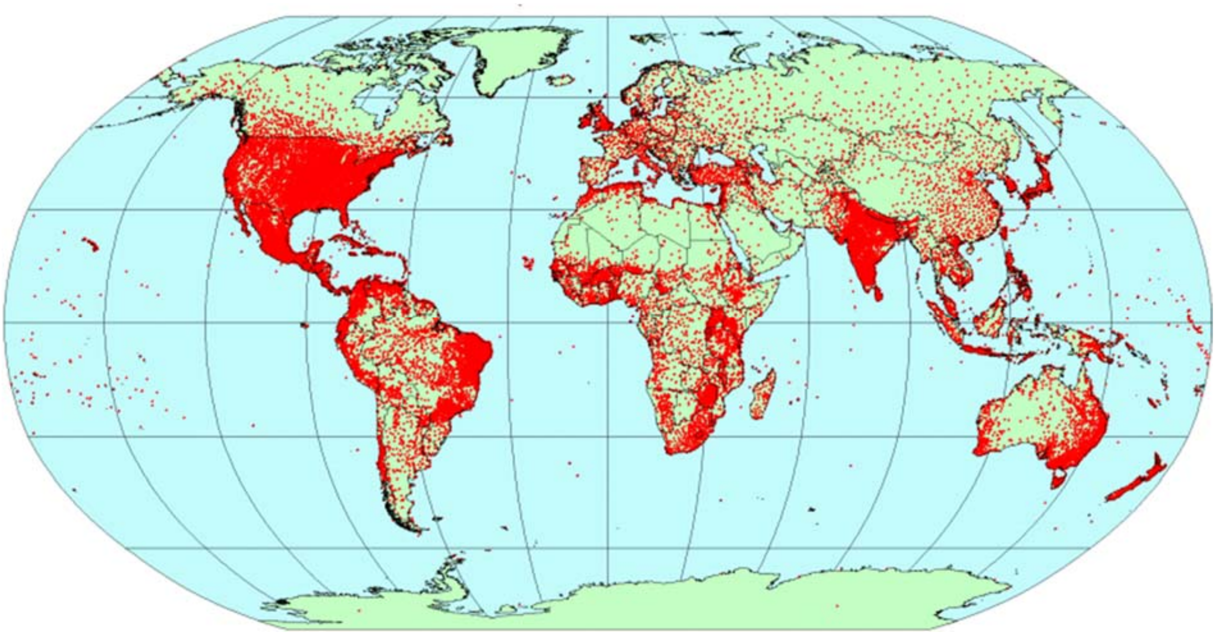

Fig. R1. Locations of the 47,554 climate stations with precipitation data used for the WorldClim climate surface (map from http://www.worldclim.org/methods)

