We are pleased that both reviewers see the value of the database that we aim to make publically available with this publication. We also appreciate the comments on the manuscript that will certainly help to improve it. We like to take this opportunity to respond to some of the comments.

Anonymous Referee #1

The manuscript presents a new and valuable dataset that is made publicly available on a well-established database (NeotomaDB) that adheres to the World Data System and FAIR principles. The dataset consists of pollen counts from traps located in various European regions (and associated metadata). It is potentially extremely useful to explore pollen-vegetation relationships, as has already been shown in a number of previous, regionally-focussed, studies. In comparison to these prior studies, the present manuscript explores this modern-pollen dataset at a scale that spans across much wider geographical (latitudinal and altitudinal) gradients. The manuscript focuses on pollen abundances of 14 pollen taxa (12 trees and shrubs genera, 2 herb families) and explores the relationships between the sum of the pollen-accumulation rates (PAR) of the 14 taxa ("total PAR") and selected environmental parameters (mean annual temperature, annual precipitation, forest cover). It also explores the relationship between tree PAR and forest cover, and the relationship between pollen deposition as a function of increasing distance to the nearest range boundary of the parent plant species. A long-distance transport threshold (LDT) is obtained that may be used to infer range-size changes based on fossil PAR values. Further a comparison between modern and fossil PAR is presented that shows PAR-inferred population-size changes for selected taxa through time at different sites located across the latitudinal gradient.

The manuscript is at times very confusing. For instance, as far as I was able to under-stand, the results indicate that total PAR is strongly related with forest biomass within a 10km radius around the traps (Table S4). The text instead reports that forest cover explains 72% of the variance of total PAR.

We realize that due to the many analyses conducted and co-authors contributions some arguments were unresolved and miscommunications occurred. Unfortunately, we were not able to obtain biomass data for European vegetation cover and used "forest cover" instead. We will work to streamline the manuscript correctly indicating the usage of "forest cover" throughout. 72% of explained variance was error in the analysis, Table S4 for updated results.

Further, Figure 3 shows how total PAR is related to latitude (besides, this relationship seems to be strongly determined by one datapoint),

We tested the influence of the data point on Spitsbergen, which of cause has a strong pull, but removing it is not changing our observations and conclusions as the following table with that point removed will show:

| | PAR | adjusted_PAR |
|---|---|---|
| latitude | 0.09 | 0.22 |
| MAT | 0.19 | 0.26 |
| Forest cover 10 km | 0.18 | 0.16 |
| latitude+MAT+Forest cover 10 km | 0.35 | 0.42 |
| latitude+MAT+Forest cover 10 km+elevation | 0.49 | 0.54 |

We would add this table with and without the point on Spitsbergen to the supplementary information of the manuscript.

but does not show how total PAR are related to forest cover. Instead, the Figure shows how tree PAR is related to forest cover.
-> It is striking to see that you explored the relationship between the sum of the pollen-accumulation rates (PAR) of the 14 taxa ("total PAR") and forest cover (text on P8L24-30 and Table S4). On which grounds would one expect a relationship between the PAR of herbs (Poaceae and of Cyperaceae) and forest cover?

We cannot expect a strong relationship between total PAR and forest cover as there may be 100% cover in a forest with very low productivity in northern Finland or a high productivity in southern Europe. By broadly accounting for differences in pollen productivity between trees and also grasses we are able to investigate a relationship between adjusted PAR and latitude and mean annual temperature. In a multiple regression model forest cover is however contributing to explain the variance in PAR and adjusted PAR. We illustrate in Fig. 3B that there is however a relationship between the minimum PAR values and the amount of forest cover. This relationship is important for interpretations as we are often interested to provide conservative interpretations on this "minimum" side of the spectrum e.g. during the initial spread of forest after the ice age.

In my opinion, one of the main weaknesses of the manuscript is that the Introduction does not convey which knowledge gaps and hypotheses are being addressed. The Introduction is strongly disconnected from the Abstract, the Results, the Discussion, and the Conclusions. It dwells on how and why the Pollen Monitoring Programme was established and lists prior, regional, studies. Some concepts that are mentioned both in the Abstract and the Discussion (CO2 fertilization, land use) are not found in the Introduction. Another concept (the importance of comparisons between modern and fossil pollen data, and of past and modern plant distributions and abundances) is only marginally mentioned in the Introduction, although it is important in the Discussion.
Introduction was improved addressing open questions (CO2 fertilization, land use, MAT). However, this manuscript also aims to introduce the database itself, which is a result of decades of research with its own history and thus development PMP network has a special subsection in the Introduction.

Moreover, links to Figures seem to be wrong in some places of the text, methods could be described better,
tables and their captions sometimes are incomplete, and the description of the results is confusing and in some places contradictory. Further, the paragraph 3.4 "Taxa specific linkage..." seems wordy and confusing, and could be substantially shortened
We attempted to place the figure close to the first mention in the text, automatic knitting process of the MarkDown document might move them one Figure per page. In a later comment you state the opposite: "The comparison between modern and fossil PARs is interesting (paragraphs 3.3 and 3.4)". Assuming "Paragraph 3.4 is enlightening" we agree that sect. 3.3 can be shortened.

(besides, I was surprised when I noticed that the taxa specific linkage is placed as supplementary material for some of the taxa. Maybe it was mentioned earlier in the Mat & Methods section?).

We selected the 8 most important taxa to be included in the main text. The remaining taxa including Cyperaceae and Poaceae, for which we assume the trapping data to be potentially biased are placed in the supplementary. We inform the reader about this in the Methods and Results section.

Some of the conclusions do not seem to be supported by the data.

Conclusions were shortened and they are supported by the data.

……

General comments:

-> Abstract:

- the collection of [...] is important...

Changed

- statement “This dataset shows that climate parameters [...] determine pollen productivity” is not supported by the data, which shows that forest cover explains a much larger share of the variance in “total PAR”.

Rephrased in the sense that total PAR is determined by forest cover.

- the statement “A signal of regional forest cover can be detected [...], while local tree cover seems more important” suggests that forest cover is substantially different from tree cover. I might have missed this difference when reading the text and suggest to better point this difference out.

We use only data from 10 km radius. We use forest cover in whole manuscript. Sentence merged with previous sentence. Tree cover deleted.

- the statement “PAR values up to (smaller than?) 30 grains [...] in fossil records should therefore be interpreted as long distance transport” should be nuanced further. I suppose this refers to PAR values of the pollen taxa that were explored in this study. There are very likely some plant species whose pollen is less well dispersed or that simply produce less pollen (e.g. Larix, insect-pollinated plants). Moreover, it seems to me that the threshold value represents the PAR at 200 km from the distribution limit (as of Figure 4b). Does this mean that you (arbitrarily) consider any distance beyond 200 km as a “long distance”? If so, please state this in a clearer way in the text.

Statement was specified by taxa used in the paper and the distance 200 km.

- the statement “Comparisons to fossil data from the same areas show comparable values” is unclear. What is meant exactly with the term “the same area”? Figures 6-13 show that the geographical distance between similar modern and fossil PAR values is often quite high.

Yes, this is unclear, because in Figures 6-13 we link individual traps or trap areas. In this sentence, we mean the comparison at level of same trap region (Fig. 5).

- L11: “may be hard to find” seems colloquial. Replace with “do not occur, were not found in this dataset”?

Replaced

- the last sentence could be replaced with a DOI link (or a Neotoma Explorer link) to the dataset in Neotoma...the link could be added on L5 after “1981 to 2017”.

Unluckily Neotoma does not release DOI links. Import was stopped for few months due to migration to another database system. New version of Tilia was released today, we will import it as soon as possible.
https://www.neotomadb.org/news/157902
Link to Neotoma Explorer added.

-> A regression model for tree PAR vs forest cover is presented. The regression is based on selected tree PAR values for 3% wide forest-cover bins (Figure 3b). The regression model suggests that an 80% forest cover within 10 km radius results in tree PAR values > 3200 (Conclusions, P23 L 15). I might miss an important point, but it seems to me that the deduction is not supported by the data presented in Figure3b. The Figure shows that values greater than 3200 tree PAR can be found even for 0% forest cover. It seems to me, instead, that tree PAR are > 20,000 for forest cover >20% (though strikingly the two sites with highest forest cover show rather low tree PAR values).

We believe that there is value in looking at the minimum values as clarified above. So yes it is well possible to obtain much higher PAR values at 80% forest cover, but values lower than 3200 grains per cm2 are unlikely.

This relationship is important for interpretations as we are often interested to provide conservative interpretations on this “minimum” side of the spectrum e.g. during the initial spread of forest after the ice age.

-> the manuscript shows decreasing pollen deposition as a function of increasing distance to the nearest distribution limit of the parent plant species (Figure 4). Based on this evidence, a long-distance transport threshold (LDT) for a distance of 200 km beyond the distribution limit is calculated. The thresholds could be used to infer range-size changes based on fossil PAR values, the manuscript reports.
- While these are interesting results and a potentially useful approach, some critical discussion of this may be useful. The distribution limits were extracted from GIS shapefiles published by Caudullo et al. (2017), which are publicly available on the figshare website with associated DataCite link (Caudullo, Giovanni; Welk, Erik; San-Miguel-Ayanz, Jesús (2017): Chorological maps and data for the main European woody species. figshare. Collection. https://doi.org/10.6084/m9.figshare.c.2918528.v5).
We cite the version we used - v2.

In the original manuscript where Caudullo et al. present the maps, they specifically mention that “Since the maps aim at representing the species general chorology at continental scale, providing a synthetic overview of distribution range, the mapped boundaries should not be considered as precise and sharp limits where the species is definitely present or absent, particularly at local level. Indeed, the first version of this dataset was created for the European Atlas of Forest Tree Species [16] to concisely outline the distribution ranges of described species, complementing information on the species biology and ecology. Errors and imprecision are partly inevitable, due to various causes, such as the quality of the original source, the geo-referencing procedure, the interpretation and the comparison of the sources in the same area and finally due to the limited precision of the manual digitalization process of the range borders (Fig. 1).”
We are aware that these distribution maps have uncertainties. However, PARs are unlikely to be much effected by the occurrence of parent trees at very low abundance. Due to plantations we did not look at the western distribution limits, but only explore latitudinal limits of tree distributions. We warn about the limited precision of these maps on the end of the Discussion.

- It is therefore highly questionable as to whether the distances to the distribution limits measured by Abraham et al. truly represent the actual distances to the species distribution limits. Thus, the precision and accuracy of the LDT values may be strongly overestimated and misleading. Instead of using one single distribution limit, a range of distribution limits may better represent the uncertainty of the mapped limits. Question is therefore: what LDT values would be obtained if the distribution limits of Caudullo et al. (2017) had an uncertainty? Say ca. +100km or even +200km?

Since we focus in these comparisons on the northern distribution limits we believe that rare occurrences of trees will be less and less likely as further north we go. So a given uncertainty in the maps would not change the relative differences between taxa nor the order of magnitude of the absolute values, which in any case are gross estimates. Still these numbers are useful as modern analogues and represent guidelines rather than hard thresholds.

- Moreover, it would be useful to show the complete data in the plots of figure 4, including the PAR values within the distribution range (thus extend the x-axes of the plots to include negative x-axis values). In theory at least, these PAR values should be greater than PAR values around the distribution limit and beyond the limits.

We included the data for all traps in Fig. 4a). Within the area of distribution area the data is represented as boxplots outside as a dots.

-> The comparison between modern and fossil PARs is interesting (paragraphs 3.3 and3.4).
- Paragraph 3.3. should be deeply revised and could be shortened. It could focus more on PAR-inferred presence/absence based on LDT limits that were presented previously (Figure 5), and on the identification of the closest modern counterparts of populations sizes and forest cover. Currently, some statements are descriptive and their relevance could be made clearer (for instance, on P13 L31 “Modern and fossil values agree for the sites in central Sweden at PARs between 1900-5600 grains...”). Some phrases could be removed (e.g. P13 L27 “As discussed in the main manuscript, ”), other ones are unclear (e.g. P13 L25 “ignoring traps from the Caucasus and Turkey”), and several statements should be supported with references to the literature (e.g. P16 L12“Picea abies is planted in many European regions outside its natural distribution”, or“ Fagus pollen occurs at fossil sites that were assumed to have never been within the distribution of the tree”), to name few examples.

Paragraph 3.3 was introduction for 3:4. It was shortened and included in the corresponding section. The summarizing sentences we removed.

- Paragraph 3.4 is enlightening. However, the term “analogue” (and “modern analogue”, which is used later in the Discussion) is not appropriate. With pollen records, modern analogues are generally referred to pollen assemblages (thus to vegetation composition). Here instead you refer to “comparable, similar, population size of one taxon”. Using the term “modern analogue” without clarifying that you are using it with a different meaning creates confusion, particularly in the Discussion where reference is made to the early-Holocene hazel maximum.

We agree that the term “modern analogue” evokes comparison of assemblages of pollen percentages (sensu Overpeck et al., 1985), however we, similarly to Overpeck et al. (1985), think that “modern analogue” should not be reduced to comparisons of modern pollen assemblages, but describes any link between a modern situation and its resemblance of a fossil find, including PAR. Also in geology and macro-ecology the term “modern analogue” is used in a broader context as exemplified by these publications: Sidder, A. (2020), Ancient sea levels in South Africa may offer modern analogues, *Eos, 101,* https://doi.org/10.1029/2020EO147001. ; Horsák, M., Chytrý, M., Hájková, P., Hájek, M., Danihelka, J., Horsáková, V., Ermakov, N., German, D. A., Kočí, M., Lustyk, P., Nekola, J. C., Preislerová, Z. and Valachovič, M.: European glacial relict snails and plants: environmental context of their modern refugial occurrence in southern Siberia, Boreas, 44(4), 638–657, doi:10.1111/bor.12133, 2015.

We did look for a different terminology but could not find a better term for what we are comparing and clarified this in the introduction.

-> There are other proxies that may be useful to determine presence of trees in fossil records (plant macrofossils and stomata). This could be mentioned in the text. Moreover, using fossil sites where such data is available could be useful to actually test the LDT limits, at least for some of the taxa. Further, how do the inferences based on the LDT limits compare with inferences made previously based on pollen percentages (or on plant macrofossils and stomata)? For instance, in a prior study (Giesecke et al., 2017 in JBiogeogr) a good agreement between the estimates of overall spread (Fig. 5a) based on different pollen percentage abundance classes was found. Some of the fossil sites were actually analysed for pollen, stomata, and plant macrofossils (e.g. Sägistalsee, Bachalpsee), but these results are not mentioned in this manuscript.

