Today’s suggestion

09:45-10:00	Café de bienvenue
10:00-12:00	Session 5: Introduction to CREST/r
12:00-13:00	Lunch
13:00-14:30	Session 6: Guided visit of *crestr*
14:30-14:45	Break
14:45-16:00	Session 7: Hands-on with a basic example
16:00-16:30	Break
16:30-18:00	Cocktail de réseautage

Sessions 05-06-07

CRESTR an R package to perform Probabilistics climate reconstructions From palaeoecological datasets

Manuel Chevalier

mchevali@uni-bonn.de

PALMOD GERMAN CLIMATE MODELING INITIATIVE

UNIVERSITÄT BONN

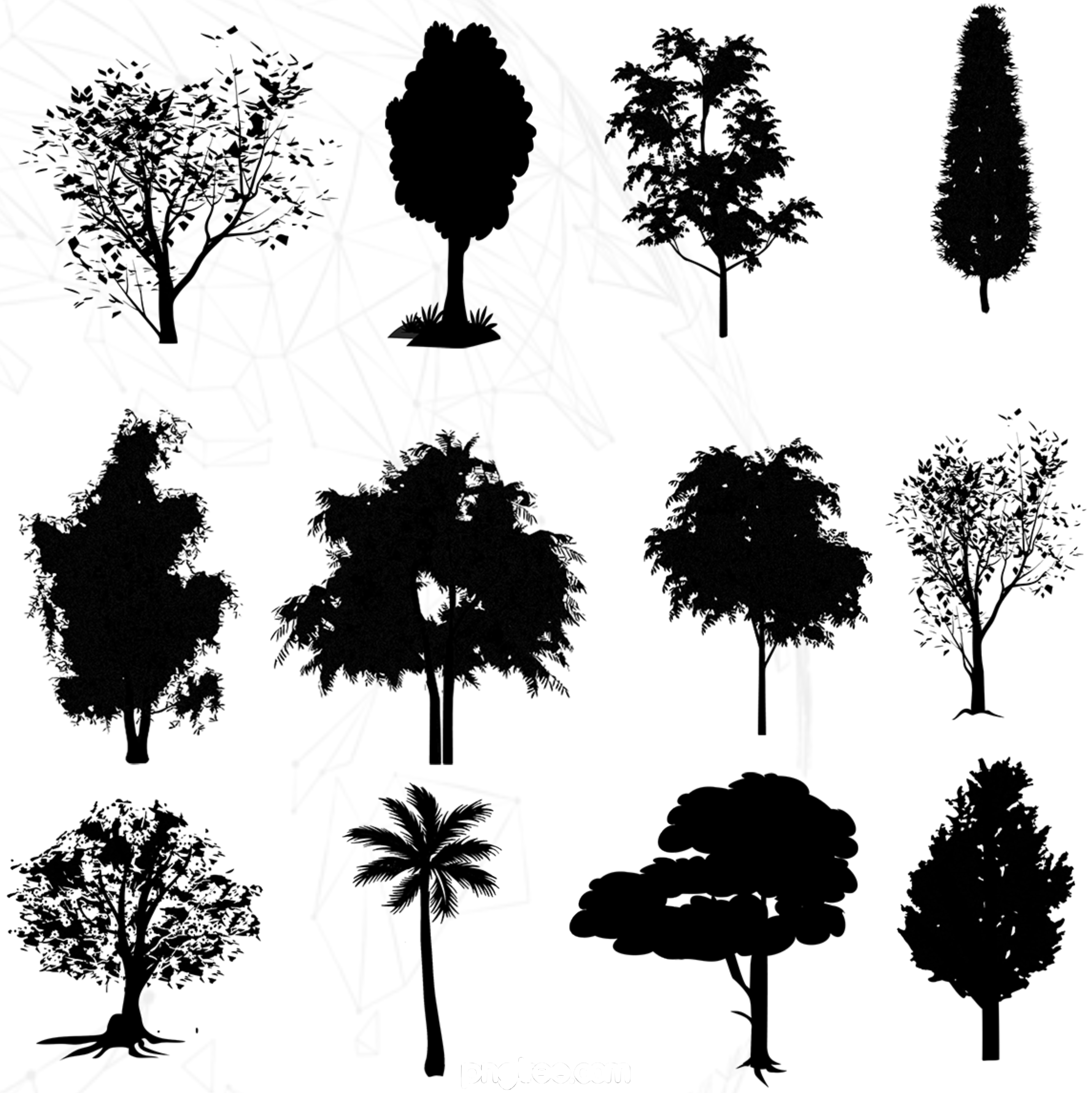
3. TABLE OF CONTENT

4. TABLE OF CONTENT

5. TWO WAYS OF RECONSTRUCTING CLIMATE FROM POLLEN

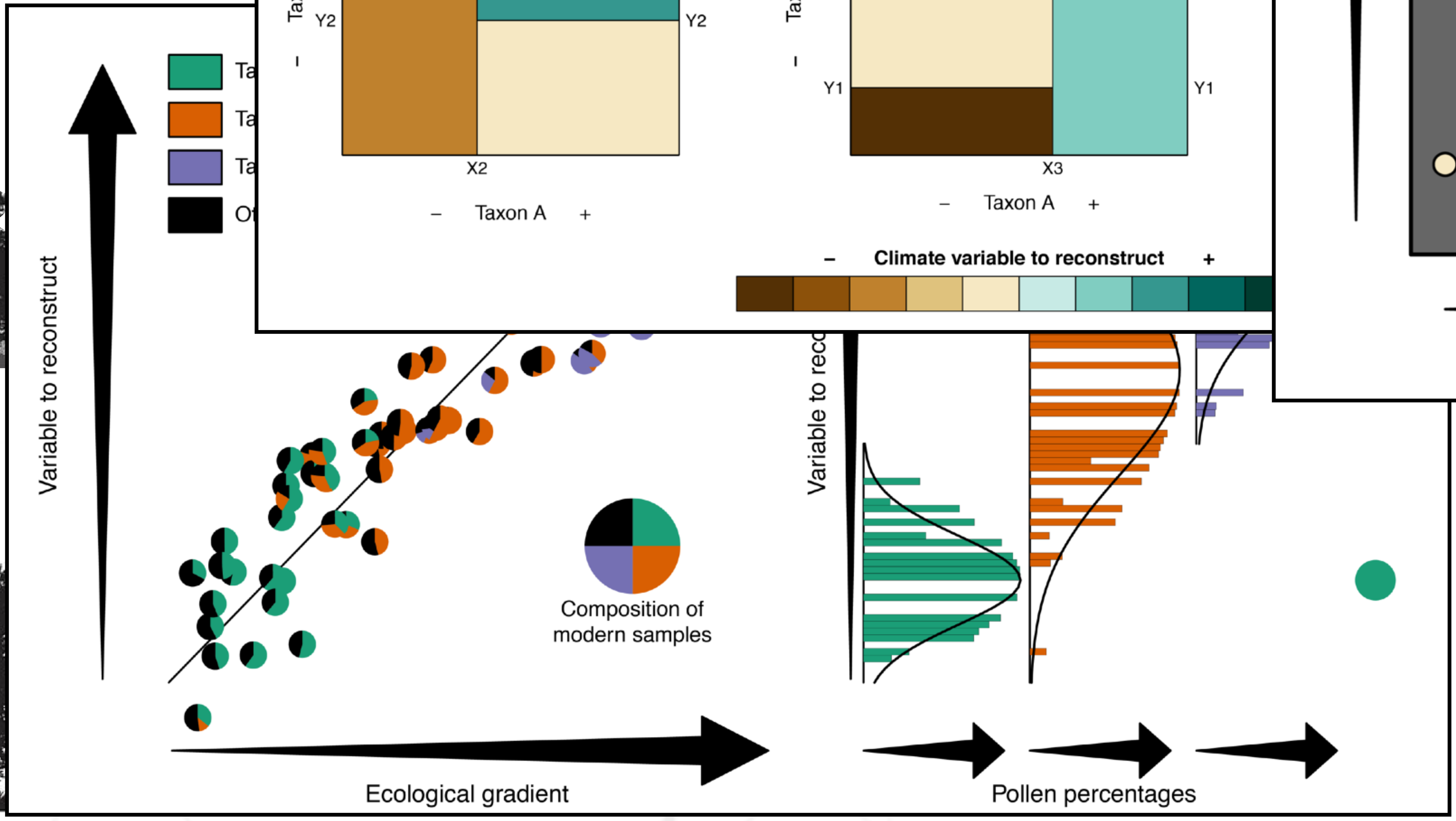
The "composition" approach = Think "Ecosystem"