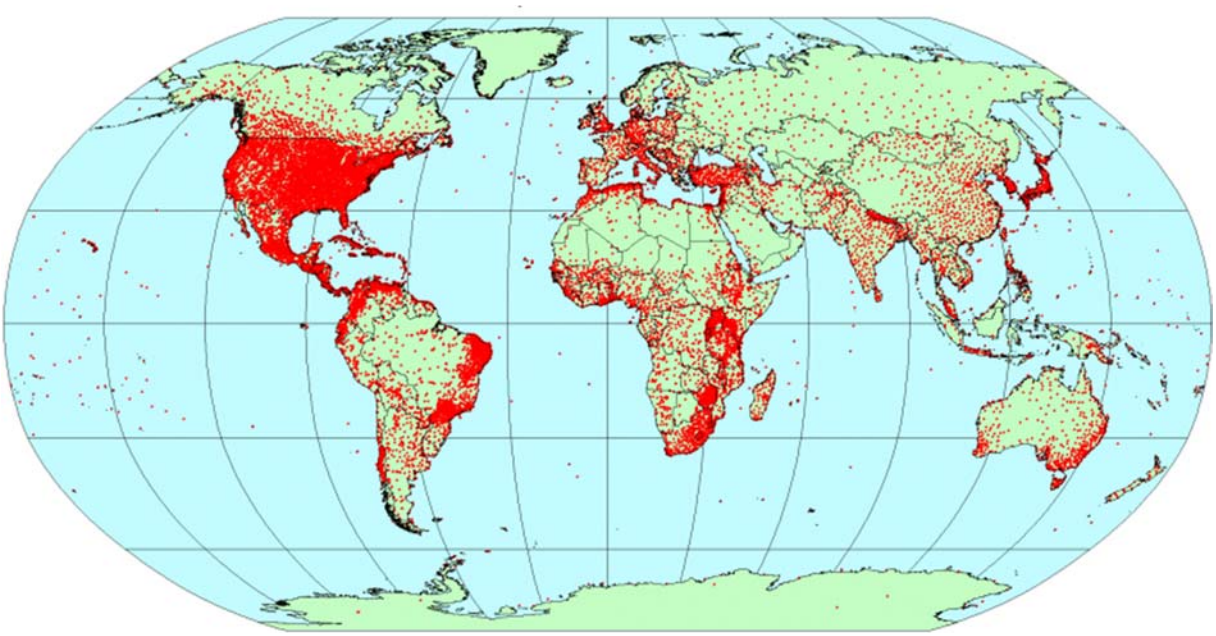

Fig. R2. Locations of the 24,542 climate stations with mean temperature data used for the WorldClim climate surface (map from http://www.worldclim.org/methods)

Line 463 What is a relict

Ignoring relicts is one guideline for CA - but I notice that even though Quan et al. Appendix B said they also removed cosmopolitan and aquatic taxa from their database, Potamogeton, Typha and Sphagnaceae are still there.

Response R4. We added the information about this inconsistent treatment (a common attribute of CA+PF studies) at the end of the section.

CA+PF papers never elaborate on why taxa were excluded or why certain monotypic genera, relicts with restricted modern distribution areas, and aquatics are included. CA+PF reconstructions can only be done in form of “joint research” with T. Utescher or another core member of NECLIME (see information on www.palaeoflora.de). Utescher informed us in 2010 that the procedure is as follows: one sends him the taxon lists, he sends the result back. Decision about in-or-out seems to be made on an assemblage-to-assemblage or study-to-study basis. A likely motivation is to keep the number of NLR > 10 and to obtain narrower intervals, see e.g. the “ohne Engel” [= without *Engelhardia*] results on “sheet 1” in the supplement of Hoorn et al. (2012). A mirrored copy can be downloaded under the following link: **http://www.palaeogrimm.org/themen/Hoorn_et_al/Hoorn_et_al_Supplement1_commented.xls**

Cyathidaceae

Li 513 In the case of Cyatheaceae, a highly diverse group of tree ferns, the association between pollen and nearest-living relatives is unproblematic. Cyathidites/ Cyatheaceae pollen is ….

Actually no - most workers would see the affinities of Cyathidites more broadly - to at least the Cyathaceae/Dicksoniaceae (e.g. Wagstaff et al. 2013) and some would add Schizaeaceae (Lygodium) and others as well (Raine, D.C. Mildenhall, E.M. Kennedy (2011). New Zealand fossil spores and pollen: an illustrated catalogue. 4th edition. GNS Science miscellaneous series no. 4. http://data.gns.cri.nz/sporepollen/index.htm")

Response R5: Has been changed and references were added.

li 514 Cyathidites/ Cyatheaceae pollen is … recorded for 15 assemblages from all three … middle arid ... This was unsurprising regarding the modern … distribution and diversity of the family.

It surprises the Hell out of me to mention Cyatheaceae and 'arid' in the same sentence…. Something very odd there.

Response R6: We included the reviewer’s point as pers. comm.

There are probably more taxa which would conflict with the palaeoclimate zone categorisation provided in Quan et al. (2012; in a later paper the authors drop this categorisation). Since this is no explicit comment to Quan et al., we don’t want to go into the detail, as it would distract from the method-application related issues; which in the case of the *Cyathidites* is even worse than we thought (violation of assumptions 1 [NLR unclear], 2 [formerly arid?, but not anymore], and 4)

li 521 for 'Steward Is' write 'Stewart Is'

Response R7: Corrected

li 522 for 'Ivercargill' write 'Invercargill'

Response R8: Corrected

Alternative Methods

Line 591 I don't see why 'alternative' MCR-NLR techniques - like 'capped MCR', are 'alternatives' to CA., Surely it's a technique to improve the database that CA uses?

