A bit of Theory

The CREST method [1] is a Bayesian approach that combines presence-only occurrence data and modern climatologies to estimate the conditional response of a given taxon to a variable of interest. Taking the form of probability density functions (*pdfs*), these links are fitted in one or two steps based on the nature of the proxy being studied. In simple cases, where fossils can be identified at species level (*e.g.* foraminifers, plant macrofossils), the pdfs are defined by unimodal and parametric functions (e.g. normal or log-normal distributions depending of the nature of the studied variable, see [1] for a more detailed discussion).

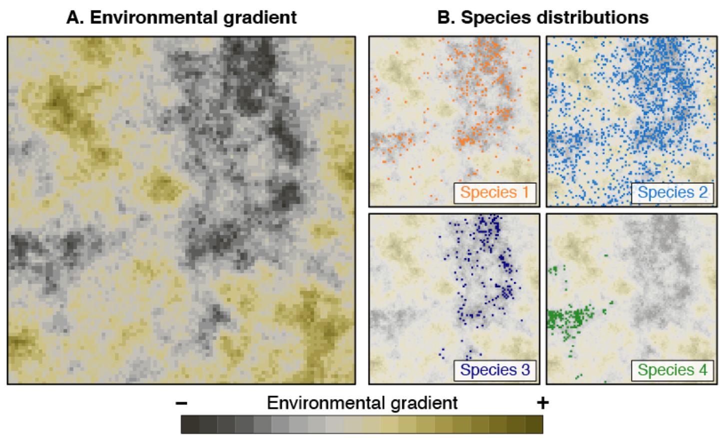

Figure 1: (left) Climate variable to reconstruct (*e.g.* mean annual temperature)

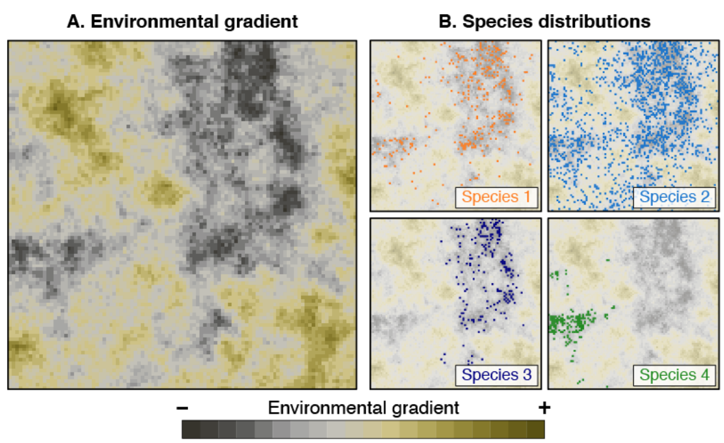

Figure 2: Caption for the picture.

The parameters (e.g. a mean and a standard deviation in the case of a normal or log-normal distribution) describing these distributions are estimated from the ensemble of climate values corresponding to the presence records, each being weighted as an inverse function of its abundance in the study area. This correction is needed to remove the influence of the heterogeneously distributed modern climate space and ensure that the optimum exhibited by the pdf truly reflects the climatic preference of the species, rather than the modern abundance of a given climate value [2-3]

When the fossils cannot be identified at the species level (e.g. pollen grains identified at the genus or family level), two steps are necessary to create the taxon-climate probabilistic link. First, following the aforementioned process, a parametric pdf is created for each of the species producing the same type of pollen grains, and then, these pdfs are grouped together to create a higher order pdf representing the pollen type, with each species being weighted as a function of the extent of its distribution. One assumption of the model is that in the absence of independent evidence, species with larger distributions are considered more likely to have produced the pollen grain observed. No additional assumptions are made concerning the shapes of these pollen pdfs, allowing them to be multimodal if different species/groupings exhibit different climate requirements. It is worth noting that because CREST uses parametric functions to define the species pdfs, the process can be used with incomplete distribution data. Recent results suggest that robust pdfs can still be obtained from truncated geographical distributions, provided that the full range of the climatic tolerance of the species is well covered in the climate space [4-7].

Finally, to estimate past climate parameters, the pdfs for each fossil taxon identified in a sample are multiplied together, each with a weight that is derived from the observed percentages. Since relative abundances, different production rates and taphonomy are important factors influencing the fossil assemblage observed, direct percentages cannot be used directly, and they need to be transformed to minimize the effect of all these factors. In CREST, it is done independently for each taxon by normalizing the raw percentages by the average of the percentages observed in a sequence (zeros excluded). A value higher (lower) than one suggests that the climate at the time of deposition was more (less) favourable (i.e. closer to the taxon's climate optimum) than the average climate in which the taxon has been observed during the studied period. The multiplication of pdfs results in a posterior distribution of probabilities along the climate gradient, from which climate estimates and uncertainties can be derived. More details about the method can be obtained from the original publication [1].

The probabilistic nature of CREST also provides the unique opportunity to obtain reconstructions in the form of posterior distributions of probabilities that describe the likelihood of all the climate values along a studied climate gradient, and not just a single 'best estimate' associated with a standard error.

```
library(crestr)
```