The "indicator species" approach = Think "Independent taxa"

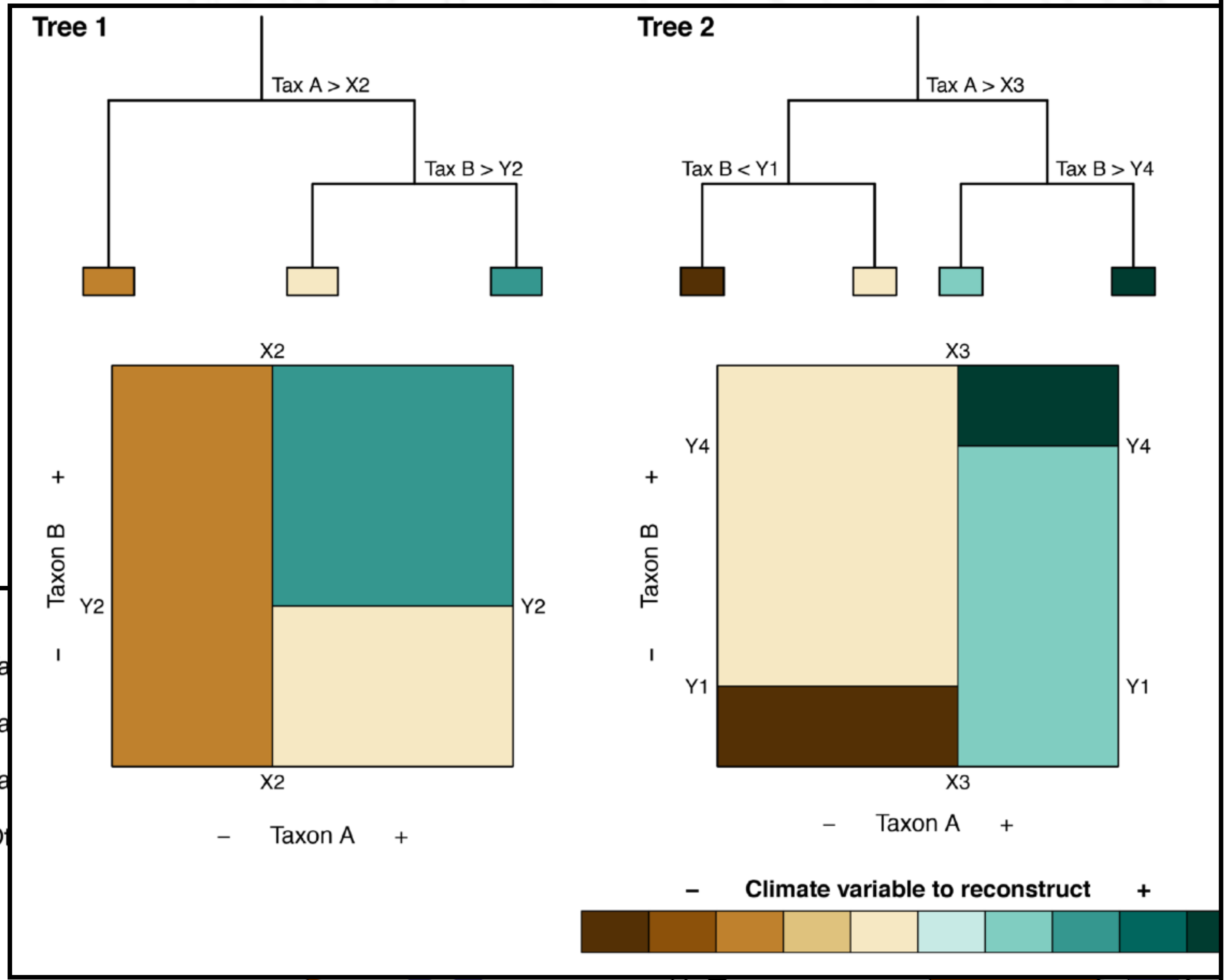