Response R9: The main principal problem, the major flaw, of CA is its vulnerability on exotic elements in a palaeoflora and its ignorance towards community information. CA shoehorns any assemblage into coexistence; if NLRs have no mutually shared climate range, some of them, the so-called 'climatic outliers' are eliminated using a simple-counting mechanism that lacks any logical basis (Grimm and Potts, 2016a). The alternatives listed here remove the bias by exotic elements to some degree, from different angles. Improving the database will not by-pass the theoretical issues of the CA (Grimm and Potts, 2016a), but would increase the usability of the database for such alternative MCR-NLR techniques.

Note that using percentiles goes back, at least to:

Kershaw, A.P. and Nix, H.A.1988. Quantitative palaeoclimatic estimates from pollen data using bioclimatic profiles of extant taxa. Journal of Biogeography, 15:589-602.

and

Sluiter, I.R.K., Kershaw, A.P., Holdgate, G.R., and Bulman, D. 1995. Biogeographic, ecological and stratigraphic relationships of the Miocene brown coal

floras, Latrobe Valley, Victoria, Australia. International Journal of Coal Geology,28:277-302.

who ought to get a look in the References.

Response R10. References added.

The South Carolina/Yunnan effect

Li 617 4.3. Do all roads still lead to North Carolina?

Nice subheading, but maybe some further explanation or rationalisation below, as some roads seem to lead to Yunnan. For example, Fig. 6 is referenced under the South Carolina subheading, but its caption quotes 'the "Yunnan effect" in the data of Quan et al. (The Highlights also mention Yunnan).

Response R11. We added Yunnan to the subheading.

Line 656 Trend to Monsoonal-Subtropical

The suggestion that the method tends towards giving an answer that is subtropical and monsoonal is particularly interesting. Partially I agree because I note the same thing – it's actually difficult to get a range of taxa that clearly indicate tropical conditions. What does bother me is their claim that a random selection of the taxa should show something other than subtropical. If one has a large database of tropical to cool temperate taxa (ignoring Arctic or Alpine), then surely, if you averaged it - you would tend to get a value somewhere in-between? So I don't see this as unexpected, and as the proponents of the CA have said somewhere else, a fossil assemblage isn't a random sample. So I think an issue may have been touched on here - just not quiet the right way.

Response R12. We clarified that no reconstruction method should get the same results for a random and non-random sample. We further added the following at the end of the paragraph: "One could argue that a fossil assemblage is never random. This is true, but any fossil assemblage will eventually include random (exotic) elements such as long-dispersed or reworked pollen or misdetermined taxa; and the coexistence interval is determined by few, the most exotic, nearest-living relatives (Fig. 4; Grimm and Potts, 2016a). Our randomisation test shows the more NLRs are included in a flora the higher will be the likelihood that coexistence is reached and a subtropical climate inferred, independent of which elements (random or genuine) constitute the assemblage."

One question - the CA database contains a lot of redundancy - e.g. Morus + Moraceae, Selaginella + Selaginellaceae, etc. So when there was random resampling for this paper, how were these dealt with?

Response R12. Following the CA protocol and practise we did not filter for redundant or inclusive taxa. We added the following information: "Quan et al's list of nearest-living relatives includes redundant (15 genera and higher-taxa with identical tolerances, e.g. *Amaranthus* and Amaranthaceae; *Juglans* and Juglandaceae) and inclusive (c. 15 higher-than-genus taxa with tolerances usually encompassing that of comprised genera included in the dataset; e.g. *Cudrania* [= *Maclura*], *Morus*, and Moraceae; but see also File S4) nearest-living relatives, which were not filtered following the practise of CA+PF studies (e.g. Grimm and Denk, 2012, ES2; Quan et al., 2012, appendix B)."