We compare LDT limits with dispersal model as required be second reviewer. Other proxies is an interesting aspect, which indeed we did not address in this manuscript. In making the database accessible we hope that questions like these will lead to additional usages of the database. We would be interested to explore it but feel it would add too much extra analysis and text and to the current manuscript.

Detailed comments:

- title: influx, or PAR? In the text you use PAR, not influx. Please homogenise terminology.

We use PAR in all text.

- Please consistently use italics for latin names.

I hope that all latin names are ok now, however translation via MarkDown is little bit tricky.

- P6 L21-22: you extracted forest cover data from all grid cells within 10km radius of the traps. Did you then calculate a mean forest cover value? Please clarify.

Yes, clarified.

- P7 L8: why 271 modern samples? The abstract mentions 2742 annual samples.

While the database contains all data that was submitted, we only considered average trap record with at least 3 years and thus obtained 271 traps that we base the analysis on. Averaging of the annual samples mentioned in the Methods. Number of traps and trap values is mentioned in the Results.

- P7 L10-ff.: am having trouble understanding what has been done, and why. Did you match fossil and modern PAR based on their ranking in the classes? What has been averaged? why should 500-year bins represent periods of long-term vegetation stability (fossil pollen records show fast population doubling times for some taxa)? What distance measure did you use for the cluster analysis (Euclidean distance)?

Rewritten. “periods of long-term vegetation stability” deleted, “distance matrix” changed to “matrix of distances”.

- P10 L111-12: total PAR is not shown in Figure 5.

Changed to “tree PAR”

- P11 L12: here the conclusion is presented before the data and then the data are presented to support the initially stated conclusion. Please reverse the line of arguments.

Removed

- P11 L13: am having trouble to understand why 92 pairs were obtained. please clarify.

We have 15 taxa in 7 regions, which make 105 cases that we investigate, however some regions lack species in the fossil and/or trap record and thus it is only possible to perform the t-test for 92 pairs of trap and fossil sites. We hope with Table S5 it is clearer.

- P11 L14: cannot find t-test and p-values on figure 6. Neither were “t-test” and “p-values” mentioned in the Material and Methods section. Please clarify why and what has been done.

We are sorry for this omission and will add it to the Methods section and add a table with p-values to the Supplementary Table S5.

- P20 L12: MAT has a very low influence on pollen deposition (see Table S4).

“strong influence” changed to “some influence”.

- P20 L17-18: Which data/analysis/result shows that biomass cannot explain the latitudinal PAR gradient?

changed to “forest cover”

- P 20 L31: why cannot the details be discussed here?

We meant that we can not discuss details of all species in all regions, our aims are general patterns. Sentence removed

- P21 L14: statement “modern PAR for Betula and Pinus are not found in fossil samples” contradicts a prior statement (P10 L12 “Maximum PAR in traps are always higher compared to fossil situations, with the exception of Corylus”).

P10 L12 refers to the whole dataset, i.e. maximum at continental scale, P21 L14 refers to the situation of individual traps. In both cases comparison results in higher trap value than fossil one. We added exact values to make it clear.

- P22 L16: the effects are possibly small in relation to the wide environmental gradient.

We added: Also the comparison of values over this large environmental gradient results in the signal being stringer than the noise.

- P22 L34-35: these are interesting notions. Please mention the usefulness of modern PARs (as listed here), and their importance for ecological and biogeographical studies, in the Introduction.

We added in Introduction: More recent investigations demonstrate the linear response of the absolute pollen deposition to absolute tree abundance (Seppä et al. 2009, Sugita et al. 2010, Matthias and Giesecke 2014), which may be used to reconstruct past standing tree biomass of different trees.

- P22 last line: remove “and it is almost surprising that...explanatory power”. It might be surprising, but is a fact that seems to contradict a prior study (Matthias et al.). Could the gradient length be an important factor here?

Sentence removed. Yes, present study uses large gradient and all trees together, whereas Matthias et al analyzed pollen-vegetation dataset by individual species.

-> Figures 6-13: - the a) and b) frames could be merged by using horizontal boxplots (instead of barplots) in a), and adding b) as an overlay;

I have tried to use horizontal boxplots (Fig. R1), but the highest values enlarge the x-axis while the lowest classes are not visible. The way of our presentation allows to appreciate the general pattern of mean values for trap areas (a) and we make the full variability of individual traps visible within the trap areas by coloured squares (b).

- font sizes are too small;

Font was enlarged to 2.5 for names of the fossil sites/trap areas and 2 for rest of the text.

-> References:

- add: Caudullo, G., Welk, E., San-Miguel-Ayanz, J., 2017. Choro-logical maps for the main European woody species. Data in Brief 12, 662–666.https://doi.org/10.1016/j.dib.2017.05.007

added

-> In Table S4:

- caption: what does “alternatively” mean here? Please clarify.

“alternatively” removed

- Forest biomass, or forest cover (as of P6 L20-21)?

Changed to forest cover.

- are the “Adjusted R2” values of the “adjusted PAR” values the ones obtained with the Andersen correction factors (P6 L23-24)? Please clarify.

Changed to: “(or logarithm of total PAR adjusted by Andersen values, Table S2)”

- please add r2 values for “tree PAR” (not total PAR) vs forest cover

Added

Anonymous Referee #2

Note to editor: The line numbers in the manuscript appear to be sometimes inconsistent. I assume this is a fault with the submission process not the authors’ fault. Apologies if any confusion arises from this.

General comments

An interesting manuscript with great potential to improve interpretations of fossil pollen records. It is clear that a lot of careful thought and work has gone into this study. While it has potential to be very useful, the structure and clarity of the work could be improved. I think in particular the structure could be refined to be more consistent throughout, and in the Introduction and Discussion sections, sub-headings introduced to clarify the development of the argument. There seem to be various strands to this paper:

1. The correlations of PAR with parameters such as forest cover and temperature as calculated from the modern samples
2. The question of long distance transportation beyond the current extent of the parent plant taxa

3. The relationship between modern PAR and fossil PAR for selected taxa.

I feel that if the first two are clearly and separately addressed, it would be easier for the authors to address the third point coherently. At the moment, the results broadly follow this structure, but the Introduction and Discussion do not, I suspect if they did, the paper would flow better.

We are grateful for this suggestion and we re-structured the Introduction and Discussion.

Specific points by section

Introduction

I was a little surprised that so much of the introduction dealt with history. Although I think this would merit its own sub-section, I think it would be better to focus on scientific questions addressed in the manuscript. Why pollen traps are an appropriate analogue for fossil records should be introduced. It would be nice to see some mention of species ranges and their possible fluctuation over time, and whether present day species distributions can be considered to be in equilibrium with climate. Another factor that could be addressed in the introduction is pollen dispersal and deposition; how far does pollen usually travel? This would set up the argument for your chosen LDT threshold.

We appreciate this suggestion to improve the organisation of the manuscript and we moved the historical aspects of the development of the network (see earlier response) to a subsection.

2 Methods

General comment: all botanical names including species epithets need to be written with their authorities the first time they are mentioned in the manuscript. Up to date authorities can be found here http://www.worldfloraonline.org/. Upon first use, a species must be written out in full even if its genus has been mentioned by name previously. This is to avoid confusion between genera that start with the same letter. So for in-stance, Pinus sylvestris L. could be shortened to P. sylvestris, but then Pinus mugo Turra needs to be written as such before it can be abbreviated to P. mugo

We do not deal any extinct organisms in the past, but palaeoecology of living, abundant and well-known plant taxa. Biogeosciences is not taxonomical journal and many papers in Biogeosciences are published without authorities. Journal´s guidelines do not mention the need for taxonomic authorities, but of course we can add them if our assoc. editor requires them.

Figure 1: As there is so much overlap on the map between modern and fossil sites, I think separating this out into two side by side maps with one showing fossil and one modern samples would make it clearer, and would also make it  easier to go back and check locations of fossil sites as I was trying to do so later in the manuscript.

Thank you for this comment. We already separated a map of only fossil sites and it is better readable.

On page 7, 2742 samples were mentioned as being in the database (Section 3.1) button line 8 only 271 are mentioned- which number was in the analysis?

We added to Section 3.1: “Considering the trap record with 3 years and more we obtained 271 mean trap assemblages.”

2.2 Data Collection

It would be interesting to see a plot of surface area of trap against PAR to test for a relationship there and potentially be able to correct if one exists.

This is an interesting point and the reviewer is invited to try this out as all this information will be available. Do you have any idea which mechanism can stay beyond?

2.3 Investigated taxa and Environmental parameter

Why was 200km chosen as the threshold for LDT?

Regional pollen is assumed to correspond to the vegetation cover in 100 by 100 km (Hellmann et al 2008) so doubling that distance represents a good rational. Additionally we considered the uncertainty of the maps.

2.4 Comparison

Page 12: Figure 5 is a bit tricky to interpret, however, once I had realised what it was supposed to be showing I saw its value. I particularly like the LDT cut-off, which will be potentially very useful in interpreting fossil records. I was surprised, however, that LDTs did not receive further attention in the discussion section as it seems that they are a tangible, useful output from the work.

We added example of LDT cut-off on the end of the section Analogues for vegetation reconstruction.

3.2 Recent and fossil PAR

Figures 6 onwards: In caption, specify that distribution of taxon is in grey.

Added: “... distribution of the species (gray, Caudullo et al., 2018 ...”

These taxon-specific figures are in general, I think, quite useful. I hope they are reproduced a little bigger (at least a page each)

Unfortunately one page per figure would surpass our APC budget.

so that the details on the maps are easily readable. If this is not possible, maps should perhaps be split into separate figures to improve legibility.

We will improve legibility by abbreviating names of fossil sites, so we make use of blank space and the map is larger now. We will enlarge all text in the figure to 2 and 2.5.

I am not sure the multi-coloured coding for the PAR values adds much to the figure-

Multi-coloured coding helps to link the fossil and trap PAR values of the same height/class. In the fossil stratigraphic plot we point out the class of interest, coloured squares in 6b) illustrate the variability of traps within one trap area.

you could probably do away with (b) and still retain the meaning of the figures.

Boxplot would be one option as proposed by reviewer 1 (see response there), black and white bar-plots for the fossil and trap record without any colours would not visualize the analysis we did.

I am also not clear on why, in the fossil graphs, only certain colours are included

- why is only the highest PAR of interest?

Only certain classes appeared in the fossil record. We picked the highest PAR class from the fossil record, because it represents also the densest population of the source plants.

3.4 Taxa specific results How were the 'main taxa' to be presented chosen for this section? It seems odd that some are arbitrarily in the SM, particularly arboreal pollen which was presented in Figure 2. Figures need to be referred to consistently in this section.

We state in Methods that they are/were dominant. In 3.3 Section we list them. We do not want to flood reader by all species. Fossil-trap links of species selected for the main text show nicely changes of distribution patterns, whereas species in the SM can suffer from certain biases and limitations (See Cyperaceae and Poaceae in section 4.2)

3.4.2 Betula

Why are traps from the Caucasus and Turkey ignored?

Reworded to "Letting aside traps from the Caucasus and Turkey...(, because those two areas have different species Betula than rest of the trap areas, but the most we deal with belong to B. pendula and B. pubescens)"

4.1 Discussion

The first sentence of the discussion doesn't seem to tally with the results- it looks, from your data, like the relationship between modern and fossil PAR is actually quite complex and variable. I don't necessarily think this is a bad thing, however; the paper presents a quantitative dataset that could potentially be used to help researchers quantify what their PARs from fossil data actually mean.

We included mention about complexity of the relationship to the first sentence: "in spite of different taphonomic processes that influence PAR values in pollen traps versus lake sediments.".

Line 22 onward: This paragraph seems to be about LDT, but that isn't explicitly stated.

Last paragraph of page 20 onwards appears to be, broadly, taxon-specific discussion of modern and fossil PARs; it would be better if this were clearly signposted and possibly split up with sub-headings.

Discussion divided into a more subsections. LDT subsection added

I would be interested to see some consideration of how these results might be useful in feeding into quantitative reconstructions of vegetation. Although PDD models tend to deal in percentages, surely this approach (on cores with appropriate chronologies) could open the door to future models being calibrated using PARs, an interesting prospect for vegetation reconstructions, particularly given your LDT estimates.

We included in the discussion possibilities of PAR vs. PDD and wider use of LTD estimates in quantitative reconstruction. See also Table S6.

4.2 Limitations

Line 21: Why are only 3 fossil sites listed here? Are the others not likely to show any bias?

The fossil PAR values at the 3 sites are rather high so we suspect that lake internal procceses may explain these values. In generall the reviwer is right these biases may also occure at other sites and we added „especially“.

Line 28: Unclear which analysis only included Poaceae and Cyperaceae

Changed to "...Poaceae and Cyperaceae are only herbs selected for our analyses."

5 Conclusions

This is succinct and mostly well-structured but would benefit from a closing statement outlining the applications and take home message of the paper.

We removed little bit vague last paragraph Conclusions and thus we make both ouputs, which are useful for palaeoecological applications – PMP database and LDT trhesholds – more visible.

Technical comments

Page 2

Line 12: replace 'case' with 'cause'

Replaced

Line 13: Could be rephrased as 'hopefully serves to improve interpretations' (or remove hopefully- I think it definitely will).

"hopefully" removed.

Page 3

Line 1: A good recent reference here would be Haselhorst 2020 (DOI:10.1111/jvs.12897) showing high interannual variability in the tropics too- strengthens general argument.

Citation added.

Line 7: remove comma after 'Although'.

Comma removed.

Page 7

Line 13: something strange happening here after 'values variation'- typo?

"variation.ignoring" replaced by "and we need to ignore"

Line 19: The sentences about fossil pollen seem out of place here as this section of the results is regarding modern pollen.

The sentences about fossil pollen placed to section of the fossil pollen.

Page 14

Line 5: ‘main text’ – where in the manuscript is being referred to?

“As discussed in the main text,” removed

Page 20

Line 31 paragraph: I think this paragraph might be better placed at the end of this section.

Paragraph placed to the end of the section.

The phrasing in Line 32 seems a little odd as you go on to give an example of linking fossil to modern PARs- perhaps delete the ‘details cannot be discussed’sentence.

Rephased and “Unfortunately, the details cannot be discussed here.” removed.

3.4.8 Line ? (Line numbers unclear) correct LTD to LDT and remove ‘threshold’

Corrected, removed.

