Report of the 5th NECLIME workshop on Climate

February, 21-22, 2022

The 5th workshop of the NECLIME working group on climate took place online and was organized by Andrea Kern and Thomas Kenji Akabane. The workshop was well-received with 37 participants including a great number of young scientists. The main purpose of this workshop was to discuss the potential and limitation of new methods for paleoclimate reconstructions for Quaternary and Neogene datasets on a global scale, especially ones applying probability density functions (pdf). Additionally, comparisons with the Coexistence Approach, non-plant proxy data as well as different pdf methods were discussed. A second scientific focus lay on the impact on CO_2 on plant-based paleoclimate reconstructions and how this can be accessed through time.

To make this workshop accessible for everyone interested in this topic, each day started with a short introduction on the methodology and potential issues associated to the used databases. Later, a series of presentations gave insights into the current results and problems of the respective techniques. The first day of the workshop showed several contributions applying an intercomparison among various climate reconstruction techniques to evaluate the applicability, precision and limitation of pdf-based climate reconstruction estimates. These included varying timescales in areas which are commonly problematic in paleoclimate reconstructions such as the tropics and high latitudes (climatic extremes), and reconstructions on a global scale as well as high-resolution records.

On the second day, the focus shifted towards the underlying datasets of the pdf climate reconstructions, particularly on the Global Biodiversity Information Facility database (GBIF). Despite containing a great and valuable collection of georeferenced data and information, certain caveats associated to erroneous datapoint entries must be considered, as these might pose a problem to the performed paleoclimate reconstructions. In this context, it was addressed how the different kinds of observation datapoints available in GBIF can represent sources of potential errors and a potential solution such as cleaning of datasets using the R package “Coordinate Cleaner”. Afterwards, an additional scientific presentation gave an overview of previous possibilities to overcome the influence of atmospheric CO_2 in paleoclimate reconstructions, particularly in low CO_2 scenarios such as the Last Glacial Maximum.

Using CRACLE and CREST for paleoclimate reconstructions – preliminary considerations:

Both CRACLE (Climate reconstruction analysis using coexistence likelihood estimation – Habert and Nixon, 2015; Habert and Baryiames, 2019) and CREST (Climate Reconstruction Software – Chevalier et al. 2014; Chevalier, 2021) are new methods using probability density functions to estimate paleoclimate with a previous focus on Quaternary records. These methods differ in in data acquisition and statistical processing, therefore results are not identical. Although we mainly considered the application on micro paleobotanical data, both methods can potentially be used with all organism groups with available georeferenced data. CRACLE and CREST are set to perform their standard analysis by collecting distribution data from GBIF (https://www.gbif.org/) and intersect this point-data information with the gridded interpolated climate maps provided by worldclim (Fick and Hijmans, 2017; https://worldclim.org/). While the CREST limits the user to GBIF and worldclim in their online R package, CRACLE offers more possibilities to freely collect and clean own datasets before running the pdf and statistical analysis. This allows the user to address concerns in the used datasets of plant distribution (already previously discussed within the NECLIME community) and also opt, if necessary, for other available databases e.g. BIEN (Botanical Information and Ecology Network).

In this workshop, we discussed the application of CRACLE and CREST through time and on a global scale in Miocene, Pliocene and Pleistocene records. While Pleistocene and Pliocene datasets mainly considered the applicability and limitations of each method, the Miocene plant assemblages allowed comparisons between pdf-methods and the Coexistence Approach. Both CRACLE and CREST can potentially perform for all the timescales considered, although a thorough consideration of data and results is recommended. Due to the focus on plant climate distribution optima, the methods often failed to rightly reconstruct extremes of the climatic parameters, such as very cold in the high latitudes (e.g. Arctic) or very wet conditions in the tropics (e.g. Amazon). In continuous high-resolution datasets, both CRACLE and CREST have the potential to reveal expected long- and short-term climatic trends associated to known events. The alternative use of pdf mean values instead of plant optima was suggested and has proven a good method on regional geographical scale. In high-resolution records, the use of CREST set for percentage and/or normalized pollen counts resulted clearly in stronger variations through time than presence/absence only (although the precision and accuracy has yet to be assessed). While in certain regions, the use of non-parametric (kernel density) distribution estimates in CRACLE have resulted in more precise paleoclimate data, in other areas results have proven less accurate and less sensitive to changes. This is potentially caused by a higher sampling bias and uneven distribution and climate data. The comparison with the Coexistence approach for Miocene estimations revealed that in certain regions climate reconstructions resulted in lower temperatures using pdf-based methods. In some records, this was caused by the low number of thermophilous taxa in comparison with temperate plants, which caused a shift towards the more abundant cooler taxa instead of calculating a mean value. In comparisons, mean annual precipitation ranges remained comparable or resulted in wetter estimates. Overall, an overlap of confidence intervals could be achieved.

One factor strongly controlling the outcome of each climate reconstruction is the pre-selected geographical range of the analysis. Limiting the geographical range of an analysis defines the outcome of each study, therefore this needs to be carefully considered before running a pdf-based paleoclimate reconstruction. While in Late Quaternary studies and geographically more isolated regions a geographic limit of the study area might be beneficial, it also might introduce a bias in deeper time studies due to an insufficient climate range for the analysis. Using no geographic limitation ("global distribution") also might introduce a bias in certain studies if some taxa have a large number of species particularly abundant in Europe or the USA (high sampling density). In this context, it is recommended a thorough evaluation of the taxa distribution and potential exclusion of cosmopolitan large taxonomic units such as wide families and species rich genera. Plants with a very wide distribution or such where distribution is not primarily climatically controlled (e.g. aquatics) might also be excluded from the analysis. The inclusion of widely distributed large groups e.g. families can often result in higher uncertainties in the reconstruction, suggesting a filtering of the taxa list is advisable prior to the analysis. Although CREST suggests a higher diversity does not cause higher precision in the reconstruction, a certain minimum number of taxa is highly recommended on the basis of our current experience. In addition, for CRACLE, higher precisions in results are accomplished with a higher number of taxa.