Morus/Moraceae, which have different tolerances according PF/Quan et al. (the Moraceae tolerances equal however those of *Cudrania* and the artificial NLR "Moraceae/Urticaceae"), would be kept as two NLRs, same for *Selaginella/*Selaginellaceae (one of the hilarious-inconsistently recorded cases; Table R1). This makes little sense regarding the basic logic of MCR techniques, but given their recorded tolerances most of these redundancies have little to no effect on the CA results, or that of any other simple MCR-NLR technique, except for those highlighted in section 4.1 and Box 1. They just inflate the number of "climatically active" NLR (according terminology of some CA+PF studies), but their presence/absence is unlikely to affect any estimate. Note: the "center value" is just the arithmetic mean between the max-tolerance of the coolest/driest NLR and the min-tolerance of the warmest/wettest NLR; all other NLRs have no function in the framework of simple MCR approach such as CA.

Table R1. Redundant (Moraceae) and highly inconsistent (Selaginellaceae) tolerance data

NLR	MAT	CMT	WMT	MAP	HMP	LMP	WMP
Cudrania[a,b]	-5.3–27.7	-25.6–27	12.9–28.8	213–3151	65–389	2–165	16–264
Morus	3.1–21.9	-11.8–13.6	15.6–28.9(!)	305–1722	82–292	0(!)–83	82–264
Moraceae[b]	-5.3–27.7	-25.6–27	12.9–28.8	213–3151	65–389	2–165	16–264
Selaginella	3.9–27.7	-17.3–26.5!	10.5–23.3!	222–1377	55–256	0–53	2–252
Selaginellaceae	10–26.9!!	-7.3–20.3!!	25–31.7!!	396!–1682	108!–343	3–43!!	69!–304

[a] = *Maclura*

[b] Recorded tolerances of *Cudrania* and Moraceae are identical

Line 739 "Avoid family-level NLRs"

This would seem to be obvious - why would anyone use them unless they had to? But more to the point why-not? They are simply another taxon. If the statement means it is better to have a few good genus-level NLRs and that including family level ones would actually lower precision, then I think this should be stated (again).

Response R13: That's not the point here. We suggest to select likely NLRs (groups of genera, species groups) rather than bulk NLRs (the entire family, genus). We clarified: "Avoid family-level or unrepresentative genus-level nearest-living relatives. Correctly established family tolerances of non-relict families will usually be very large and hardly representative for the fossil; the same holds for widespread, diverse genera such as *Pinus* or *Quercus*."

Realistic Precision

Line 760 'Sensible' values - Start with +- 2 C. Yes, but one person's sensible is another's silliness. Where does this figure come from? Why not 3, or 4, or 5?

Response R14: Too true. The 2° figure was just a shot-in-the-blue. Anything below 2°, which translates roughly into 400 m altitude range, is hardly realistic given the lack of gridded distribution data for many taxa/density of climate records for critical areas. And people will not stop using 0.1 °C if one doesn't provide an alternative. Nevertheless, the critique is valid and we fused points 4 and 5 and just provide references how to get more sensible tolerance data than provided by Palaeoflora's 4-6 handpicked climate stations and get a feeling about the error range, but refrain from giving another (indeed silly) cut-off value.

Plea to make work reproducible.

Line 766. Document all steps, etc….

Surely this should be directed at editors? If a paper is based on undescribed, unillustrated material, - Don't publish it!

Response R15: Good point. We added: **"**Editors should ensure that critical data is properly documented as outlined by Grimm and Denk (2012) and Utescher et al. (2014; see also Table 1)." See also **Response E3**.

Table 5

Quan et al. Give the NLR of Cupanieidites as Sapindaceae/Myrtaceae and the discussion is about the fact that these are distantly related. True, but the error is in saying Sapindaceae/Myrtaceae in the first place. A far as I know, the distinction between Cupanieidites (Sapindaceae) and Myrtaceidites (Myrtaceae) is sculpture. The two should not be confused.