# Patterns in recent and Holocene ~~pollen influxes~~ Pollen Accumulation Rates across Europe; the Pollen Monitoring Programme Database as a tool for vegetation reconstruction

Vojtěch Abraham[1], Sheila Hicks[2], Helena Svobodová-Svitavská[1, 3], Elissaveta Bozilova[4], Sampson Panajiotidis[5], Mariana Filipova-Marinova[6], Christin Eldegard Jensen[7], Spassimir Tonkov[4], Irena Agnieszka Pidek[8], Joanna Święta-Musznicka[9], Marcelina Zimny[9], Eliso Kvavadze[10], Anna Filbrandt-Czaja[11], Martina Hättestrand[12], Nurgül Karlıoğlu Kılıç[13], Jana Kosenko[14], Maria Nosova[15], Elena Severova[14], Olga Volkova[14], Margrét Hallsdóttir[16], Laimdota Kalniņa[17], Agnieszka M. Noryśkiewicz[18], Bożena Noryśkiewicz[19], Heather Pardoe[20], Areti Christodoulou[21], Tiiu Koff[22], Sonia L. Fontana[23, 24], Teija Alenius[25], Elisabeth Isaksson[26], Heikki Seppä[27], Siim Veski[28], Anna Pędziszewska[9], Martin Weiser[1], and Thomas Giesecke[29]

[1]Department of Botany, Faculty of Science, Charles University; Benátská 2; CZ-128 01; Prague; Czech Republic
[2]P.O. Box 8000, FI-90014 University of Oulu, Finland
[3]Institute of Botany v.v.i.; Czech Academy of Sciences; Zámek 1; CZ-252 43 Průhonice; Czech Republic
[4]Laboratory of Palynology, Department of Botany, Faculty of Biology, Sofia University, 8 Dragan Tsankov blvd., Sofia 1164, Bulgaria
[5]Lab. of Forest Botany, Faculty of Forestry and Natural Environment, Aristotle University of Thessaloniki, P.O. Box 270, 54124 Thessaloniki, Greece
[6]Museum of Natural History Varna, 41 Maria Louisa Blvd. 9000 Varna; Bulgaria
[7]University of Stavanger, Museum of Archaeology, Peder Klows gate 31A, PB 8600 Forus, NO-4036 Stavanger, Norway
[8]Institute of Earth and Environmental Sciences, Maria Curie-Sklodowska University; al. Krasnicka 2d; 20-718 Lublin; Poland
[9]University of Gdańsk, Faculty of Biology, Department of Plant Ecology, Laboratory of Palaeoecology and Archaeobotany, ul. Wita Stwosza 59, 80-308 Gdańsk, Poland
[10]Georgian National Museum, Purtseladze Str.3, Tbilisi 5, Georgia 0105.
[11]Faculty of Biological and Veterinary Sciences, Geobotany and Landscape Planning, Nicolaus Copernicus University in Toruń, 87-100 Toruń, ul. Lwowska 1; Poland
[12]Department of Physical Geography, Stockholm University, SE-106 91 Stockholm, Sweden
[13]Department of Forest Botany, Faculty of Forestry, Istanbul University-Cerrahpaşa; Bahçeköy; TR-34473; Istanbul; Turkey
[14]Depertament od Higher Plants, Moscow State University; Leninskie Gory, 1, 12, Moscow, 119234, Russia
[15]Main Botanical Garden RAS; Botanicheskaya, 4, Moscow, 127276, Russia
[16]Laugarnesvegi 87 íbúð 105, 105 Reykjavík, Iceland
[17]Department of Geography and Earth Sciences, University of Latvia; Jelgavas Street 1, LV-1004; Riga, Latvia
[18]Institute of Archeology, Faculty of History, Nicolaus Copernicus University in Toruń; Szosa Bydgoska 44/48; 87-100 Toruń; Poland
[19]Faculty of Earth Sciences and Spacial Managment, Nicolaus Copernicus University in Toruń; Lwowska 1, 87-100 Toruń; Poland
[20]Department of Natural Sciences, National Museum Wales, Cathays Park, Cardiff, CF10 3NP, U.K.
[21]Department of Forests, Ministry of Agriculture, Rural Development and Environment, P. Box 24136, 1701 Nicosia, Cyprus
[22]Tallinn University, School of Natural Sciences and Health, Institute of Ecology, senior researcher. Uus Sadama 5, 10120 Tallinn, Estonia
[23]Cátedra de Palinilogía, Facultad de Ciencias Naturales y Museo, UNLP, Calle 64 n°3, 1900 La Plata, Argentina

[24]Faculty of Resource Management, HAWK University of Applied Sciences and Arts, Büsgenweg 1a, 37077 Göttingen, Germany
[25]Turku Institute for Advanced Studies (Department of Archaeology), FI-20014 University of Turku
[26]Norwegian Polar Institute, Fram Centre, N-9296 Tromsø, Norway
[27]Department of Geosciences and Geography, University of Helsinki, Gustav Hällströmin katu 2, 00014, Helsinki, Finland
[28]Department of Geology, Tallinn University of Technology, TalTech, Ehitajate tee 5, 19086 Tallinn, Estonia
[29]Palaeoecology, Department of Physical Geography, Faculty of Geosciences, Utrecht University, P.O. Box 80115, 3508 TC Utrecht, The Netherlands.

**Correspondence:** Vojtěch Abraham (vojtech.abraham@gmail.com)

**Abstract.** The collection of modern spatially extensive pollen data ~~are~~ is important for the interpretation of fossil pollen diagrams. Such datasets are readily available for percentage data but lacking for pollen accumulation rates (PAR). Filling this gap has been the motivation of the pollen monitoring network, whose contributors monitored pollen deposition in modified Tauber-traps for several years or decades across European latitudes. Here we present this monitoring dataset consisting of 351 trap locations with a total of 2742 annual samples covering the period from 1981 to 2017. This dataset shows that ~~climate parameters correlating with latitude~~ total PARs are influenced by forest cover and climate parameters, which determine pollen productivity ~~. A signal of regional forest cover can be detected in the data, while local tree cover seems more important.~~ and correlate with latitude. Pollen traps situated beyond 200 km of the distribution of ~~the parent tree~~ a given tree species are still collecting occasional pollen grains of ~~the tree in question~~that species. PAR's of up to 30 grains $cm^{-2}$ $y^{-1}$ in fossil diagram should therefore be interpreted as long distance transport ~~. Comparisons to fossil data~~ from beyond 200 km from the area of distribution. Comparisons between modern and fossil PAR from the same ~~areas~~ regions show comparable values. Comparisons often demonstrate that similar high values for temperate taxa in fossils sites are found further south or downhill. While modern situations comparable to high PAR values of some taxa (e.g. *Corylus*) ~~may be hard to find~~do not occur, $CO_2$ fertilization and land use may ~~case~~ cause high modern PAR's that are not documented in the fossil record. The modern data is now publically available in the Neotoma Paleoecology Database ~~and hopefully~~ (https://apps.neotomadb.org/explorer/) and serves improving interpretations of fossil PAR data.

*Copyright statement.* The article is distributed under the Creative Commons Attribution 4.0 License. Unless otherwise stated, associated published material is distributed under the same licence.

## 1 Introduction

~~The interpretation of past vegetation composition from pollen analytical results hinges on the concept of uniformitarianism. Pollen percentagesfrom modern samples of mosses, soil litter and the top sediment of lakes are essential for understanding how the vegetation and other environmental conditions are represented in fossil pollen assemblages (**???**). The same understanding is required in the interpretation of pollen~~

## 1.1 The need for a dataset of modern absolute pollen deposition

Pollen analytical results are usually reported as percentages, which are a simple and robust representation of the data, while they have a number of well-known biases. The problem of separating locally produced from long distance transported pollen has received attention from the very beginning of modern pollen analysis (?). Particularly at the tree-line the interpretation from pollen percentage data may be misleading as local treeless vegetation (e.g. tundra) is often low pollen procuring, while distant woodlands (e.g. consisting of boreal trees) may produce much pollen. In such situations absolute pollen data are very informative. Pollen accumulation rates (PAR) or the ~~absolute numbers~~ number of pollen grains deposited ~~per unit surface area per unit time , only here collecting modern reference values is more difficult~~on the sediment surface over a set period of time is better suited to differentiate between long distance transported and locally produced pollen as already realized by (?) and (?). In a seminal publication (?) document the power of using absolute pollen deposition for the interpretation of vegetation change during the postglacial afforestation around Rogers Lake in Connecticut, USA. Another bias in percentage data is the interdependence of values obscuring the quantification in the amount of change of a single taxon. PARs are therefore required when studying the population dynamics of individual trees (?).

However, obtaining PAR from sediments cores requires accurate chronologies. While sedimentation in lakes occurs generally without abrupt changes, lake internal processes such as re-deposition and sediment focusing and also catchment erosion may bias the resulting values [?; (?); ?; (?). These are some reasons why advances in interpreting PAR have been slow. The other reason is that collecting modern PAR values is not as simple as collecting mosses, soil litter or the top sediment of lakes for obtaining modern pollen percentages for a particular vegetation type. Modern rates of pollen accumulation can be obtained from monitoring pollen deposition using pollen traps (?) as well as from carefully sampling the top sediments of lakes that are either annually laminated or precisely dated (?). ~~Monitoring modern absolute deposition of pollen has been conducted since the development of modern pollen analysis (?) already allowing Welten (?) to draw on this information for the interpretation of the first pollen accumulation rate reconstruction from the laminated sediments of Faulenseemoos.~~ Due to the high ~~interannual~~ inter-annual variability in pollen production ~~(?)~~ (??) it is necessary to conduct pollen monitoring over several years to enable comparisons ~~to be made~~ with estimates from sediment cores (??). ~~The~~

For these reasons there are numerable studies investigating the pollen vegetation relationship using percentage data and comparably few using absolute pollen deposition. Nevertheless, investigations using pollen traps yielded invaluable insights in the mode of pollen transport (?). Also the construction of representation factors for common Europe trees by Andersen (?) , which are still used, were based on pollen monitoring data from pollen traps. In this way pollen monitoring studies have contributed to the development of models of pollen dispersal and deposition (?).

More recent investigations demonstrate the linear response of the absolute pollen deposition to absolute tree abundance (?, ?, ?), which may be used to reconstruct past standing tree biomass of different trees. However, rather than changes in tree biomass, changes in PAR at an annual time resolution correspond to weather conditions of the previous, as well as the year of flowering (?, ?, ?). Thus the question arises: If weather is determining annual pollen production, could climate or site conditions not determine average rates? Comparing *Pinus* PAR between two pine dominated forest regions in central Sweden and north east Germany shows much higher values in the south, suggest that PAR may not correspond to tree biomass alone (?) . The relationship between pollen production and weather suggests that more pollen is produced when the primary productivity of the tree is higher. Therefore, climate, soil fertility and even the amount of $CO_2$ in the atmosphere may determine the pollen productivity of a tree at a given site. Already Welten (?) interpreted the first fossil PAR in this way, suggesting that climate deteriorations may not immediately lead to a decline in forest cover but to the amount of pollen produced. This interpretation of changing PAR was however forgotten. If climate and $CO_2$ determine pollen productivity than the postglacial increase in PAR at Rodgers Lake could also be due to a change in these parameters. It is therefore impotent to investigate the possible relationship between climate and average PAR in more detail. Also the initial question on the amount of pollen that may arrive at a site from long distance sources has not been addressed in a systematic way using modern absolute pollen deposition data.

Spatially extensive modern pollen percentage datasets provide the possibility of searching for modern analogues for fossil pollen proportions and in this way reconstructing past vegetation and environmental conditions assemblages (???). Modern datasets of absolute pollen deposition are hitherto rarely used as references helping to reconstruct past tree abundances or environmental conditions. By using a network of pollen traps across the latitudinal ~~treeline in Finland (?) demonstrated that these modern analogues can be applied directly to interpret past vegetation changes (?), and this~~ tree-line in Finland Hicks et al. (?) show that average modern PAR values can be obtained representing the gradual transition from the boreal forest to tundra. These "modern analogues" were successfully applied reconstructing Holocene shifts of the latitudinal tree line (?). This idea of building a modern dataset of absolute pollen deposition that can be used as a reference to interpret fossil PAR was the motivation for the establishment of the Pollen Monitoring Programme (PMP, ?; ?; ?)~~.~~).