The underlying data of CRACLE and CREST were further considered regarding potential errors. Despite the magnitude of data available on databases like GBIF, errors are always evident. Within the GBIF database; known issues are missing or imprecise coordinates, reversed coordinate entries, the occurrence outside of the organism natural environment (invasive species, anthropogenic influence in e.g. botanical gardens, historical agricultural products), while further problems can occur with missing data in poorly sampled regions and/or taxa. Although CREST and CRACLE address these issues to some extent, a careful consideration for all taxa is advised at the current status. The discussed R package Coordinate Cleaner allows addressing wrong and imprecise data entries and therefore represents a strong tool to help to reduce noise and avoid unnecessary bias in the paleoclimate reconstruction. Furthermore, remaining problematic taxa should be excluded and explained likewise to the Coexistence Approach.

Useful considerations:

R packages

- rGBIF https://www.gbif.org/tool/81747/rgbif
  https://cran.r-project.org/web/packages/rgbif/index.html
- BIEN R
  https://cran.r-project.org/web/packages/BIEN/index.html
- Coordinate Cleaner
  https://www.rdocumentation.org/packages/CoordinateCleaner/versions/2.0-20
- CRACLE
  https://github.com/rsh249/cRacle
- Crest
  https://github.com/mchevalier2/crestr

Important to report/record before publication

- R version
- Version of the R packages used e.g. CRACLE or/and CREST version

- DOI of the GBIF download (be careful to have the deletion date of your download extended to make sure you have time to publish and have the GBIF recognize the publication before this is deleted)
- All parameters used in the functions e.g. extraction() and dens_obj() (CRACLE), area considered for the analyses (e.g. a continent, a drainage basin, worldwide extension) (CREST and CRACLE).
- Climate data version and resolution
- If MCR, if it is weighted or unweighted
- If taxa were excluded and why

References

Chevalier, M., Cheddadi, R., Chase, B.M., 2014. CREST (Climate Reconstruction SofTware): a probability density function (PDF)-based quantitative climate reconstruction method. Clim. Past 10, 2081–2098. https://doi.org/10.5194/cp-10-2081-2014.

Chevalier, M., 2019. Enabling possibilities to quantify past climate from fossil assemblages at a global scale. Glob. Planet. Change 175, 27–35. https://doi.org/10.1016/j.gloplacha.2019.01.016

Fletcher, T., Feng, R., Telka, A.M., Matthews, J.V., Ballantyne, A., 2017. Floral dissimilarity and the influence of climate in the pliocene high arctic: biotic and abiotic influences on five sites on the canadian arctic archipelago. Front. Ecol. Evol. 5, 1–13. https://doi.org/10.3389/fevo.2017.00019.

Fletcher, T.L., Csank, A.Z., Ballantyne, A.P., 2019. Identifying bias in cold season temperature reconstructions by beetle mutual climatic range methods in the Pliocene Canadian High Arctic. Palaeogeogr. Palaeoclimatol. Palaeoecol. 514, 672–676. https://doi.org/10.1016/j.palaeo.2018.11.025.

Gibson, M., McCoy, J., O'Keefe, J., Nuñez Otaño, N., Warny, S., Pound, M.. 2022. Reconstructing Terrestrial Paleoclimates: A Comparison of the Co-Existence Approach, Bayesian and Probability Reconstruction Techniques Using the UK Neogene. Paleoceanography and Paleoclimatology. 37. 10.1029/2021PA004358.

Harbert, R.S., Baryiames, A.A., 2020. cRacle: R tools for estimating climate from vegetation. Appl. Plant Sci. 8, 641183. https://doi.org/10.1002/aps3.11322.

Harbert, R.S., Nixon, K.C., 2015. Climate reconstruction analysis using coexistence likelihood estimation (CRACLE): a method for the estimation of climate using vegetation. Am. J. Bot. 102, 1277–1289. https://doi.org/10.3732/ajb.1400500.

Schedule of the workshop

Monday, February, 21st, 2022

3.00 – 3.20 pm	Welcome and Introduction
3:20 – 3.35 pm	**Thomas K. Akabane:** CREST and CRACLE for the late Quaternary of Central America
3.35 – 3.50 pm	**Tamara Fletscher**: A Beaver Pond frozen in time: Multi-proxy analysis in the Pliocene High Arctic
3.50 – 4.00 pm	Break
4.00 – 4.15 pm	**Martha Gibson**: CREST and CRACLE for Middle Miocene temperature and precipitation reconstructions
4.15 – 4.30 pm	**Andrea K. Kern**: Comparison of Miocene climate estimates from Austria using the Coexistence Approach and CREST: first results
4.30 – 5.00 pm	Discussion

Tuesday, February, 22nd 2022

3.00 – 3.10 pm	Introduction and recap of day 1
3.10 – 3.25 pm	**Angela A. Bruch**: Determining climate requirements of plant taxa – a comparison of GBIF and Palaeoflora datasets
3.25 – 3.40 pm	**Krister Smith**: GBIF, R and CoordinateCleaner – pitfalls and opportunities
3.40 – 4.10 pm	Discussion

4.10 – 4.25	Break
4.25 – 4.35	**Nils Weitzel**: Accounting for changing CO_2 concentrations in pollen-based climate reconstructions
4.35 – 5.00 om	Final discussion