

Workflow paper about climate reconstructions using a multi-method approach based on two biomisation processes in Europe since the LGM (Reference fossil site)
No best fundamental methods, only pros and cons!

Pollen data are commonly used as paleoclimate proxies and offer the most spatially extensive reconstructions in unglaciated terrestrial regions for the Quaternary period. However, significant scientific challenges → such as the lack of fossil sites (particularly across spatial and altitudinal gradients) and the need for comparison of output methods ([Cao et al., 2019](#) ; [Song et al., 2021](#) ; [Brierley et al., 2020](#)). Many modern datasets under-sample specific spatial areas, to extract the nearest climate space relative to past vegetations, assuming a constant relationship pollen-climate over time (*e.g.*, [Jackson and Williams, 2004](#)).

The calibration of local biomes remains a topic of debate. Additionally, no publicly accessible programming language for the community, meaning that the local PFT-Biome relationship (as defined by [Prentice et al., 1998](#)) has not yet been tested on large datasets in Europe. Despite many climate reconstructions using multiple variables from modern datasets, few papers evaluate pollen methods, for testing and comprising between different methods and paleoclimatic proxies ([Birks et al., 2010](#) ; [Blois et al., 2013](#)). Hot, cool and cold biome distinctions are rarely described in literature, whereas the biome classes show important effects on climate reconstruction.

Transfer functions and climate results depend on the abundances and spatial distributions within the calibration dataset, but also the taxonomical resolution which complexify the pollen-climate reconstructions/responses ([Kühl et al., 2002](#) ; [Chevalier et al., 2019](#)). MAT and WAPLs methods describe just the 'best estimate conditions' to find unique or pollen assemblages along a climate gradient. To mitigate geographical constraints and take into account for the taxonomic diversity of modern pollen, exhaustive/comprehensive modern calibration datasets are essential.

All previous biomization studies consider different hypothesis, by (i) selecting dominant biomes with the highly affinity scores to fit individual climate datasets from each biome and (ii) taking precedence the least PFT-rich biome when the affinity values for more biomes are identical ([Harrison et al., 2010, 2017](#) ; [Dallmeyer et al., 2019](#)). We propose to test these assumptions to isolate its climate effects.

(Introduction), a general overview of climate quantifications based on pollen assemblages explain how these methods work and the fundamental assumptions involved. Due to the unimodal and Gaussian climate responses of vegetation, climate reconstructions from the genus to family pollen level are often uncertain.

The probabilistic approach (likelihood distributions) from CREST (Climate REconstruction SoftWare, [Chevalier et al., 2020](#)) offers one promising avenue to (i) reduce the heterogeneity influence of modern climate, (ii) focus on data-data comparison and (iii) potentially implement data in Earth System Models paleo-simulations (ESMs, [Dallmeyer et al., 2022](#)). However, the pollen-climate proxy relationship established by the CREST method has only

been applied to tropical sites thus far ([Chevalier et al., 2014](#) ; [Chevalier et al., 2019](#)). We implement the biomization method in R to CREST approach in Europe.

The present methodological development evaluates the reliability of traditional approaches and identify their limitations. Comparisons between data models remain challenging and are particularly difficult to explain in the context of LGM cooling.

I- Study sites, 14C (Bacon R) chronologies and vegetation dynamics – taxonomic resolution (record quality)

This section contains a brief description of data and codes in free access at https://github.com/GabrielFens/Climate_reconstructions_since_the_LGM_based_on_pollen_assemblages.git

Dataset sources – Code and data availability

The pollen database used in this study derives from the European Pollen database (EMPD; <https://empd2.github.io> ; www.europeanpollendatabase.net/, last access: January 21, 2024). The total number of sites is 8746 that represent modern surface samples. The quaternary fossil datasets including five sites, are in open-access from ACER 1.0 database (<https://doi.org/10.1594/PANGAEA.870867>, last access: September 22, 2022; [Sánchez Goñi et al., 2017](#)) and the PANGAEA Event. We expanded the biomization tools (Taxonomic harmonization, the lists of taxa-PFTs and PFTs-megabiome assignments and PFTs-Biome dataset) from [Binney et al., 2017](#) ; (<https://hdl.handle.net/21.11116/0000-0001-B800-F>, last access: May 16, 2024; [Dallmeyer et al., 2019](#) ; [Cao and Tian, 2021](#) ; <https://www.r-project.org/>, last access: May 10, 2023; R Core Team, 2020).

Chronological control

Dating information (i.e., radiocarbon chronologies, 14C dating) was re-calibrated to construct new Bayesian age-depth models from conventional ages of each fossil sites, using the CLAM v. 1.2 (BACON routine in R, [Blaauw, 2010](#) ; [calibration curve of Reimer et al., 2021](#)). For extant age models, we assume a linear variation of accumulation rate through time. One figure shows the temporal distributions of pollen sequences. Most of the pollen records have more than 8 dates. Only the Fùramoos site derived from 40-Ar ages. All used ages in reconstruction are calibrated years before present (cal BP).

II- Methods of two biomization processes: PFTs Assignment and megabiome associations

PFTs methods

We employed the biomization method as an algorithm following the square root procedure to assign pollen taxa to plant functional type (PFTs) ([Prentice et al., 1996](#) ; [Prentice and Webb, 1998](#)). The biomization affinity scores was deduced taking into account the PFT vectorizations between steppic and desertic conditions. Assignment of pollen i.e., biome-PFT-taxa relationship is adapted to the descriptions of bioclimatic zone, life form, and leaf

phenology ([Williams et al., 2004](#)) and European pollen records. We defined 25 and 30 PFTs in two classifications from sets of taxa allows the floristic and functional heterogeneity within key biomes. And, we converted pollen taxa into megabiomes using the previous procedure of [Prentice et al., 1996](#) and standardized categories of [Harrison et al., 2010](#) ; [Dallmayer et al., 2019](#).

The main aim of PFTs-megabiomes is to follow signals of key land-cover changes (e.g., wetlands extent, conifer dominance) without their being masked by classification into biomes. All affinity scores from pollen taxa were calculated only for pollen taxa percentages greater than 0.5% abundance to increase signal-over-noises (e.g., [Chen et al., 2010](#)). Pinus abundance is corrected by a factor 5 to limit their prevalence in temperate open-forests and then compensate their high pollen productivity. When affinity scores are equal, Biome with the biome with the fewest PFTs is selected. The megabiomes “dotmaps” in each site (assigned age) is mapped according to a colour-coded key, i.e., key time-slices in Europe-only maps at a larger scale.

- III- *General principles of the three pollen methods:*** calibration, regression and pdf techniques using EMPD2 2020 2021
- IV- *Results:*** the spatial distributions of biomes, climate spaces to reconstruct and climate reconstructions (TANN, PANN and GA, may be MTCM and MTWM)
- V- *Discussions:*** multi-sites, multi-samples discontinuities, the method evaluations – signal-to-noise ratio, altitudinal gradients in vegetation responses
- VI- *Conclusions, perspectives***