### 1.2 The Pollen Monitoring Program (PMP)

The program was launched in August 1996 at a meeting in Finland, bringing mainly European researchers together. Members of the network changed over the years and monitoring experiments were discontinued or newly started. Although ~~,~~ pollen monitoring studies were and are carried out on other continents (e.g. ?) the PMP had little success in attracting researchers working outside Europe. The standardisation of the monitoring protocol allowed for easy comparisons between the results in different regions which were discussed at INQUA in 1999 and led to a special volume published in 2001 (?) collecting results based on several initial time series (????) as well as a first comparative study (?). More individual results were published in the following years (e.g. ??????) and comparative studies followed in a second special volume published in 2010 (?). The data produced by contributors to the PMP were analysed for different questions, including weather parameters determining the amount of pollen production (?) and its correlation to masting years in *Fagus* (?). ~~However, no study has hitherto explored to what extent the PMP data collected provides modern comparisons to fossil situations as was originally intended. Since this dataset extends across the distribution limits of some major European trees, it can also provide information on the quantity of pollen dispersed over long distances, as an analogue for the interpretation of regional or local presence, which is difficult to achieve based on pollen percentages (?).~~

The programme established a database collecting the original data for individual years, as well as general information on the pollen traps installed in the different regions (Fig. 1). The database was developed offline and was thus difficult to access by

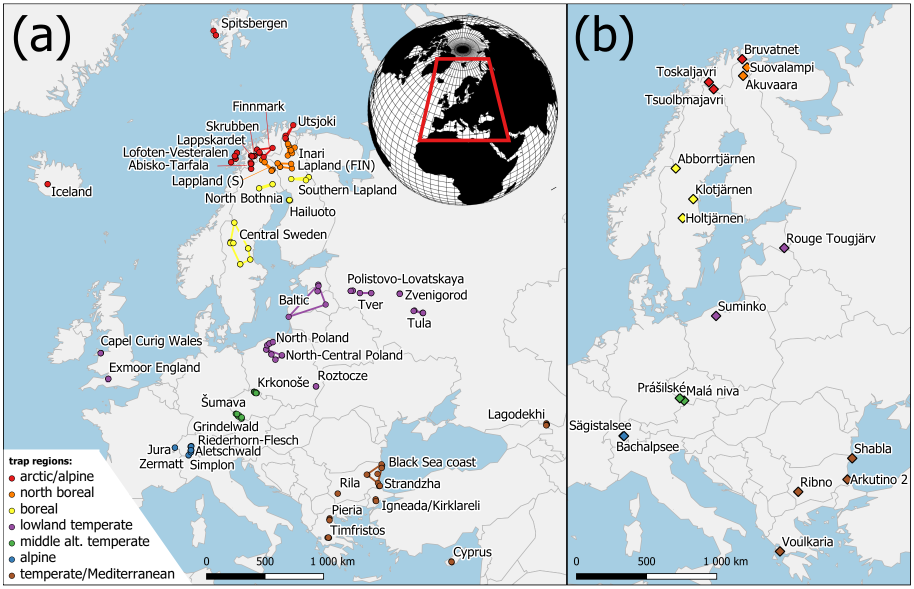

**Figure 1.** Map of the study area ~~. Trap dataset is divided into~~ with trap ~~regions (colours) and trap~~ areas (~~labels~~a) ~~.~~ and Holocene pollen sites ~~selected for the comparison~~ (~~target symbol, italic labels – see Tab.1~~b). Both datasets are divided into trap regions (colours). – Colours correspond to Fig. 5.

individual researchers. The paleoecology database Neotoma (**?**) offers a platform to store the PMP data and make it available to researchers worldwide, allowing them to interrogate the data and potentially identify modern analogues to interpret fossil pollen accumulation rates. ~~The aim~~

The overall purpose of this manuscript is to present an overview of the data in the PMP database and ~~explore its potential to provide analogues for fossil situations. We collected fossil~~ to interrogate this continental scale dataset of modern PAR with the following aims:

a) Examine the hypothesis that climate as well as regional forest cover explain the variability in PAR.

b) Study the absolute amount of long distance dispersed pollen encountered in pollen traps beyond the known distribution limits of the parent trees.

c) Compare modern and fossil pollen accumulation rates by collecting fossil datasets with estimates of ~~pollen accumulation rates~~ PAR from the same regions where the pollen traps were installed~~to explore the relationship between modern and fossil pollen accumulation rates.~~. For the most abundant pollen types we explore how the modern situations can provide a reference for the interpretation of the fossil data.

## 2 Methods

### 2.1 Study area

Sites in the PMP database were divided into 7 ‘trap regions’ according to the latitude and altitude. These regions were further divided into ‘trap areas’, by grouping 2-10 trap locations according to their spatial proximity. The arctic/alpine region includes distant trap areas in Spitsbergen and Iceland, northernmost traps in Finland (Utsjoki) and Norway (Lofoten-Vesterålen) and traps in the Scandinavian mountains above the local tree-line (Finnmark, Abisko-Tarfala, Skrubben and Lappskardet). The landscape around these traps is often treeless or covered by sparse birch shrubland with *Betula nana* and *B. pubescens* in some locations. The north boreal region includes traps in Northern Lapland in Finland, Norway and Sweden with a vegetation dominated by *Betula* and *Pinus sylvestris*. The northern limit of *P. sylvestris* occurs between traps from this region, which are situated at altitudes below 500 m a.s.l. The boreal region includes trap areas situated in Southern Lapland, around the Bay of Bothnia and in Central Sweden. The vegetation is dominated by *Picea abies*, *P. sylvestris* and *Betula* species, with the occurrence of *Alnus incana*. Northernmost populations of *Alnus glutinosa* occur near some sites and the southernmost traps in Central Sweden are situated near northern outpost populations of *Corylus avellana* and *Ulmus glabra*. Traps from the temperate lowland region have the widest longitudinal extent including the British Isles, Poland, the Baltic countries and European Russia. Vegetation at trapping locations below 500 m a.s.l. is characterised by *Quercus* and *Fraxinus excelsior* in the West and *P. sylvestris*, *P. abies* and *Betula* with an admixture of *Quercus*, *Tilia*, *Ulmus* and *C. avellana* in the East. *Fagus sylvatica* and *Carpinus betulus* occur in Poland, *Abies alba* only in south-eastern Poland. Trapping locations in the temperate region at elevations above 500 m (mid elevation) were separated and include the Krkonoše and Šumava Mountains. Traps in both areas are placed on an elevation gradient from 500 m and 1200 m a.sl. The lower slopes of the mountains are dominated by *F. sylvatica*, while the traps are situated on a gradient from *P. abies*-dominated forest to the onset of alpine vegetation in Krkonoše. In the Alps and Jura Mountains, traps were placed at even higher elevations, between 1200 m to 3000 m, crossing the altitudinal treeline. Trapping locations in the Southeast represent diverse landscapes and vegetation types including grasslands, evergreen and deciduous forests. Some traps are situated in high mountain regions around treeline situations or within the upper mountain forests including Rila (Bulgaria), Pieria (Greece), Timfristos (Greece), Lagodheki (Georgia) and Cyprus. Traps at lower elevations are situated near the Black Sea coast, within the low Strandzha Mountains and European Turkey.

### 2.2 Data collection

The pollen traps used in the PMP network generally consist of a bucket or bottle large enough to contain the annual surplus in precipitation on a surface of usually 19.6 $cm^2$ (5 cm diameter opening) or similar. Many traps had a sloping collar inspired by the design of pollen traps by Tauber (**?**), although few collars were truly aerodynamic. The collection of the trap content was generally carried out annually and any special circumstances potentially affecting the annual pollen deposition were noted and stored in the database. For the analyses presented here and data overview we excluded traps where the pollen record is 2 years or less, as averages may be affected by high inter-annual variability. The only exception is the trap situated in Spitsbergen where there is a two-year record. Pollen accumulation from the two-year record shows little variation and, being the only analogue for a truly arctic and treeless environment, provides important information on long distance pollen transport. We also excluded annual samples with records shorter than 8 months and, in addition, traps or years with spurious values due to particular events or local conditions (Table S1).

Most of the traps in the PMP network were placed in the open vegetation or in forest openings in order to avoid an unrepresentative contribution of individual trees e.g. due to anthers dropping into the traps. Traps were generally installed at ground level mimicking collection conditions relevant for sedimentary archives. In consequence tall herbs or grasses might overgrow or cover some of the traps potentially leading to higher pollen deposition. Traps not equipped with a mesh occasionally trapped pollen collecting insects, leading to enormous counts of insect pollinated taxa e.g. *Calluna, Erythranthe guttata*. The presence of insects in the traps is usually noted for the collection year so that careful evaluation of the information in the database can also inform on herbaceous pollen types (**?**). Including this information in comprehensive database queries is currently not possible and a manual screening of datasets is required when analysing herbaceous pollen types. This problem does not seem to occur in tree pollen taxa. The occurrence of phytophagous insects in the traps were not accompanied by unusual peaks in tree pollen taxa, indicating that the insects inadvertently trapped were primarily collecting pollen from the herbaceous vegetation around the traps. Comprehensive database queries restricted to tree pollen, Poaceae and Cyperaceae should therefore not be affected by the occurrence of insects in the trap and mainly represent pollen transport via wind, the rainout of pollen from the atmosphere and the gravity component.

Concentrating the content of the traps was carried out either using filter paper or centrifugation and decanting the supernatant. In many cases the trap content was washed onto a paper filter, which was later digested using acetolysis. Pollen quantity was assessed by adding *Lycopodium* spore tablets (**?**) to each trap before processing. Pollen concentration was obtained from the ratios between pollen grains counted to *Lycopodium* spike counted and *Lycopodium* spike added. Details about *Lycopodium* spike data, as well as details of the pollen trap such as the exact size of the opening are stored in the database. The PMP database was created in the PostgreSQL database system. Names of pollen taxa were unified using the accepted variable names from the European Pollen Database (**?**).

### 2.3 Investigated taxa and environmental parameters

We selected the common tree and shrub taxa of Europe. Pollen taxa generally refer to all the species within the genus. Pollen taxa allowing higher taxonomical resolution, which were consistently separated and excluded from the genus in the whole dataset are marked as “excl.”. Pollen taxa potentially including pollen grains from another genus are indicated by “incl.”: *Abies*, *Alnus* (excl. *A. viridis*), *Betula* (excl. *B. nana*-type), *Carpinus*-type (incl. *C. orientalis*/*Ostrya*-type), *Corylus*, *Fagus*, *Fraxinus* (incl. *F. ornus*), ~~*Juniperus*-~~*Juniperus*~~type-~~-type (incl. *Cupressus*, *Tetraclinis*, *Thuja*), *Picea*, *Pinus* (excl. *P. cembra*-type), *Tilia*, *Quercus* (incl. *Q.robur*-type, *Q. cerris*-type and *Q. ilex*-type). Pollen accumulation rates of trees and shrubs were summed as arboreal pollen accumulation (hereafter as “tree PAR”). We also included pollen from the plant families Cyperaceae and Poaceae (excluding cereals). For the purpose of the analysis in this paper we refer to sum of tree PAR plus Cyperaceae and Poaceae as “total PAR”.

The climate parameters Mean Annual Temperature (MAT) and Annual Precipitation (APrecip) for the trapping locations were obtained from WorldClim 2 (**?**). For site altitude we used the information supplied by the individual investigator. Comparisons between PAR and forest cover were conducted using the data of the Forest Map of Europe (**?**), which has a grid resolution of 1 km$^2$. Forest cover was extracted ~~at a spatial level~~ as a mean of all grid cells within a 10 km radius. We used regression analysis to explore whether individual or combinations of these environmental parameters describing the trapping location can explain the variance in average pollen accumulation of the traps. To balance the contribution of high and low pollen producers in the assessment of the total PAR we applied correction factors (Table S2, **?**).

Pollen deposition beyond the distribution area of the parent plant was studied by merging the distribution maps of the relevant species included in the pollen type described above ~~(**??**)~~(**??**). These comparisons were not suitable for *Alnus*, *Betula*, Cyperaceae, *Juniperus*, *Pinus* and Poaceae as these taxa are widely distributed in Europe and few traps are located beyond their distribution area. We compared the amount of pollen accumulating from these taxa in pollen traps at 200 km from their mapped distribution limits. Pollen traps in the UK are situated beyond the natural distribution limits of several of these trees but were excluded from the comparison as the taxa considered may be planted in the area.

For each trap location and each pollen taxon we calculated the distance to the nearest area of distribution using GIS (**?**). Linear regression between this distance and the decadic logarithm of PAR was used to predict the threshold of long-distance transport (hereafter also as “LDT”).

### 2.4 Comparison between modern and Holocene PAR

~~We searched for Holocene pollen records containing estimates of PAR in proximity to the locations of pollen traps. From the available sites~~ To enable the comparison of modern with fossil PAR values the pollen trap data was extracted from the PMP database with above described constraints and all annual samples were averaged within traps. Per each trap region we selected at least one and a maximum of three Holocene PAR records ~~per trap region~~ (Table 1). Holocene PAR estimates often show high variation between samples due to changes in the sedimentary environment. To reduce this effect ~~of the here conducted~~ in this comparison Holocene data were averaged in 500-year bins.

Site and sample compilation resulted in ~~PAR values from 354 Holocene and 271 modern samples~~a fossil dataset containing 354 Holocene samples.

We compared trap and fossil PAR datasets in two ways. First, we compared the distribution of logged PAR values and their mean for all taxa in all regions. Differences were measured by t-test at 5% level of alpha. Second, we searched for Holocene pollen records containing estimates of PAR in proximity to the locations of pollen traps, so we compared trap and fossil PAR at level of individual sites or trap areas. Average trap and fossil PAR values per taxon were submitted to one-dimensional clustering using the R-package Ckmeans.1d.dp (**?**). The classes produced were used to facilitate the comparison between trap and fossil data and to match the ~~modern~~ trap values with analogous situations in the past. The aim of this comparison was to find traps with similarly high values for individual taxa that compared to the highest average fossil PAR. However, averaging did not smooth out all spuriously high values variation~~.ignoring~~ , thus we need to ignore individual high fossil values ~~as described above~~ (Table S3). Thus, we aimed to find modern analogues for fossil situations represented by several bins (more than 500

**Table 1.** Fossil sequences including type and size of the deposit.

| country | region | site | deposit | latitude | longitude | (m) a.s.l. | area (ha) | reference |
|---|---|---|---|---|---|---|---|---|
| FIN | arctic/alpine | Bruvatnet | lake | 70.17933 | 28.39998 | 119 | 60 | (?) |
| FIN | arctic/alpine | Toskaljavri | lake | 69.19177 | 21.44841 | 704 | 100 | (?) |
| FIN | arctic/alpine | Tsuolbmajavri | lake | 68.68915 | 22.05235 | 256 | 14 | (?) |
| FIN | north boreal | Suovalampi | lake | 69.58333 | 28.83333 | 104 | 16 | (?) |
| FIN | north boreal | Akuvaara | lake | 69.125 | 27.68333 | 170 | 4 | (?) |
| SWE | boreal | Abborrtjärnen | lake | 63.88333 | 14.45 | 387 | 3 | (?) |
| SWE | boreal | Klotjärnen | lake | 61.81667 | 16.53333 | 235 | 1 | (?) |
| SWE | boreal | Holtjärnen | lake | 60.65 | 14.91667 | 232 | 1 | (?) |
| EST | lowland temp | Rõuge Tõugjärv | lake | 57.73904 | 26.90515 | 114 | 4.2 | (?) |
| POL | lowland temp. | Suminko | lake | 53.72556 | 17.77278 | 115 | 0.12 | (?) |
| CZE | middle alt. temp. | Prášilské | lake | 49.07551 | 13.40002 | 1079 | 3.7 | (?) |
| CZE | middle alt. temp. | Malá niva | peatbog | 48.90789 | 13.81982 | 754 | 65 | (?) |
| CHE | mountain temp. | Sägistalsee | lake | 46.68139 | 7.9775 | 1935 | 7.2 | (?) |
| CHE | mountain temp. | Bachalpsee | lake | 46.66944 | 8.020833 | 2265 | 8 | (?) |
| BGR | temp/medit. | Shabla | lake | 43.58333 | 28.55 | 1 | 1.51 | (?) |
| BGR | temp/medit. | Arkutino 2 | lake | 42.3299 | 27.72363 | 0 | 40 | (?) |
| BGR | temp/medit. | Ribno | lake | 42.20682 | 23.32346 | 2184 | 3.5 | (?) |
| GRC | temp/medit. | Voulkaria | lake | 38.866667 | 20.833333 | 0 | 10000 | (?) |

years)~~representing a period of long-term vegetation stability~~. We linked these periods with high fossil PARs to the closest pollen trap, using a ~~distance matrix~~ matrix of distances between fossil sites and pollen traps. We selected the comparisons for the 8 most abundant taxa to be presented in the main text and show the remaining 7 taxa in the supplementary. All statistical analysis and data visualizations were produced in R (?).

