

Interactive comment on “Compilation of pollen productivity estimates and a taxonomically harmonised PPE dataset from Northern Hemisphere extratropics” by Mareike Wieczorek and Ulrike Herzschuh

Marie-Jose Gaillard (Referee)

marie-jose.gaillard-lemdahl@lnu.se

Received and published: 31 March 2020

General Comment

This compilation is useful as a summary of the studies that have attempted to estimate pollen productivity of plant taxa with the aim to apply models of pollen-vegetation relationships such as REVEALS and LOVE (Sugita et al. 2007a and b, The Holocene). However, the way in which the values of relative pollen productivity (RPP, earlier abbreviated PPEs) found in these publications have been handled is neither adequate

C1

nor useful. My major concern with this synthesis is that it handle RPPs as they were measurements, but they are estimates calculated using a model (ERV model) that has assumptions. When these assumptions are badly met in the studies, the results do not make sense. Unfortunately several RPP studies were published although they are theoretically not sound. This is due to the fact that few palynologists understand the theory of the ERV model, its sub-models (1,2,3), and the likelihood method used to estimate RPPs. Therefore reviewers did not notice that these studies present values of RPP that are not correct.

The compilation of M. Wieczorek and U. Herzschuh does not take into account earlier expert evaluations of the RPP estimates (e.g. Mazier et al., 2014 for Europe; Li et al., 2018 for China). It therefore disregards careful evaluations that were meant to help palynologists in their choice of values to be applied. It implies that the database includes a mix of reliable and unreliable values of pollen productivity. The database also excludes reliable values for plant taxa that have been harmonized into higher taxonomic groups and, therefore, mixed with values representing a larger number of species or genus and often different ones. Moreover, the different ERV sub-models used are not explained, and the reason why the results from all ERV sub-models are not included in the database is not provided. In the excel file that can be uploaded in PANGAEA, the column “model” includes a mix of information on the ERV model submodels (1, 2, 3) and on models of dispersion and deposition (e.g. Langrarian model), although these are two different things. There should be one column for the ERV sub-model chosen by the authors of the original publication (1, 2 or 3), one column for the dispersion and deposition model (Gaussian Plume Model or Lagrangian model), and a third column for the vegetation distance weighting model (1/d or Prentice’s model (bog)).

In my opinion, a database should either include all RPP values published (i.e. values obtained with all ERV sub-models and all distance-weighting methods used by the authors) OR a selection of RPP values based on a proper expert evaluation. The database as it stands is neither nor. If it is a database including all values, the user

C2

should be referred to former evaluations and warned on not using the database as a “black box”, but as a source of data for further evaluation and testing of the values. If it is a database including only part of the published values, the reason for including and excluding values need to be argued for on scientific/theoretical grounds. I am of course fully open to new and different evaluations from those by Mazier et al. (2012) and Li et al. (2018). However, these evaluations should be based on expert knowledge, which is not the case in the current compilation submitted to ESSD. The database is now including values from studies that are theoretically not sound because most of the assumptions of the ERV model are not met. For instance, Li et al 2011 and Han et al. 2017 (included in the compilation) are theoretically not sound studies (see evaluation in Li et al., 2018). In contrast, Zhang et al. 2017 (not included in the compilation) is a study performed following the correct standards and that meets the most important assumptions of the ERV model (see evaluation in Li et al., 2018). The same can be said of several European studies that are either included or excluded without relevant reasons. It should be also noted that there is a new synthesis and evaluation of RPPs in Europe soon to be submitted (Gaillard et al., in progress), and a synthesis and evaluation of RPPs in N America-Canada-Alaska (Dawson et al. in progress). These studies will further help palynologists to choose RPP values for applications in those regions.

I do not want to reject this paper, because it would not be constructive. I require instead, for the sake of high quality science, that the revisions I am suggesting be considered with care. In order to be useful for the scientific community, the database should include all published RPP values. The authors then have to warn the users on the importance of choosing their RPP values after thorough evaluation of the RPP studies, considering the aim of their application (reconstruction of local, regional, or continental vegetation cover), and having strong arguments for their choice. One has to remember that most published RPP values have not been tested/validated. There is therefore an enormous need of test/validation studies using various alternatives of RPP values (see validations in Hellman et al., 2018 a and b; Cui et al., 2014; Mazier et al., 2015). For the sake of

C3

such studies, a complete database of RPP values would be most valuable.

There is, however, one part of the paper that cannot be accepted: the “harmonized RPP dataset” of means of RPPs throughout the northern hemisphere. There are some limits to mixing the RPP values over geographical space!. When there are strong reasons to think that differences between obtained RPPs are due to differences in species represented in different continents (such as for *Artemisia* and *Pinus* in China versus in Europe/N America), it does not make sense to use a mean of those values in all parts of the Northern hemisphere. If we have used a mean of values within NW Europe (Trondman et al., 2015) and within temperate China (Li et al., 2020), it was motivated by the fact that there were too little RPP values to demonstrate that the difference between RPPs within Europe (or within China) could be explained by climate and/or vegetation composition alone. However, it is clear that the RPP values of e.g. *Pinus* and *Artemisia* are generally higher in temperate China than in NW Europe (based on the theoretically soundest studies). Therefore, it is not appropriate to mix them.

In summary, I request the following major revisions:

1. Delete the “harmonized RPP dataset” with mean RPP values for the entire N Hemisphere, it does not make sense, as long as it is not tested/validated to produce realistic REVEALS reconstructions of land cover. The danger is to let people believe that this RPP dataset is the best possible for the N Hemisphere and can be used right away for the best possible results. This neither true nor tested.
2. Add to the database all RPP values available in all RPP studies that are not included so far.
3. Indicate in the database the studies and RPP values that were evaluated as not reliable due to theoretical problems (often assumptions of ERV model not met, and/or strange behaviour of log likelihood values while estimating RPP values) in earlier expert evaluations.

C4

4. Warn the user: the database should not be used as a “black box”. Values should be selected carefully based on sound arguments.

Detailed comments

Please read comments made directly in the pdf file and implement the required revisions.

I am open to discuss the issues I explain above directly with the authors in case something is unclear or seems not correct in the view of the authors.

Marie-José Gaillard, March 21st 2020

Please also note the supplement to this comment:

<https://www.earth-syst-sci-data-discuss.net/essd-2019-242/essd-2019-242-RC1-supplement.pdf>

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-242>, 2020.