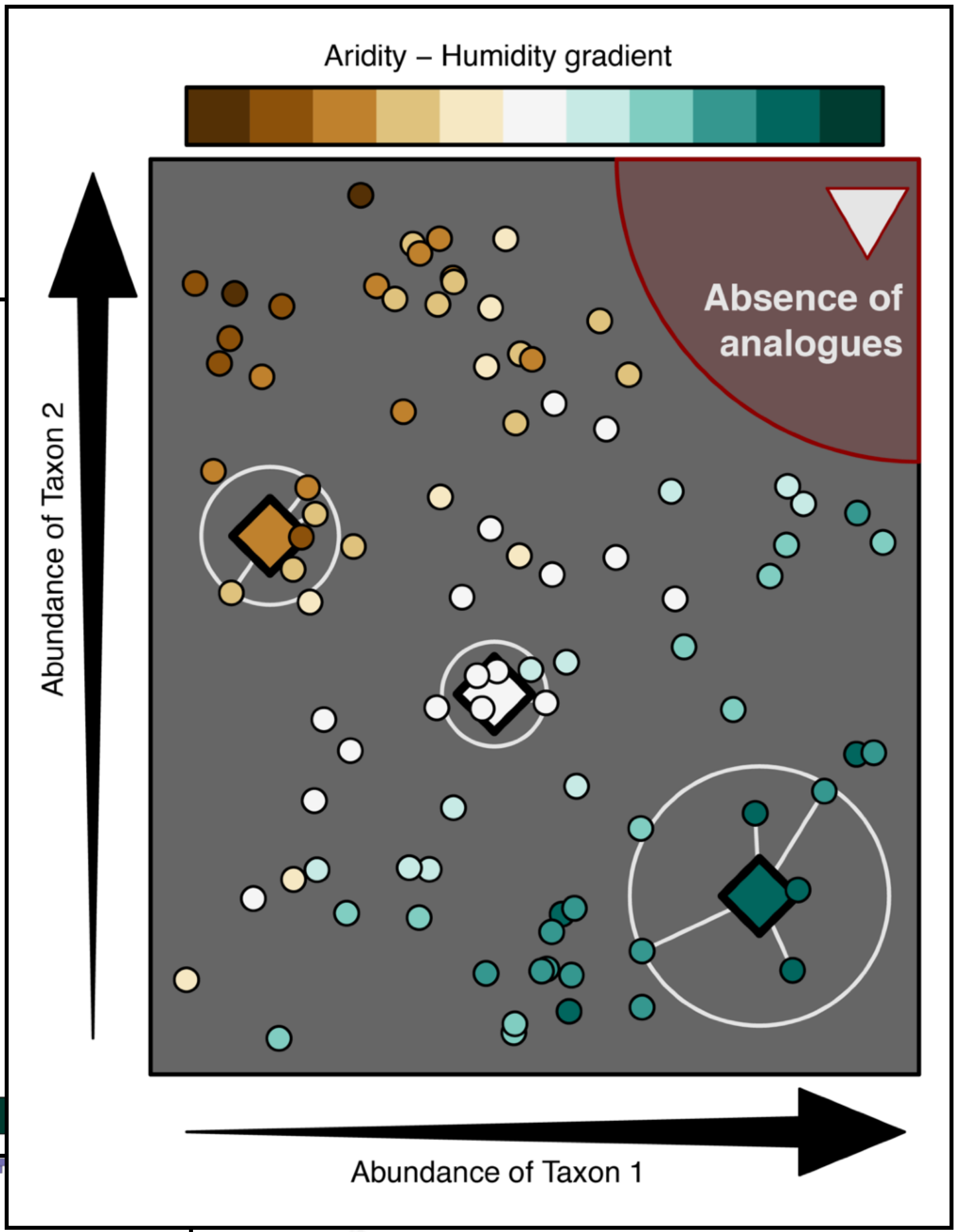
6. Two Ways of Reconstructing Climate from Pollen (1/2)