Response R16: In both families (Sapindaceae, Myrtaceae) syncolpate pollen can show psilate sculpturing under LM (e.g. Van der Ham, 1990; Thornhill et al., 2012a, b, c) or can be sculptured (in Sapindaceae reticulate, striate, to striatoreticulate, microrugulate; in Myrtaceae rugulate to verrucate under SEM). Potonié (1960) mentions in his original description of *Myrtaceidites* "..zart granulat oder mit feiner Infra-textur, niemals deutlich retikulat" [faintly granular with minute infrastructure, never distinctly reticulate], which would fit both families under LM. Krutzsch (1959) mentions in his original description of *Cupanieidites* not a word about sculpturing, only that this form genus does not show "Polfelder" [polar fields]. We hence did not change the affinities in the text. But this uncertainty could be easily be solved by combined LM and SEM investigation.

Figure Captions

In general I would ask for a little more explanation of the complex figures. For instance, Figure 4 has me confused - what is the 'x' axis component to the data points?

Response R17a: We added information to legends of Figs 1, 4 and 5. In one the explanation of the abbreviations was missing, in the other important information about what is shown in A–D.

Plate 1 needs its point stated so it is stand-alone. I.e.'They all look the same'.

Response R17b: We added "Lack of diagnostic morphological characteristics in extant members of subfamily Taxodioideae (Cupressaceae)" to the legend.

[This review was signed (non-anonymous); name removed]

Reviewer #2

Manuscript Number: PALBO3161

Title: Fables and foibles: a critical analysis of the Palaeoflora database and the Coexistence approach for palaeoclimate reconstruction

Dear editor,

Many thanks for the possibility to review this manuscript which is an extremely important contribution to the ongoing discussion if the coexistence approach in its current form is a valid method to reconstruct palaeoclimate.

Response R18: Thanks for accepting to review. We were however not aware that there is an ongoing discussion on the Coexistence Approach. Its supporters consider it robust and its results valid beyond any doubt (see Utescher et al., 2014), and this opinion is obviously shared by many editors and reviewers. We have shown that it is as a pseudo-science based on poor data (Grimm and Denk, 2012; this study), badly applied (Grimm and Denk, 2012; this study; Grimm, 2015), and generally fallacious (Grimm and Potts, 2016a). See also **Response E2**.

CA has never been really validated. The only published validation (see also Utescher et al., 2014) is the original study (Mosbrugger and Utescher, 1997), which reconstructs wrong precipitation values for one of three modern floras, and uses erroneous tolerance data (cf. Utescher et al., 2014, table 2) for their single fossil case flora. Another test can be found in the thesis of Klotz (1999; supervised by V. Mosbrugger), who, however, did not use the Palaeoflora data but other (better) climate tolerance data and found that a simple (unweighted) MCR method such as CA cannot even resolve the dramatic climate fluctuations during the Pleistocene-Holocene. This agrees with the findings of Thompson et al. (2012), who showed that unweighted MCR fails to reconstruct the c. 5 °C increase in CMT since the last glacial maximum in North America, in contrast to a weighted MCR method. Indeed, the MCR part of a MCR-NLR method does not need any validation: the way we define tolerances (if done properly) ensures that any modern flora (given that all elements are correctly identified) will have 100% coexistence and produce a MCR interval which naturally comprises the real value. That this is not the case for CA+PF, is just because their climate tolerance data is poorly established. The critical question is, however, how precise the MCR will be if we only have 10, 20, etc genera at hand. This has not been studied for any MCR method so far. The only taxon-based method studied in this respect has been the “taxonomic calibration” technique suggested by Boyle et al. (2008), which showed a similar precision using species- and genus-level data.

General remarks:

- *Despite minor problems, I encourage publication but would like to ask the authors to use a strictly rational language (no ironic comments).*

Response R19: We went through the text again to check for irony. Some phrases remain that may still sound ironic, but note that these merely state the obscure logic used by CA+PF.

- *Title: I would shorten the title, also to avoid provocations:*

"A critical analysis of the Palaeoflora database and the Coexistence approach

for palaeoclimate reconstruction"

Response R20: The title should provoke, such as that of the follow-up paper (Grimm and Potts, 2016a, Fallacies and Fantasies...), which has been published in January this year (see also Grimm and Potts, 2016b). CA+PF publications are to a substantial fraction fables, fairy tales revolving around minute climate shifts which, even according to Utescher et al. (2014), are beyond the resolution capacity of the CA. Furthermore, they are full of foibles from the very first publication onwards (see e.g. Grimm and Denk, 2012, ES2 and ES4; this study).