## 3 Results and interpretations

### 3.1 ~~Spatial pattern~~ Overview of ~~absolute pollen deposition~~the PMP database and the environments sampled

The PMP database version 02.02.2020 contains data from 351 trap locations with a total of 2742 annual samples covering the period from 1981 to 2017. Considering the trap record with 3 years and more we obtained 271 mean trap assemblages. Trapping sites cover a range of altitudes from 0 to 3000 m a.s.l. with annual precipitation ranging from 402 to 1549 mm. Mean annual temperature (MAT) for the sites fall between -5.7 to 14.1 °C. The forest cover within a 10 km radius of the trapping sites ranges from 0 to 98%. This range of environmental situations has yielded tree pollen accumulation rates from 5 to 86000 grains cm$^{-2}$ y$^{-1}$, with a median value of 5400 grains cm$^{-2}$ year$^{-1}$ (Fig. 2). An overview of the taxonomic composition of the traps (Fig. S1) shows a dominance of pollen from *Pinus* and *Betula* in the traps from boreal and hemiboreal environments, with *Betula* as the taxon with the highest PARs overall. In most northern traps from open environments Cyperaceae is the dominant NAP pollen type while Poaceae are dominant in traps from open environments in the south, where they also contribute much higher absolute amounts. The diversity of landscapes and forest types in central and southern Europe is well represented in the pollen composition of traps from this area. Differences in pollen composition and abundance between the high mountain forests of central and southern Europe and the boreal forest in the north are also noticeable~~. The pollen composition of the fossil sequences selected represents the same regional patterns as the traps~~ (Fig. S2). ~~Temporal changes are most pronounced in both fossil sequences from the Alps, documenting forest compositional changes as well as the suppression of the tree-line during the Holocene.~~

### 3.2 Dependence of variation in PAR on regional forest cover and climate

Total PAR is generally lower at high latitudes, with the lowest values in the arctic alpine region (trap area Spitsbergen). However, the highest absolute values are not from the southernmost traps but from the lowland temperate region (trap area Tver; Fig. 2). Nevertheless, latitude alone explains about 11% of the variance in log transformed ~~total~~ tree PAR, while MAT and forest cover within 10 km explain 21% and ~~72~~19% respectively. In combination, these three variables explain ~~76% and with~~ 37% of the variation in log-transformed absolute tree pollen deposition. The addition of elevation ~~73% of the total variance in log transformed total PAR~~increased the amount of variance explained to 50% (Table S4a).

Large differences in the pollen productivity between different trees effect this relationship. Adjusting the PAR from individual taxa by Andersen factors reduces the bias of differential pollen production between different plants and ~~thus~~ makes it possible to consider the total amount of pollen deposition including grasses. This adjustment increases the amount of variance explained by the regression model with all 4 explanatory variables to ~~82~~56% (Table ~~S4). This adjustment reduced the individual~~ S4b). Due to the inclusion of grasses the explanatory power of forest cover within 10 km ~~, due to the inclusion of grasses in the total PAR. On the other hand,~~ is reduced, while latitude alone explains ~~38~~37% of the Andersen adjusted log transformed total PAR (Fig. 3a, Table S4b).

The regression models consider the full range of the data, while due to local factors there is often a spread of average trap values for different traps in the same region. ~~Traps with maximum values per region~~ The traps with the highest regional values do not follow a latitudinal pattern, the distribution of the minimum average trap values are more informative (Fig. 3a). These lower values closely follow a latitudinal trend. The average PAR south of 62° latitude and below the altitudinal treeline or close to forests are generally higher than 1000 grains cm$^{-2}$ y$^{-1}$. An area with low PAR in the south is the coastal grassland in northern Bulgaria. The generally low PAR in this area can be explained by the sparse vegetation cover on thin rendzina soils formed on limestone rock. Adjusting the PAR values by Andersen factors increases the values for this region so that they fit the general latitudinal trend ~~.~~(Fig. 3a). Traps with minimum average PAR values per region also correspond well to the forest cover within 10 km (Fig. 3b). ~~Taking the traps with minimum PAR within~~ Considering 3% wide bins of the forest cover~~data provides~~, traps with the lowest PAR per each bin of the forest cover provide a regression model predicting a tree PAR of 3200 grains cm$^{-2}$ year$^{-1}$ at 80% forest cover within 10 km of the trap.

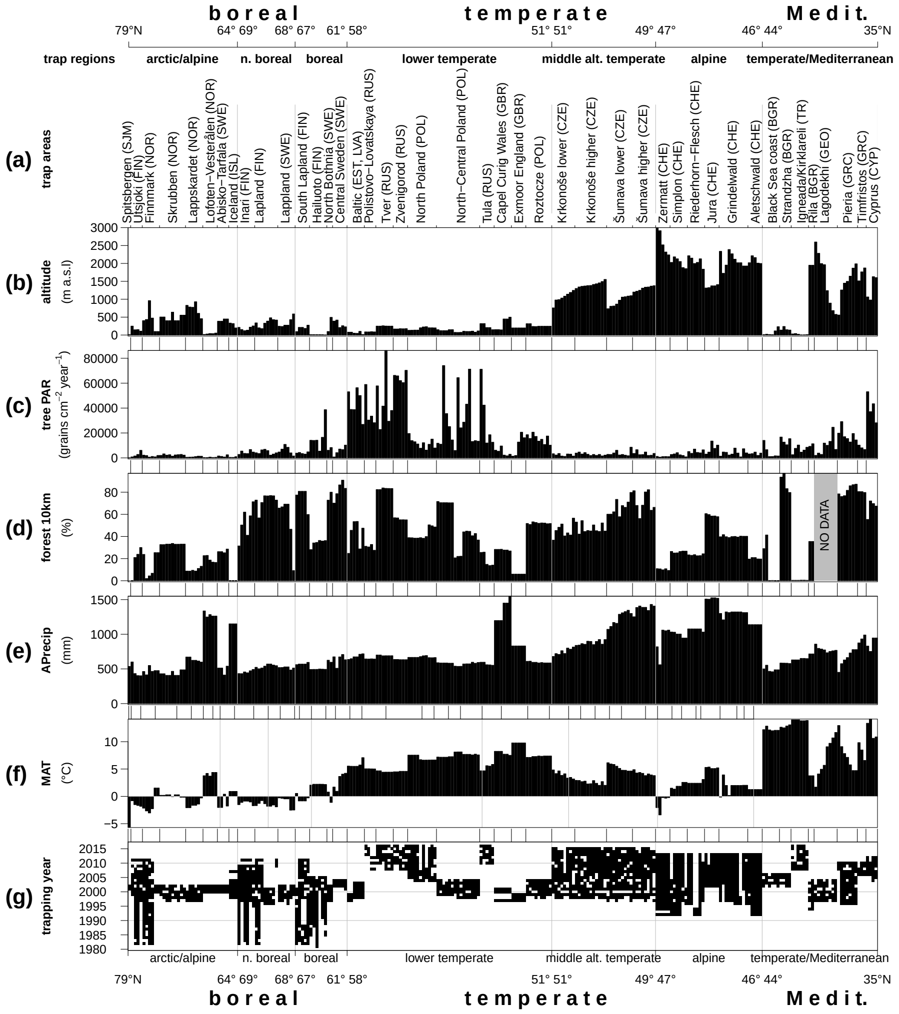

**Figure 2.** Environmental setting of the trap dataset. (a) trap areas ordered from north (left) to south (right), (b) altitude, (c) mean annual tree PAR, (d) forest cover within 10 km radius, (e) annual precipitation, (f) temperature: MAT - Mean Annual Temperature, (g) temporal coverage of the PMP database.

~~Environmental setting of the trap dataset. (a) trap areas ordered from north (left) to south (right), (b) altitude, (c) mean annual tree PAR, (d) forest cover within 10 km radius, (e) annual precipitation, (f) temperature: MAT - Mean Annual Temperature, (g) temporal coverage of the PMP database.~~

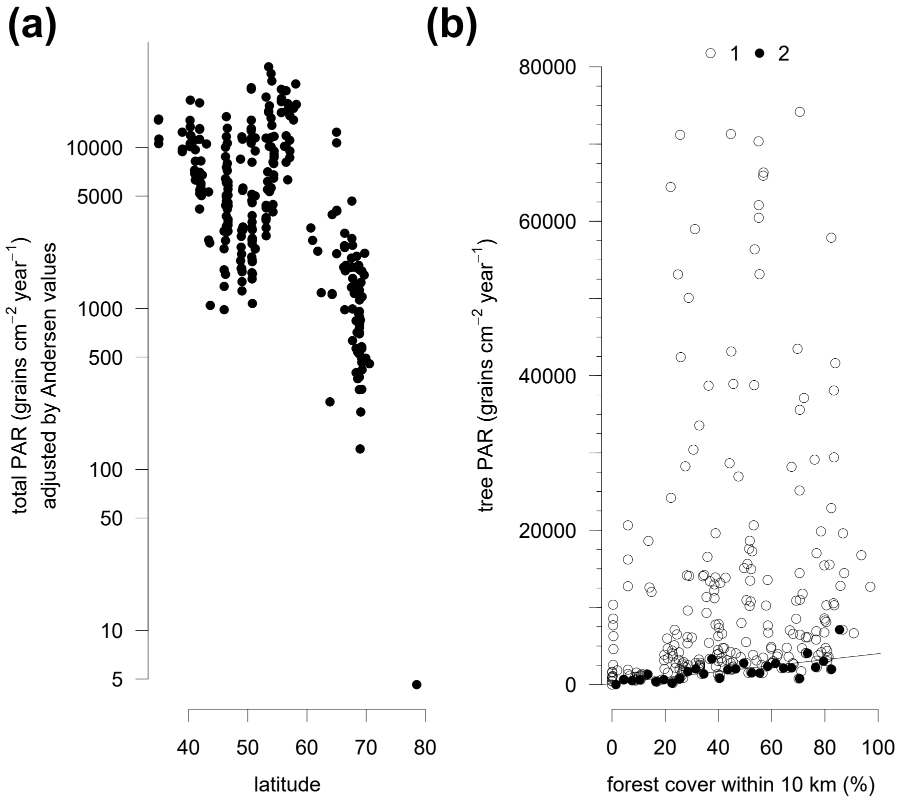

**Figure 3.** Latitudinal gradient in pollen accumulation rates of major tree taxa and Poaceae and Cyperaceae (total PAR) adjusted by Andersen values (see Tab S2) of pollen representation (a). Relationship between forest cover within 10 km radius and tree PAR (b). All trap sites (1) and minimum tree PAR per every 3 % of forest cover (2).

### 3.3 Long distance dispersed pollen

The comparison of PAR with the distribution limit of different tree taxa shows that PAR generally decline with distance (Fig. 4). A gradual decline is best documented for *Quercus* where average values at the distribution limit scatter around 100 grains cm$^{-2}$ y$^{-1}$. This analysis also documents the long-distance transport of many tree pollen, including the heavy pollen of *Picea*. ~~Where pollen numbers are very low in this comparison there is a degree of uncertainty since the likelihood of encountering a grain depends on chance and is related to the number of pollen grains counted, which we did not consider in this analysis.~~ For better comparison of the absolute values between taxa we used regression analysis to estimate the amount of pollen at 200 km from the distribution limit (Fig. 4b). This comparison indicates that less than 30 grains cm$^{-2}$ y$^{-1}$ of *Carpinus, Corylus, Fagus,*

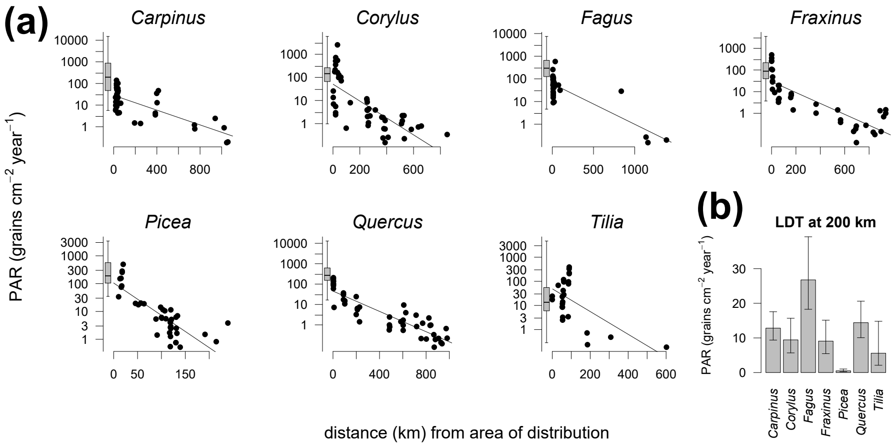

**Figure 4.** Relationship between the distance from the trap site to the nearest area of species distribution and PAR for selected trees. Zero distance represents edge of distribution area. Traps within the distribution area are aggregated in boxplots (a). PARs of the long-distance transport (b) calculated from linear regression at 200 km (Fig. 4a).~~Traps within the area of species distribution were excluded.~~

*Fraxinus, Quercus* and *Tilia* are deposited beyond 200 km of the distribution of the parent trees. Only *Picea* shows less than 1 grain cm$^{-2}$ y$^{-1}$ at 200 km of the distribution range.