The "composition" approach



The 'tree-based' approaches



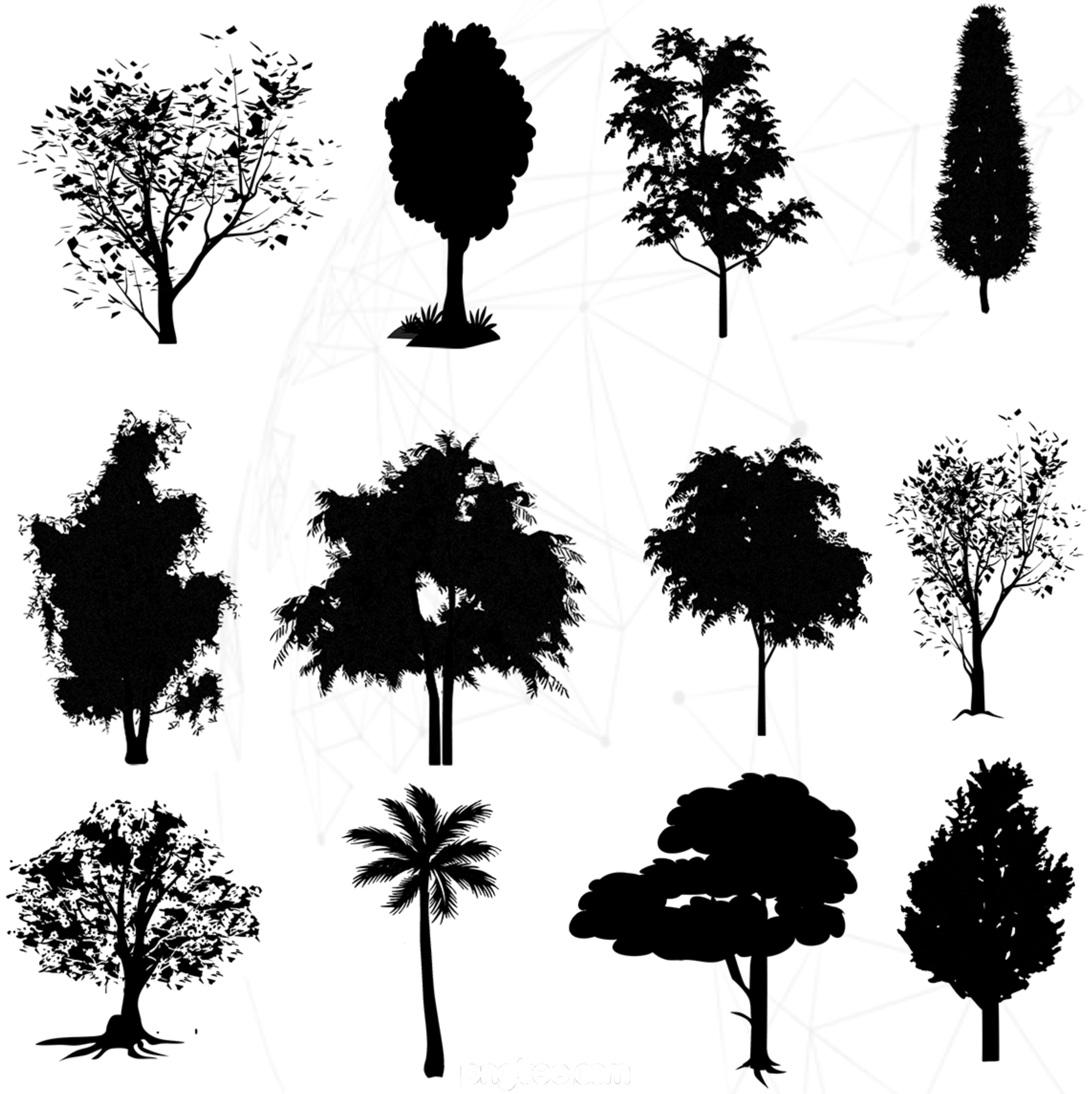