- *Very often palynological problems, e.g., uncertain botanical affinities, are a good argument. This should be emphasized more often, e.g., by giving more examples. A classical one is Arecipites which is not only produced by palms but also by many other monocotyledons. Also the genus Cyathidites might be not only produced by Cytheaceae but probably also a number of other families, e.g., Dicksoniaceae. During the Eocene the botanical affinities of countless pollen types are even completely unclear.*

Response R21: The two suggested form genera and the associated references have been added as examples.

-Very often the authors assume that a certain fact is general knowledge which is not always the case. More references and brief explanations are needed in respective areas (see annotated text)

Response R22: See **response R24.**

-Another aspect which should be emphasized is that many current taxa might occupy certain ecological niches at the margin of their actual optimum because they are weak competitors, e.g., many conifers.

Response R23: This is a principal problem of all NLR approaches, and in particular for MCR-NLR techniques, and has been noticed with-in the follow-up paper (Grimm and Potts, 2016a). We added the reference to that paper where appropriate.

For details please read the attached annotated pdf!

Response R23: We followed most suggestions, except for the following:

Re: *The authors should mention another problem: esp. the studies in the Lower Rhine Embayment (Utescher et al. 2000...) do not reveal the actual pollen data used. There is no way to evaluate the climate reconstructions. It is very likely that old/methodologically outdated data from the GD NRW in Krefeld were used.*—Most likely. Such problems linked to the poor documentation were already addressed in our 2012 paper, and we see no point in repeating it; it's a characteristic of pseudoscience to cloud the origin of data used for reconstructions. Apparently our first paper stirred things up a bit and Utescher et al. (2014)

now recommend that fossil lists, fossil-NLR associations, and NLR tolerance data must be documented, but don't adhere to their own guidelines (e.g. Tang et al., 2015; Utescher et al., 2015). This comes close to fraud, but we're not allowed to call it by the name as this would be too personal (see **Responses E1, E2**)

Re: *The authors should write a sentence or two about the fact that during the Eocene many current genera were not even present.*—Regarding northern hemispheric trees, an increasing number of genera are found to be actually present in the Eocene (e.g. most Fagales genera can be traced back to the Eocene), but since only LM-studied pollen genera are listed, we cannot tell if the pollen represent extinct lineages (which would need to be excluded in case of any NLR approach) or members of modern genera. But the more important issue is that even if the genera are present in the Eocene, we have no idea if their realised niche back then fits to their modern one. And this is a problem for all NLR methods. Theoretically one opens Pandora's Box, but technically, the problem could be handled (see Grimm and Potts, 2016a): we simply should not aim to reconstruct ancient palaeoclimates using a technique that already fails for the last 10,000 years, but go back in time step-by-step. This may allow identifying the genera, which underwent substantial niche shift (e.g. dragon trees; Denk et al., 2014), i.e. the taxa that increasingly violate the nearest-living-relative principle.

Re: *another problem: taphonomy. Many records might be biased due to differential pollen preservation*—Sure. But this is a principle problem of any pollen-based approach and not a practical foible of CA+PF. We also cannot test the amplitude of this effect, since only very few CA studies provide images of their pollen.

Re: "extinct climate" *Direct quote? What is an extinct climate anyway :)?*—Yes. Direct quote. Bare of any logic (see Grimm and Potts, 2016a, fig. 7 and according text passages), Utescher et al. (2014) consider it a "strength" of the CA to be able to reconstruct climate situations not found today, which they refer to as "extinct climates" in various CA+PF papers (the term has not been used anywhere else to our knowledge, another characteristic of pseudoscience). The notion of being able to reconstruct a climate situation not found today using an actuo-palaeontological method makes any outsider laugh out loud or shake the head in disbelief. An "extinct climate" such as a narrow coexistence interval directly evidences a violation of the basic assumption (illustrated in Grimm and Potts, 2016a). Apparently, editors and reviewers of CA+PF papers don't see this the same way.