### 3.4 ~~Recent~~ Ranges of modern and fossil PAR values~~at continental and regional level~~

~~Comparison~~ The here presented comparison of modern and fossil PAR values ~~for the dataset presented here produces good agreement with similar peaks in total PAR~~shows good agreement in tree PAR. The highest frequency of tree PAR values ranges between 2000 and 10000 grains cm$^{-2}$ y$^{-1}$ in both datasets (Fig. 5). Maximum PARs ~~in the pollen traps are always~~ of the trap dataset are higher (often ten times~~higher) compared to fossil situations~~) for all species compared to maximum PARs in the fossil dataset, with the exception of *Corylus*. The frequency distribution of PARs is log normal for *Alnus*, *Tilia* and *Fraxinus*. A bimodal distribution of values occurs for some taxa, which is particularly clear for the fossil values of *Abies*, *Picea* and Poaceae. In the modern samples such a bimodality can be recognized for *Pinus* with a trough at around 1000 grains cm$^{-2}$ y$^{-1}$. The frequency distribution of modern and fossil PARs corresponds best for *Corylus*, with most values falling between 100 - 300 grains cm$^{-2}$ y$^{-1}$ and recent and fossil maxima at around 3000 grains cm$^{-2}$ y$^{-1}$. The greatest difference in the distribution of modern versus fossil PARs occurs for *Juniperus*, where maximum values are around a hundred times larger in the traps. Minimal PARs are about ten times higher in the traps for Poaceae and Cyperaceae in particular and the right side of the distribution is shifted upwards. For *Fagus, Quercus* and *Carpinus* the fossil PARs show a local maximum in the frequency of low values, which does not occur in the traps. These frequent low values range below the threshold indicating long distance transported pollen.

~~There is a poor correspondence between modern and Holocene PARs for the individual trapping regions.~~ Using the 15 selected taxa in seven trap regions and the occurrence of the pollen types we obtained 92 pairs; of these 31 pairs are similar based on a t-test and a p-value > 0.05 (Fig. ~~6~~5, Table S5). In this regional comparison *Betula* shows the best agreement between modern and fossil values. Values are similar in four regions across the gradient, with highest values of 5400 grains cm$^{-2}$ y$^{-1}$ in the lowland temperate region and the lowest values of 34 grains cm$^{-2}$ y$^{-1}$ in the temperate/Mediterranean region, where the parent trees are generally absent. Modern and fossil PARs generally correspond well for the lowland temperate region where, in addition to *Betula*, *Alnus*, *Carpinus*, Cyperaceae, *Fraxinus*, *Picea*, *Pinus* and *Quercus* also have similar values. Although *Corylus* has a good overall agreement, the regions with similar modern to fossil data are shifted, with Holocene values in the boreal region corresponding to modern PARs in the lowland temperate region.

## 3.5 ~~Recent and fossil~~ Taxa specific linkage of the highest average PAR ~~values~~ at ~~the levels of~~ fossil sites ~~, traps and~~ with individual trap ~~areas~~values

~~The~~

### 3.5.1 Introductory notes

To facilitate the comparison of modern and fossil PAR the combined taxa specific values were submitted to a one-dimensional cluster analysis~~distinguished~~, which resulted in between 5 and 9 classes of PAR values per taxon ~~.~~ (Fig. S1-S2). Comparing the maximum averages of fossil PARs to modern trap data on a site by site basis shows that it is possible to find modern comparisons for all fossil situations. ~~Fossil PARs from sites in the arctic and boreal region often match with modern values in pollen traps situated south of the sites. The frequently high modern values in pollen traps from northern and central Poland in particular provide frequently good matches for fossil situations. The spatiotemporal pattern for *Juniperus* stands out, with highest fossil values in the Early Holocene and core top samples having analogues at the latitudinal and altitudinal treelines.~~ We demonstrate the linkage of the highest PAR clusters per each fossil site with individual trap records ~~on example of main~~ for the example of selected tree taxa (*Abies*, *Betula*, *Corylus*, *Fagus*, *Picea*, *Pinus*, *Quercus*, *Tilia*; Fig. 6). Detailed description of the ~~rest of taxa see~~ remaining of taxa are presented in the supplementary (*Alnus*, *Carpinus*, Cyperaceae, *Fraxinus*, *Juniperus*, Poaceae, arboreal pollen; Fig. S3).

## 3.6 ~~Taxa specific linkage of the highest average PAR at fossil sits with individual trap values~~

### 3.5.1 *Abies*

Modern PAR of 490-3900 grains cm$^{-2}$ y$^{-1}$ are observed in Roztocze, Jura, Rila and Timfristos. These values are produced by the different species: *Abies alba* in Roztocze and the Jura mountains and *A. cephalonica* in the Rila mountains and on Timfristos. Generally high fossil values occur in the two Alpine lakes Sägistalsee and Bachalpsee around 7000 years, in Ribno in the Rila mountains around 5000 years and in the two sites in Šumava between 4000 and 1000 years. In all these regions with fossil evidence of high *Abies* populations the modern PAR values in pollen traps are comparably low documenting that the populations have much declined. The *Abies alba* populations in Roztocze provide modern analogues for how dense *Abies* forests may have been in the Alps and Šumava although the forests in Roztocze occur at much lower elevations. Pollen traps

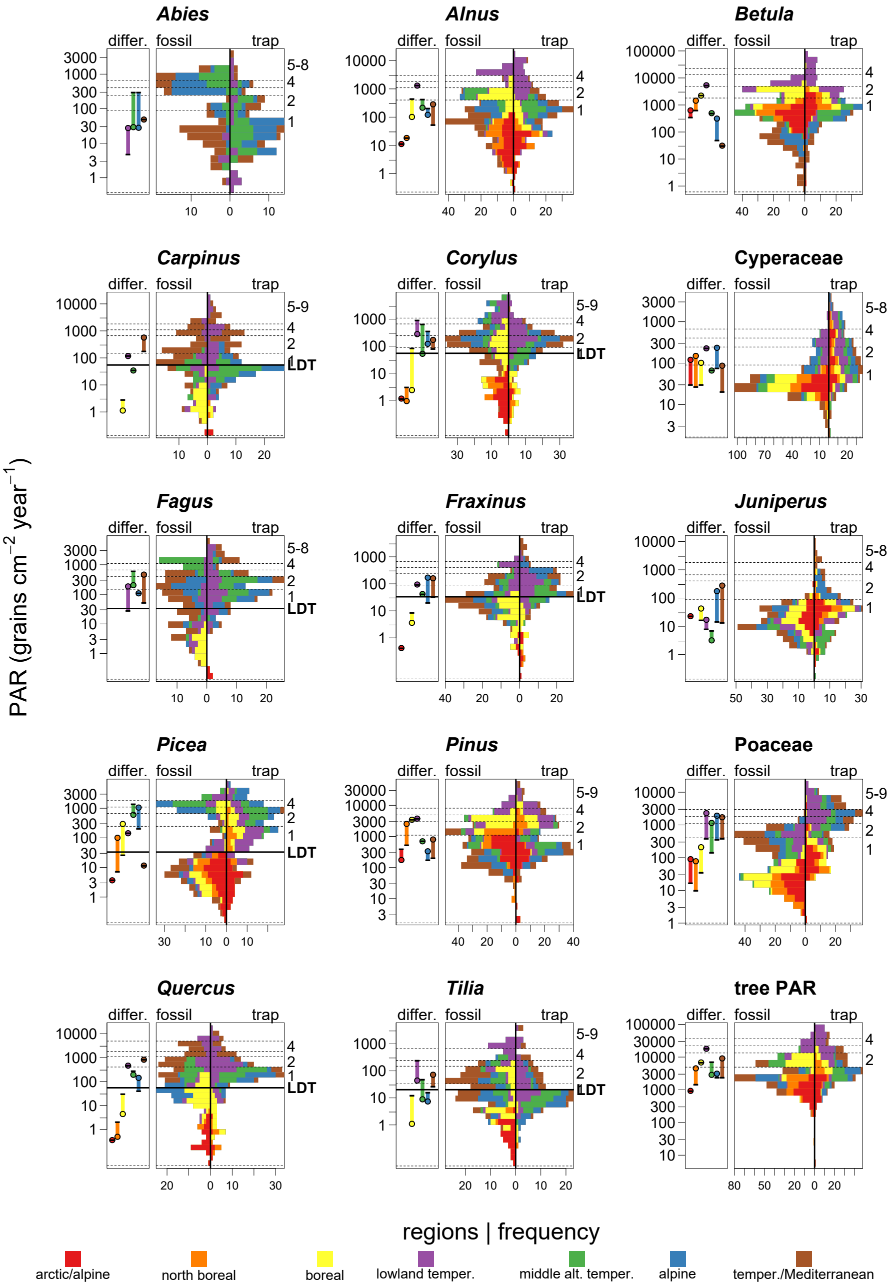

**Figure 5.** Difference (differ.) between the mean fossil (-) and the mean trap (o) ~~pollen influx~~ PAR per trap region is shown by length of the vertical segment. Paired histograms of mean annual ~~pollen influx~~ PAR from fossil record (on the left) and from traps (on the right). Colours denote different trap regions and correspond to Fig. 1. Note logged y-axis. Horizontal lines and numbering on the secondary y-axis denote classes of PAR, for more detail see Fig. 6b) and d). LDT is threshold for long distance transport.

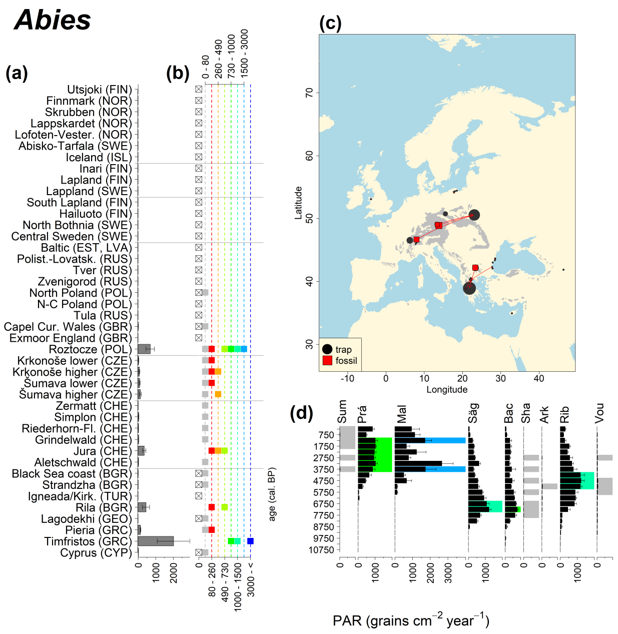

**Figure 6.** Mean modern PARs averaged for each trap area (a). b) Range of mean individual trap values classified by one-dimensional clustering. Crossed squares indicate that pollen of the taxon was not found in any trap from the area. c) Map of Europe with the distribution of the species (gray, Caudullo et al., ~~2018~~2017, San-Miguel-Ayanz et al. 2016) falling within the pollen taxa, size of symbols shows classes of PAR in recent and the highest PAR per each fossil record. Arrows show the closest trap with the same class of PAR. d) Fossil PAR values with the highest PAR class per each record (see Table 1 for full name) highlighted by the corresponding colour for the class (see b) Note the scale of the x-axis corresponds to the x-axis scale of graph a).

on Timfristos mountain provide analogues of the density of Middle Holocene *Abies* forests in the Rila mountains. Pollen traps located far from the distribution limits of *Abies* in Wales, northern Poland, Georgia and Cyprus registered occasional *Abies* pollen grains with PARs of up to 80 grains cm$^{-2}$ y$^{-1}$.

### 3.5.2 *Betula*

~~Ignoring~~ Letting aside traps from the Caucasus and Turkey, trap sites cover the distribution of *Betula pendula* and *B. pubescens*. The highest modern values between 11900-73900 grains cm$^{-2}$ y$^{-1}$ are found in Estonia and Russia as well as in one pollen trap from Hailuoto. ~~As discussed in the main manuscript these~~ These modern PARs exceed values from fossil examples. The highest fossil PAR of around 10000 grains cm$^{-2}$ y$^{-1}$ in this comparison come from northern Poland and Estonia and are thus consistent with the area of high modern values. Whether high abundance of *Betula* is a characteristic of the eastern European forests or a result of frequent disturbance at the forest ecotone or due to recent land-use change is difficult to evaluate based on the available data. Modern and fossil values agree for the sites in central Sweden at PARs between 1900-5600 grains cm$^{-2}$ y$^{-1}$.

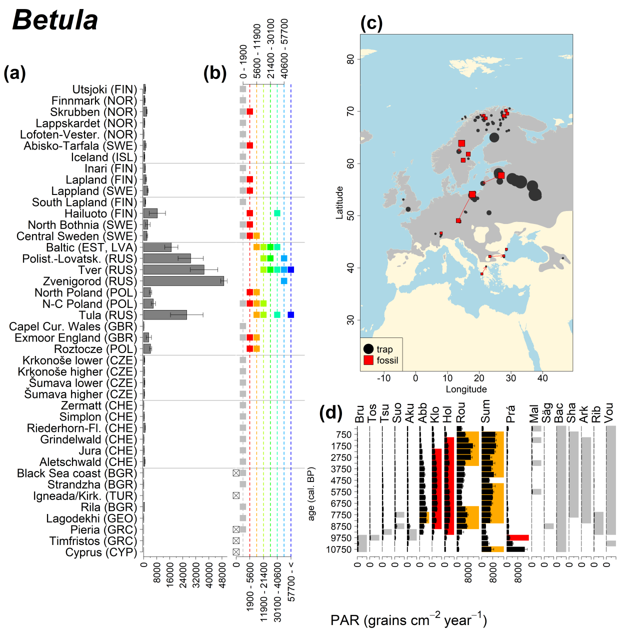

**Figure 7.** See caption Fig. 6

### 3.5.3 *Corylus*

Modern PARs stay below 2800 grains cm$^{-2}$ y$^{-1}$ except for two traps in north-central Poland. Values above 610 grains cm$^{-2}$ y$^{-1}$ are also found in pollen traps from the Baltic, Wales, Turkey and Georgia. The pollen type mainly comes from *Corylus avellana*, while *C. maxima* occurs in Greece, *C. colchica* in Georgia and *C. colurna* in Turkey and in plantations or as ornamental trees in the rest of the Europe. As discussed in the main text, the selection of fossil sites did not include studies from western and lowland central Europe where values of 10000 grains cm$^{-2}$ y$^{-1}$ are common for the Early Holocene. The highest fossil PARs from the chosen examples were estimated for the Early Holocene from Prášilské situated at 1000 m a.s.l. Thus the high modern values in north-central Poland provide analogues for several fossil situations. Occasional grains and small PARs of *Corylus* pollen are common in traps from boreal regions as well as two traps from the arctic region.