Re l. 472: *yes, that is another problem with Utescher´s approach. They calculate very high precipitation for the Lower Rhine Embayment during the Neogene. In fact this was a huge swamp with an extremely high water table-the regional precipitation might have been significantly lower. This point is quite interesting and you should elaborate a bit.*—This would go beyond the scope of the paper. Effectively one needs to discuss if azonal elements would need to be excluded per se, but this reduces significantly the number of NLRs. As for many other assemblages, it's probably the unrepresentative NLR tolerance data of the taxodioids and engelhardioids that inform CA reconstructions for the Lower Rhine Embayment. Otherwise, if the modern plant still thrive in swamps and is independent of local climate, this should be captured by the modern distribution (subcosmopolitan) and reflected by extremely large tolerance ranges (as in the case of e.g. *Alnus*). The follow up paper (Grimm and Potts, 2016a) discusses in more depth the problem of the realised vs. actual niche, and the reference has been added.

Re l. 504 *Reference needed. Reliability of the study? How young? Salas-Leiva et al. 2013: No older than 12 Mio a. Compare Condamine et al. 2015.*—References (Nagalingum et al., 2011; Salas-Leiva et al., 2013) added. GWG must agree that both referenced studies have notable data issues, but principally it's safe to state that the modern species don't root deep (although

12 Ma is probably much too young; all dating studies on Cycadales use poor age constraints), but their genera do. In other words, there may have been *Cycas* (or other dragons) in the Eocene of China, but we have no idea if their niche is the cumulative one of the modern species of the same genus and how close they were related to their modern-day counterparts.

Re l. 533 *You might mention that U. is also known from the Cretaceous, so it is a rather ancient lineage with many "dead ends". Also Ulmaceae are notoriously difficult to differ from each other if you only have the pollen... Ulmus and Planera are genetically very close anyway and can even be hit by the same diseases :)*—We don't not want to discuss Cretaceous pollen determined via LM.

Re l. 555 *This point is very strong and you might want to make it No 1 in your line of arguments.*—The current order is simply alphabetical. And following 19th [century] style (see introductory comment of the editor), we like the increasing dramatic storyline, with the most obscure example (and most important for purported CA+PF estimates in general) at the end.

Re l. 671 Removal of "highly alarming"—We'd like to keep this. Anything weaker (e.g. "worrying") suggests we might be worried without cause. The foible level of CA+PF studies is alarming, most of these results should have never been published if the peer-review system would have worked properly.

Re l. 674 *Name taxa*—The pairs are named in Table 5.

I recommend a minor revision.

References

Anonymous Referee #2, 2016. Interactive comment on "Fallacies and fantasies: the theoretical underpinnings of the Coexistence Approach for palaeoclimate reconstruction" by G. W. Grimm and A. J. Potts. Clim. Past Discuss., 11: C2884–C2888.

Boyle, B., Meyer, H.W., Enquist, B. and Salas, S., 2008. Higher taxa as paleoecological and paleoclimatic indicators: A search for the modern analog of the Florissant fossil flora. Geol. Soc. Am. Sp. Paper, 435: 33–51.

Denk, T., Güner, H.T. and Grimm, G.W., 2014. From mesic to arid: Leaf epidermal features suggest preadaptation in Miocene dragon trees (*Dracaena*). Rev. Palaeobot. Palynol., 200: 211–228.

Grimm, G.W., 2015. Interactive comment on "Strong winter monsoon wind causes surface cooling over India and China in the Late Miocene" by H. Tang et al. Clim. Past Discuss., 11: C81–C86.

Grimm, G.W. and Denk, T., 2012. Reliability and resolution of the coexistence approach — A revalidation using modern-day data. Rev. Palaeobot. Palynol., 172: 33–47.

Grimm, G.W. and Potts, A., 2016a. Fallacies and fantasies: the theoretical underpinnings of the Coexistence Approach for palaeoclimate reconstruction. Clim. Past, 12: 611–622.

Grimm, G.W. and Potts, A., 2016b. Response to anonymous reviewer 2. Interactive comment on "Fallacies and fantasies: the theoretical underpinnings of the Coexistence Approach for palaeoclimate reconstruction" by G. W. Grimm and A. J. Potts. Clim. Past Discuss., 11: C2936–C2939.