### 3.5.4 *Fagus*

PARs of more than 3300 grains cm$^{-2}$ y$^{-1}$ are found in pollen traps in Poland and Georgia and values above 2100 occur in traps from the Strandza, Šumava and Jura mountains, in southern Bulgaria, Czechia and Switzerland respectively. Two species of *Fagus* contribute to the pollen type with *Fagus sylvatica* as the dominant tree across much of Europe and *F. orientalis* occurring in southern Bulgaria, Turkey and Georgia. Fossil PAR increase around 8000 cal. BP at the Bulgarian Black Sea coast, around 7000 in the Šumava Mountains and 1500 years ago in northern Poland. High values in Šumava and northern Poland match the values in adjacent traps. Hardly any *Fagus* pollen occurs in pollen traps outside its modern distribution except sporadic

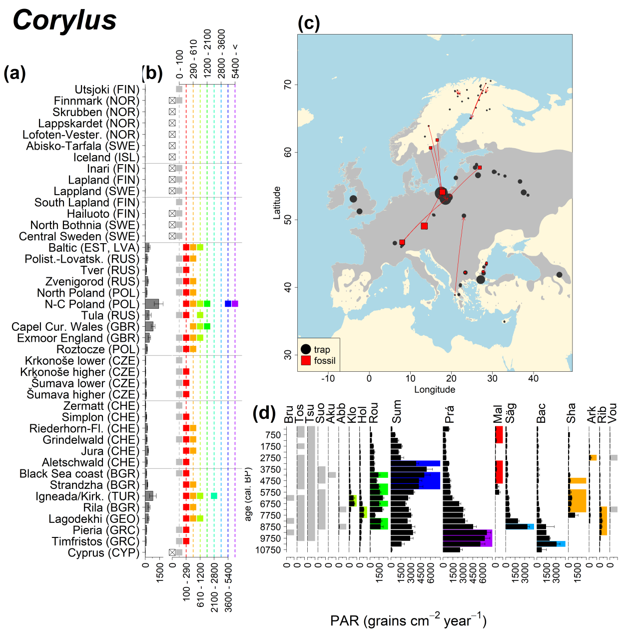

**Figure 8.** See caption Fig. 6

appearances in Tula and the two northernmost trap areas. However, *Fagus* pollen occurs regularly at fossil sites that were assumed to have never been within the distribution of the tree, such as the sites in central Sweden, where *Fagus* grains occurred regularly over the last 3000 years.

### 3.5.5 *Picea*

Modern *Picea* PARs above 2800 grains cm$^{-2}$ y$^{-1}$ are present in traps from Central Sweden, the Baltic region, Russia, Šumava mountains and the Alps. While some of the modern *Picea* pollen may be released by planted non-native *Picea sitchensis* and *P. pungens*, most of the pollen comes from Picea abies, which is also planted in many European regions outside its natural distribution. The highest fossil values in the selection of sites come from the Šumava mountains and are comparable to high trap values from the same region although at lower elevations. Noteworthy are also the generally low fossil ~~Picea~~ *Picea* PARs for sites in Central Sweden and Tver region, where the tree is dominating or co-dominating the forest for the last 2000 years. Most pollen traps from beyond the distribution area of ~~Picea collect individual Picea~~ *Picea* collect individual *Picea* pollen.

### 3.5.6 *Pinus*

Highest modern PARs exceeding 43600 grains cm$^{-2}$ y$^{-1}$ are observed in traps on Cyprus, while the values in traps from the northern boreal forest often stay below 5400 grains cm$^{-2}$ y$^{-1}$. *Pinus* PAR values increase from Finnmark (2000 grains cm$^{-2}$ y$^{-1}$) and central Sweden (5000 grains cm$^{-2}$ y$^{-1}$) to the Baltic and north-central Poland with 35000 grains cm$^{-2}$ y$^{-1}$. In northern Europe nearly all *Pinus* pollen comes from *P. sylvestris*, while southern European trapping sites have a higher diversity of trees within

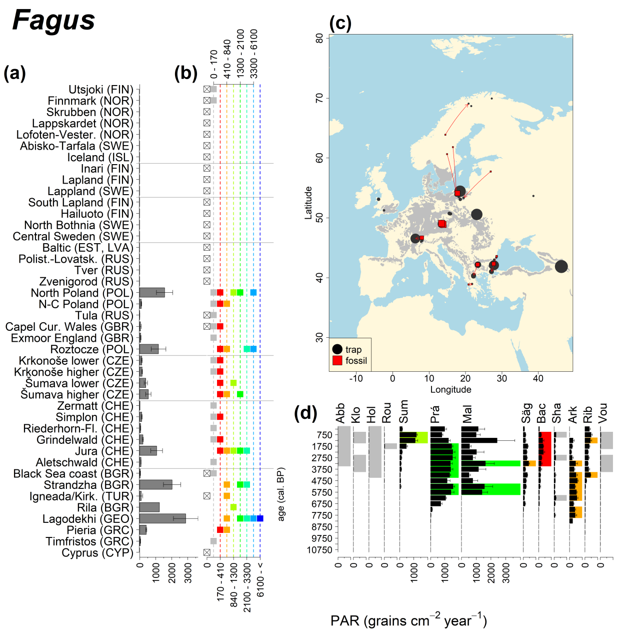

**Figure 9.** See caption Fig. 6

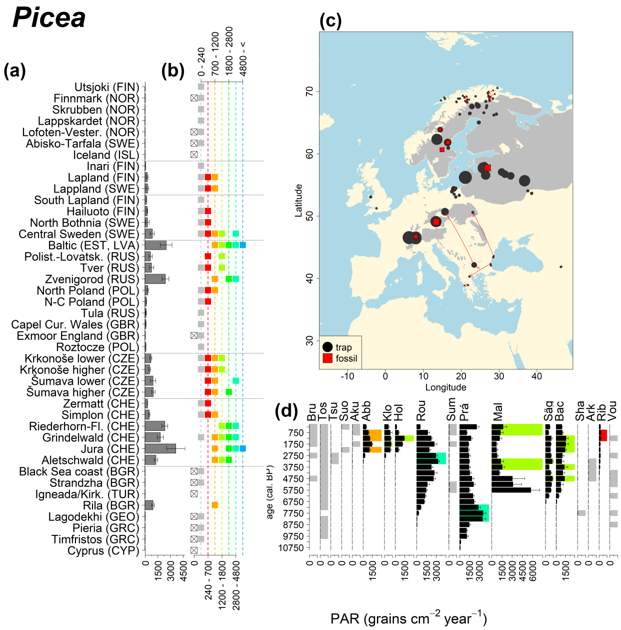

**Figure 10.** See caption Fig. 6

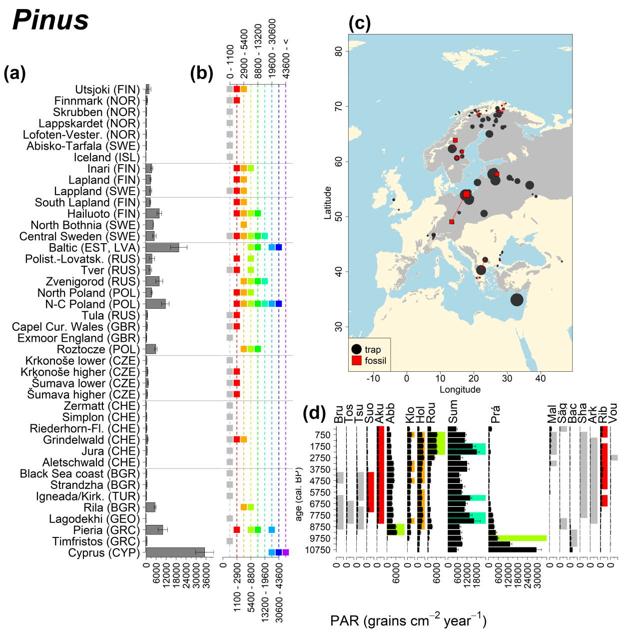

**Figure 11.** See caption Fig. 6

the subgenus Diploxylon including *P. mugo*, *P. nigra*, *P. brutia* and *P. halapensis*. Pollen of both Haploxylon pines *P. cembra* in Alps and *P. peuce* in Rila was separated. High fossil *Pinus* PAR values are estimated for Suminko in northern Poland, matching modern trap values from the same region. In the Šumava mountains *Pinus* was possibly the dominant forest tree during the Early Holocene and declined thereafter, so that modern values from north-central Poland provide the nearest analogue to the Early Holocene situation. The lowest *Pinus* PARs (<125 grains cm$^{-2}$ y$^{-1}$) are found in taps from Iceland (2-5 grains cm$^{-2}$ y$^{-1}$), Lagodheki and most of the traps from Lofoten-Vesterålen (38-125 grains cm$^{-2}$ y$^{-1}$) and single traps from Lappskardet, Exmoor England and Zermatt.

### 3.5.7 *Quercus*

Modern PARs of *Quercus* within the area of distribution of a parent tree in the genus range between 620-15000 grains cm$^{-2}$ y$^{-1}$. Highest modern values (6000-15000 grains cm$^{-2}$ y$^{-1}$) are found in the traps from the UK, Poland and the southern Balkan. The first two areas host only species belonging to the *Q. robur*-type, whereas the two latter also include species from *Q. cerris*-type and *Q. ilex*-type. The highest fossil values from Suminko seem too high for the region compared to other fossil samples, nevertheless modern values in pollen traps from the region provide comparable high PARs. Also, PARs in recent sediments from some lakes in north-eastern Germany show similar values (**?**). PARs to about 1300 grains cm$^{-2}$ y$^{-1}$ are estimated for Rõuge Tõugjärv near the distribution limit of *Quercus robur* where the tree only became abundant after 7000 years ago. *Quercus robur* reached its maximum abundance even later in central Sweden where the PAR at Holtjärnen around 3200 years ago suggest its presence. The pollen seems to disperse well and is found in small amounts in most pollen traps beyond its distribution area

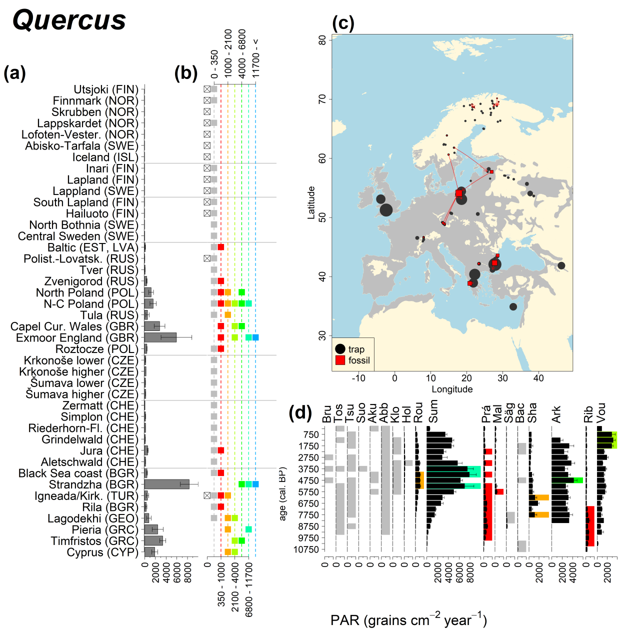

**Figure 12.** See caption Fig. 6

and fossils samples far to the north of the distribution have also collected *Quercus* pollen during the Holocene, which may be partly due to the abundance of trees of this genus in Europe.

### 3.5.8 *Tilia*

The highest modern PAR 1500-4700 grains cm$^{-2}$ y$^{-1}$ was measured in Poland, European Russia (Tula) and Lagodheki in Georgia, however, this high range has no comparison in the fossil record. Lower values 120-1500 grains cm$^{-2}$ y$^{-1}$ were found in traps from European Turkey and the Alps. Highest fossil PAR 270-1500 was measured in middle altitude and lowland temperate zone during the Middle Holocene. The closest trap analogues to them are in North Poland. Sites on the current edge of *Tilia* distribution and in the mountains (Central Sweden, Alps and Rila) show highest PAR range 120-270 grains cm$^{-2}$ y$^{-1}$, also during the Middle Holocene. Trap records corresponding to the lowest class and above the ~~LTD threshold~~ LDT (5-40 grains cm$^{-2}$ y$^{-1}$) appear in sites within distribution limit of *Tilia* (Baltic, European Russia, Poland), on its edge (Central Sweden, Wales, Black Sea Coast, Greece) or in the mountains (Šumava, Alps).

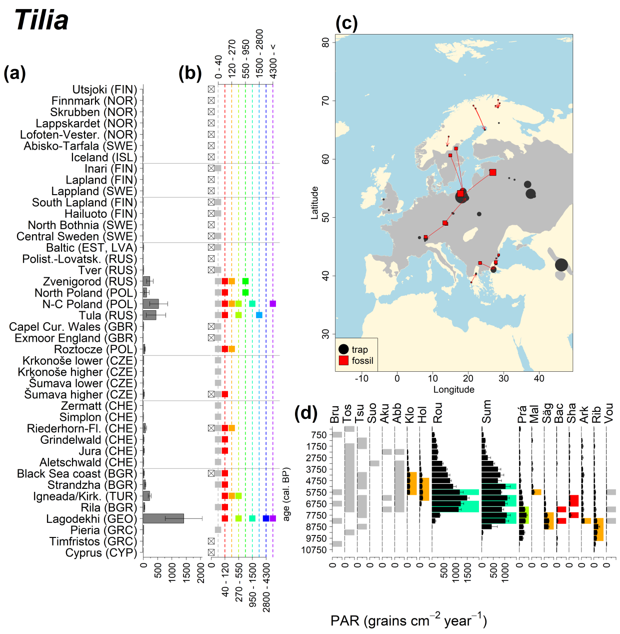

**Figure 13.** See caption Fig. 6

## 4 Discussion

### 4.1 ~~Analogues for vegetation reconstruction~~Trap PARs and environmental conditions

This overview of European pollen trap data collected by the PMP network demonstrates that modern PARs provide comparable values to fossil records and can thus help interpreting the fossil signal in spite of different taphonomic processes that influence PAR values in pollen traps versus lake sediments. The dataset extends across the European latitudinal and altitudinal range and ~~even if it lacks representation of western European vegetation types, it documents some~~ documents general patterns. The latitudinal gradient in PAR is clearly visible in this dataset. Although data on plant biomass and primary productivity are not available for all trapping locations the regression analysis indicates that mean annual temperature has ~~a strong~~ an influence on the quantity of pollen deposition. The July temperature of the previous year determines the amount of pollen production in *Pinus* near the tree-line (**??**). Evidence from other European regions (**??**) suggests that also growing season warmth and other climate variables ~~also~~ explain the interannual variability of pollen deposition. On the regional scale PAR corresponds to plant biomass of the parent tree ~~(**??**)~~(**??**). However, differences in ~~plant biomass~~ forest cover cannot explain the latitudinal gradient in PAR described here, which may, at least in part, result from the latitudinal gradient in primary productivity of trees (**?**)~~as previously suggested (**?**)~~. An increase in primary productivity and pollen production has been shown in a carbon dioxide fertilization experiment (**?**), which supports the interpretation that PAR of the same species may vary due to environmental parameters determining its productivity.