Hoorn, C., Straathof, J., Abels, H.A., Xu, Y., Utescher, T. and Dupont-Nivet, G., 2012. A late Eocene palynological record of climate change and Tibetan Plateau uplift (Xining Basin, China). Palaeogeogr. Palaeoclimat. Palaeoecol., 344–345: 16–38.

Klotz, S., 1999. Neue Methoden der Klimarekonstruktion - angewendet auf quartäre Pollensequenzen der französischen Alpen. Tübinger Mikropaläontologische Mitteilungen. Institut & Museum für Geologie & Paläontologie [now: Institute for Geosciences], Eberhard Karls University, Tübingen.

Krutzsch, W., 1959. Mikropaläontologische (sporenpaläontologische) Untersuchungen in der Braunkohle des Geiseltales. Geologie, 8, Beiheft 21/22: 1–425.

Mosbrugger, V. and Utescher, T., 1997. The coexistence approach — a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. Palaeogeogr. Palaeoclimat. Palaeoecol., 134: 61-86.

Nagalingum, N.S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P. and Mathews, S., 2011. Recent synchronous radiation of a living fossil. Science, 334: 796–799.

Potonié, R., 1960. Synopsis der Gattungen der Spore disperse. III. Beihefte zum Geologischen Jahrbuch, 39: 1–189.

Quan, C., Liu, Y.-S.C. and Utescher, T., 2012. Eocene monsoon prevalence over China: A paleobotanical perspective. Palaeogeography, Palaeoclimatolgy, Palaeoecology, 365–366: 302–311.

Salas-Leiva, D.E., Meerow, A.W., Calonje, M., Griffith, M.P., Francisco-Ortega, J., Nakamura, K., Stevenson, D.W., Lewis, C.E. and Namoff, S., 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. Ann. Bot., 112: 1263–1278.

Tang, H., Eronen, J.T., Kaakinen, A., Utescher, T., Ahrens, B. and Fortelius, M., 2015. Strong winter monsoon wind causes surface cooling over India and China in the Late Miocene. Clim. Past Discuss., 11: 63–93.

Thompson, R.S., Anderson, K.H., Pelltier, R.T., Strickland, L.E., Bartlein, P.J. and Shafer, S.L., 2012. Quantitative estimation of climatic parameters from vegetation data in North America by the mutual climatic range technique. Quaternary Sci. Rev., 51: 18–39.

Thornhill, A.H., Hope, G.S., Craven, L.A. and Crisp, M.D., 2012a. Pollen morphology of the Myrtaceae Part 1: Tribes Eucalypteae, Lophostemoneae, Syncarpieae, Xanthostemoneae and subfamily Psiloxyloideae. Austr. J. Bot., 60: 165–199.

Thornhill, A.H., Hope, G.S., Craven, L.A. and Crisp, M.D., 2012b. Pollen morphology of the Myrtaceae. Part 2: tribes Backhousieae, Melaleuceae, Metrosidereae, Osbornieae and Syzygieae. Austr. J. Bot., 60: 200–224.

Thornhill, A.H., Hope, G.S., Craven, L.A. and Crisp, M.D., 2012c. Pollen morphology of the Myrtaceae. Part 3: tribes Chamelaucieae, Leptospermeae and Lindsayomyrteae. Austr. J. Bot., 60: 225–259.

Utescher, T., Bondarenko, O.V. and Mosbrugger, V., 2015. The Cenozoic Cooling – continental signals from the Atlantic and Pacific side of Eurasia. Earth Planet. Sci. Lett., 415: 121–133.

Utescher, T., Bruch, A.A., Erdei, B., François, I., Ivanov, D., Jacques, F.M.B., Kern, A.K., Liu, Y.-S.C., Mosbrugger, V. and Spicer, R.A., 2014. The Coexistence Approach—Theoretical background and practical considerations of using plant fossils for climate quantification. Palaeogeogr. Palaeoclimat. Palaeoecol., 410: 58–73.

Van der Ham, R.W.J.M., 1990. Nephelieae pollen (Sapindaceae): Form, function, and evolution. Rijksherbarium/Hortus Botanicus, Leiden.