## 4.2 Long Distance Transport

Modern PARs from traps near the latitudinal limit of *Pinus* and *Betula* have been used previously to reconstruct past changes in the northern distribution limits of these trees (**?**). Here we evaluated larger distances and therefore had to ignore *Pinus* and *Betula*, while suggesting some general thresholds for other dominant European trees. The upper value of around 30 grains cm$^{-2}$ y$^{-1}$ for *Picea* agrees well with the fossil PAR value for the tree of 50 *Picea* grains cm$^{-2}$ y$^{-1}$ found in a sample at Klotjärnen just after the occurrence of the first *Picea* bud scale (**?**). However, these modern thresholds estimated here are likely to depend on the abundance of the parent tree in the larger region rather than properties of the pollen types. A larger threshold would be expected for *Corylus* compared to *Fagus*, based on the fall speed of pollen. However, *Corylus* is not very abundant near its northern limit in these more continental areas, while *Fagus* often dominates forests near its presumed limits. Similarly, the occurrences of taxa outside the mapped natural distribution can bias this estimate.

~~Linking the fossil to modern PAR values facilitates interpretation of the fossil record of individual sites. Unfortunately, the details cannot be discussed here. However, the central Swedish sites Holtjärnen and Klotjärnen provide excellent examples. These sites are situated north of the modern distribution of~~ Characteristic radius is a useful measure showing pollen transport predicted by pollen dispersal models. It represents theoretical proportion of pollen loading at different distances from the source plants (**?**) and thus can be easily compared with our empirical values. PAR of ~~*Tilia*~~*Carpinus*, *Corylus*, ~~*Quercus*~~*Fraxinus* and ~~near the limit of *Alnus glutinosa*. The fossil PAR values are higher for these taxa than those found in pollen traps at or near these lakes (**??**). Modern reference values for the PARs of these taxa can be found in northern Poland and Estonia.Moreover, this analogue matching indicates that 3000 years ago the PAR values for *Quercus* at Holtjärnen were sufficiently high to indicate the occurrence of small populations near the lake~~*Quercus* at 200 km from the distribution area represent 5-10 % of the median PAR within the distribution area (Fig. 4). Considering the range of standard terminal velocities of those taxa and Gaussian Plume model with wind speed 3 m.s$^{-1}$ (**?**), we get 2-13% of pollen loading at the characteristic radius 200 km (Table S6).

## 4.3 Analogues for vegetation reconstruction

While the dataset of modern PARs presented here provides analogues for the selection of fossil sites, reliable fossil PAR records were not available from the trapping locations in the UK. Moreover, no sites with fossil PAR from western Europe or from low elevations in the Alps are included in this comparison. Consequently, the large quantities of *Corylus* pollen deposited in many west European sites during the early Holocene was not considered. There are limited modern analogues forthe highest early Holocene values of around 7000 *Corylus* grains cm$^{-2}$ y$^{-1}$ at Prášilské. Average Early Holocene *Corylus* PAR at Soppensee in northern Switzerland are 12000 grains cm$^{-2}$ y$^{-1}$ (**?**) and at Meerfelder Maar (**?**) in western Germany 18000 grains cm$^{-2}$ y$^{-1}$. Judging from pollen percentages even higher Early Holocene values should be found in more oceanic situations and the *Corylus* PAR at Hockham Mere in eastern England may be as high as 40000 grains cm$^{-2}$ y$^{-1}$ for the early Holocene (**?**). Modern values in pollen traps from Wales at around 2000 grains cm$^{-2}$ y$^{-1}$ are far below these early Holocene figures and it is likely that modern analogues of sites with high *Corylus* PARs no longer exist in Europe.

Conversely, the high modern PAR values for *Pinus* and *Betula* from Poland and Latvia are not found in the fossil examples. *Pinus* PAR values around 30000 grains cm$^{-2}$ y$^{-1}$ were also obtained from $^{210}$ Pb dated modern lake sediment samples in north eastern Brandenburg (**?**). This study evaluated the PARs for the years 1993 and 2009. The increase in *Pinus* PAR values between the first and the second sampling period corresponded with an increase in the amount of standing pine volume in the region. Forestry practices aimed at increasing yield could account for the high *Pinus* values. *Pinus* was extensively planted after the 1950s, even on soils where trees with a lower pollen production would have grown naturally. The fertilization due to increased nitrogen deposition, as well as increased atmospheric carbon dioxide, increase the pollen production not only of *Pinus*. A carbon dioxide enrichment experiment of 19-year old *Pinus taeda* resulted in a twofold probability of reproductive maturity after 3 years (**?**). The continued experiment also showed that carbon dioxide fertilization increased the number of pollen cones and therefore pollen grains produced per tree (**?**).

In the case of *Pinus*, the modern dataset includes trap data from Cyprus, where *Pinus brutia* dominates at 1600 m a.s.l., resulting in even higher *Pinus* PAR values compared to those found in the Polish and Baltic regions. The highest *Betula* values come mainly from Russia, where values frequently exceed 30000 grains cm$^{-2}$ y$^{-1}$. We previously discussed such high fossil PAR values for *Corylus*, which is assumed to produce a similar amount of pollen. However, fossil *Betula* PARs in the examples considered here are consistently below 6000 grains cm$^{-2}$ y$^{-1}$ and published early Holocene values rarely exceed 6000 grains cm$^{-2}$ y$^{-1}$ (but see **?**). Pollen diagrams from the forest steppe ecotone in European Russia are often characterized by high *Betula* percentage (**??**). However, there are no suitable diagrams with reliable PAR estimates from that region. It is thus difficult to judge whether high modern trap values are associated with recent land-use change or are characteristic of eastern European forests.

The comparison of regional PARs between traps and fossil estimates indicates higher fossil PAR of *Picea*, *Fagus* and *Abies* in middle altitudes of the temperate zone (Fig. 5), which, in the case of *Abies*, represents the Europe-wide decline in *A. alba* (**?**). *Picea* and *Fagus* dominate central European forests today and *Picea* is planted much beyond its natural range. However, both trees start flowering rather late in their lives and harvesting the trees at a young age may contribute to lower modern PAR values. Fossil and modern PARs for these trees in Šumava are similar, while only the highest trap values match the Holocene high values. On average there are lower modern PAR values; this may be explained by a lowering of the treeline over the last millennia. This interpretation agrees with REVEALS reconstructions for this region, indicating a decline in the cover abundance of *Picea* and *Fagus* (**??**). Within a 60 km radius of the fossil sites *Picea* decreased in abundance from 70% during the Middle Holocene to 43%, compared to modern abundance. *Fagus* and *Abies* declined from Late Holocene values of 22% and 3% to currently 20% to 1% respectively (**?**). The abundance of *Abies* in the Roztocze region (SE Poland; Fig. 6) provides a good analogue for the past abundance of the tree in Šumava with maximum PAR of 1000-3000 grains cm$^{-2}$ y$^{-1}$. *Abies* disappeared from the Czech Republic during the Mediaeval Age due to forest management methods (**?**), which were not practiced in SE Poland.

Linking the fossil to modern PAR values facilitates interpretation of the fossil record of individual sites. Unfortunately, the details cannot be discussed here. However, the central Swedish sites Holtjärnen and Klotjärnen provide excellent examples. These sites are situated north of the modern distribution of *Tilia*, *Corylus*, *Quercus* and near the limit of *Alnus glutinosa*. The fossil PAR values are higher for these taxa than those found in pollen traps at or near these lakes (??). Modern reference values for the PARs of these taxa can be found in northern Poland and Estonia. Moreover, this analogue matching indicates that 3000 years ago the PAR values for *Quercus* at Holtjärnen were sufficiently high to indicate the occurrence of small populations near the lake. Those taxa at both sites also nicely illustrate the potential use of LDT for the interpretation of the fossil record. Fossil PAR values of *Tilia*, *Corylus* and *Quercus* in boreal region result above LDT threshold, whereas all trap PAR values are below (Fig. 5).

### 4.4 Limitations and problems

There are significant differences between the accumulation of pollen in traps and on peatlands and lakes (**??**). Differences in pollen trap design and placement in the landscape will influence the values. Trap values are also affected by modern processes that have no impact on the fossil signal. ~~The~~ These biases appear minor as indicated by the large consistency of the data collected in the PMP database~~indicates that these effects may be small.~~. Also the comparison of values over this large environmental gradient result in the signal being stringer than the noise. Nevertheless, some traps or individual years have unusual values and were removed from the comparison (Table S1). Despite this, the uncertainty of fossil PAR values is much greater than pollen traps, which is primarily due to the added uncertainty coming from sampling a sediment core, combined with the uncertainty of the age model (**?**). PAR from lake sediments has additional biases due to differential sedimentation of pollen grains in lakes (**?**), sediment ~~redeposition~~re-deposition, focussing and catchment erosion (**??**). Although we carefully selected the best available fossil sites~~PARs~~, PARs especially from lake Suminko and Rõuge Tõugjärv may be biased by lake internal processes and the addition of stream borne pollen respectively. Nevertheless, their fossil PAR estimates are in the range of values found in pollen traps. Where detailed knowledge on the sedimentation process is available the bias of sediment focussing may be reduced as in the example of Hockham Mere cited above (see also **??**). Peatlands may thus seem the better choice for obtaining fossil PAR, which may be the case in northern Scandinavia (**??**), but frequent changes in the rate of peat growth lead to difficulties assessing the time represented in individual samples at many sites.

The problem of traps ~~being contaminated by~~ collecting high amounts of herbaceous pollen brought by insects and small animals was ~~mentioned~~ discussed in the method section and for this reason only Poaceae and Cyperaceae ~~pollen were included in this comparison~~are the only herbs selected for our analyses. However, pollen from these two families is also often overrepresented in the pollen traps (**?**), as the plants may overhang the trap opening and their pollen may fall directly into the trap. Reduced PARs in the trap may be caused by overgrowth of the vegetation or leaves temporally blocking the opening, while proximity to the forest edge would increase values compared to large open peatlands or lakes. These effects have not been systematically evaluated so far.

Detailed comparisons of vegetation data to PARs hold potential for a better understanding of the spatial representation and processes shaping the pollen signal (**?**) and allow estimates of absolute pollen productivity (**?**) or test pollen dispersal models. However, ~~forest~~ for this continental scale dataset available vegetation data have limited precision. Forest inventory data with the detail essential for this type of study is not available for all traps. The forest cover data presented here ~~presented here~~ has a resolution of 1 km$^2$, ~~so that it was insufficiently detailed to consider the distribution~~ which is insufficient as the abundance of trees within hundreds of meters of the traps is important. Moreover, without information on standing volume or age structure, the percentage cover used here is a crude measure of the vegetation producing the pollen~~and it is almost surprising that this variable has some explanatory power~~. Forestry practices like harvesting trees that start flowering at a later age (e.g. *Picea* 30-40 years) reduce the number of trees producing pollen (**?**) and bias the search for modern analogues. Also, the available mapped distribution limits of trees have large uncertainties precluding more detailed assessments of the quantity of long distance transported pollen using this continental dataset.

## 5 Conclusions

Comparison of the mean annual PAR from traps and fossil sites showed similar ranges for *Abies*, *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Fagus*, *Fraxinus*, *Picea*, *Pinus, Quercus* and *Tilia* at the continental scale. ~~It~~ This indicates that there are no ~~significant~~ major biases hampering the application of the PMP Database data as a modern reference to interpret the fossil record. The dataset clearly shows that climate parameters that correlate with latitude determine pollen productivity. The effect of regional forest cover is discernible~~, while cover within hundreds of meters is likely to be orders of magnitude stronger, but could not be assessed here~~. Minimum values suggest that an 80% forest cover within 10 km of the trap results in PARs above 3200 tree pollen grains cm$^{-2}$ year$^{-1}$.

Assessment of long-distance transport indicates that ~~threshold PAR values range from 1 to~~ values below 30 grains cm$^{-2}$ y$^{-1}$ for *Carpinus*, *Corylus*, *Fagus*, *Fraxinus*, *Picea*, *Quercus* and *Tilia* may originate from beyond 200 km of a sampling site~~and values up to~~. This number of 30 grains cm$^{-2}$ y$^{-1}$ ~~should therefore be regarded as potentially of~~ may therefore be used as a general threshold indicating long distance origin of pollen. The application of these threshold values holds potential to refine and adjust reconstructions of tree distributions.

~~Comparison of fossil and trap datasets in individual regions with LDT threshold provided evidence for reconstruction of changes in distribution of selected species during the Holocene. Matching the target periods at individual fossil sites to trap records revealed the nearest analogous populations and provided insights into past vegetation. History of species in particular regions viewed by both ways of vegetation reconstruction show certain similarities. When ranges of trap and fossil PARs are similar within the region, distances to the nearest analogous traps are short.~~

*Code availability.* Primary trap and fossil pollen data are availible in Neotoma Palaeoecology Database https://www.neotomadb.org/. Analysis are based on the WorldClim 2 dataset of Fick and Hijmans (2017), Chorological data for the main European woody species, version 2 by Caudullo et al. (2018), European atlas of forest tree species, 2016th ed. by San-Miguel-Ayanz et al. (2016), which are availible online: http://worldclim.org/version2 , https://data.mendeley.com/datasets/hr5h2hcgg4/2 and http://www.euforgen.org/, respectively. Forest Map of Europe of Kempeneers et al. (2012) is availible on request on authors.
Code for analysis, derived data and code for figures are available in the https://github.com/vojtechabraham/PMPdatabase.

*Author contributions.* Design of the analysis: TG and VA. Data analysis: VA, TG and MW. Manuscript draft: VA and TG. Manuscript comments: HP, SH, CEJ, HSS, HS, SV, SH, MN, MW, SP, ST, TK, AMN and JŚM. Data collection: all authors except VA and MW.

*Competing interests.* The authors declare no competing interests.

*Acknowledgements.* We thank to all collaborators of pollen trapping, especially to those who contributed to the PMP dataset: Pim van der Knaap, Jacqueline van Leeuwen, Karl-Dag Vorren, Heather Tinsley, Lena Barnekow, Małgorzata Latałowa, Enikő Magyari, Larissa Savelieva, Elena Pavlova, Heidi Hyyppä, Mari Kuoppamaa and Gunhild Rosqvist.