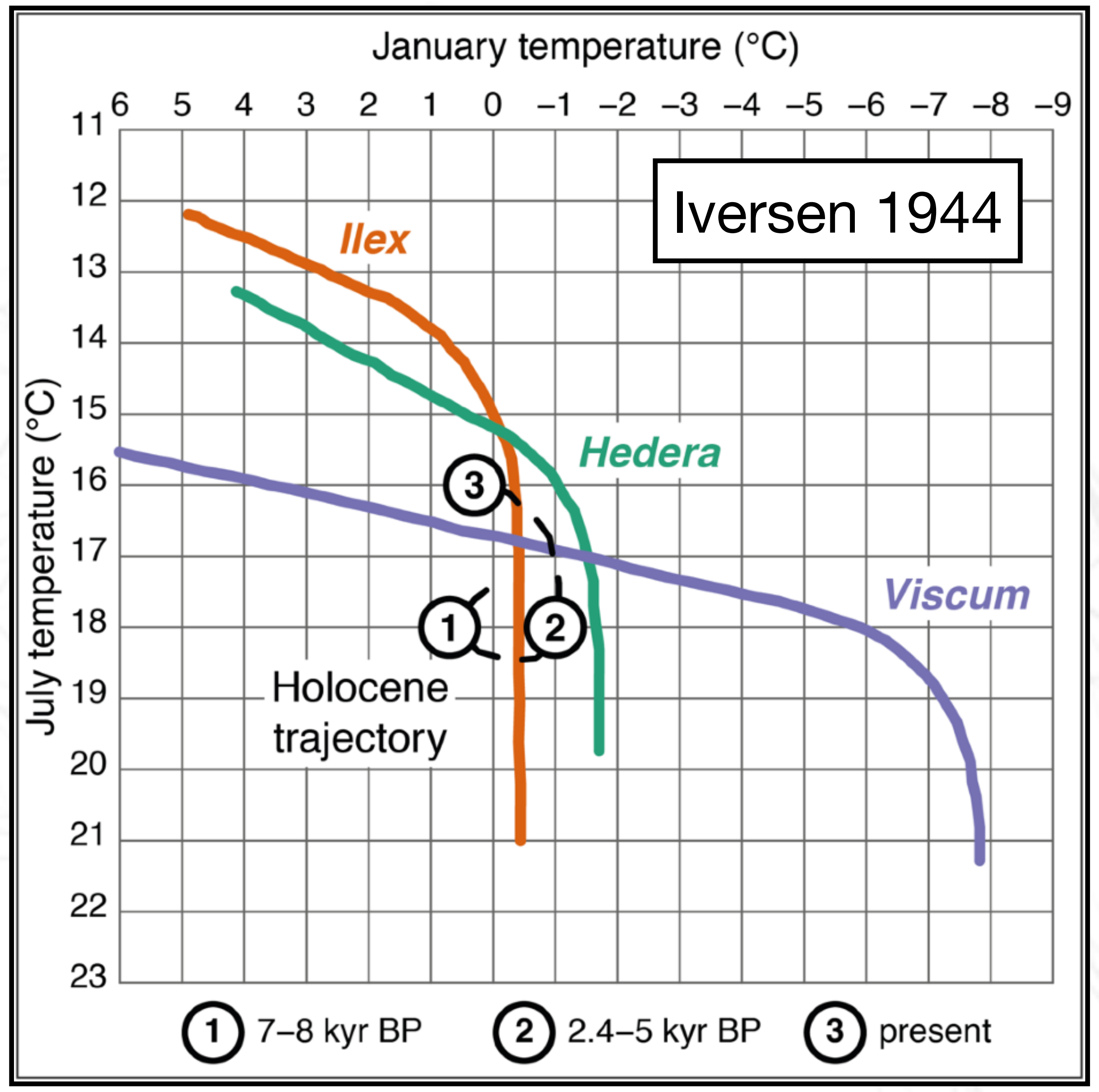
The 'modern analogue' approach

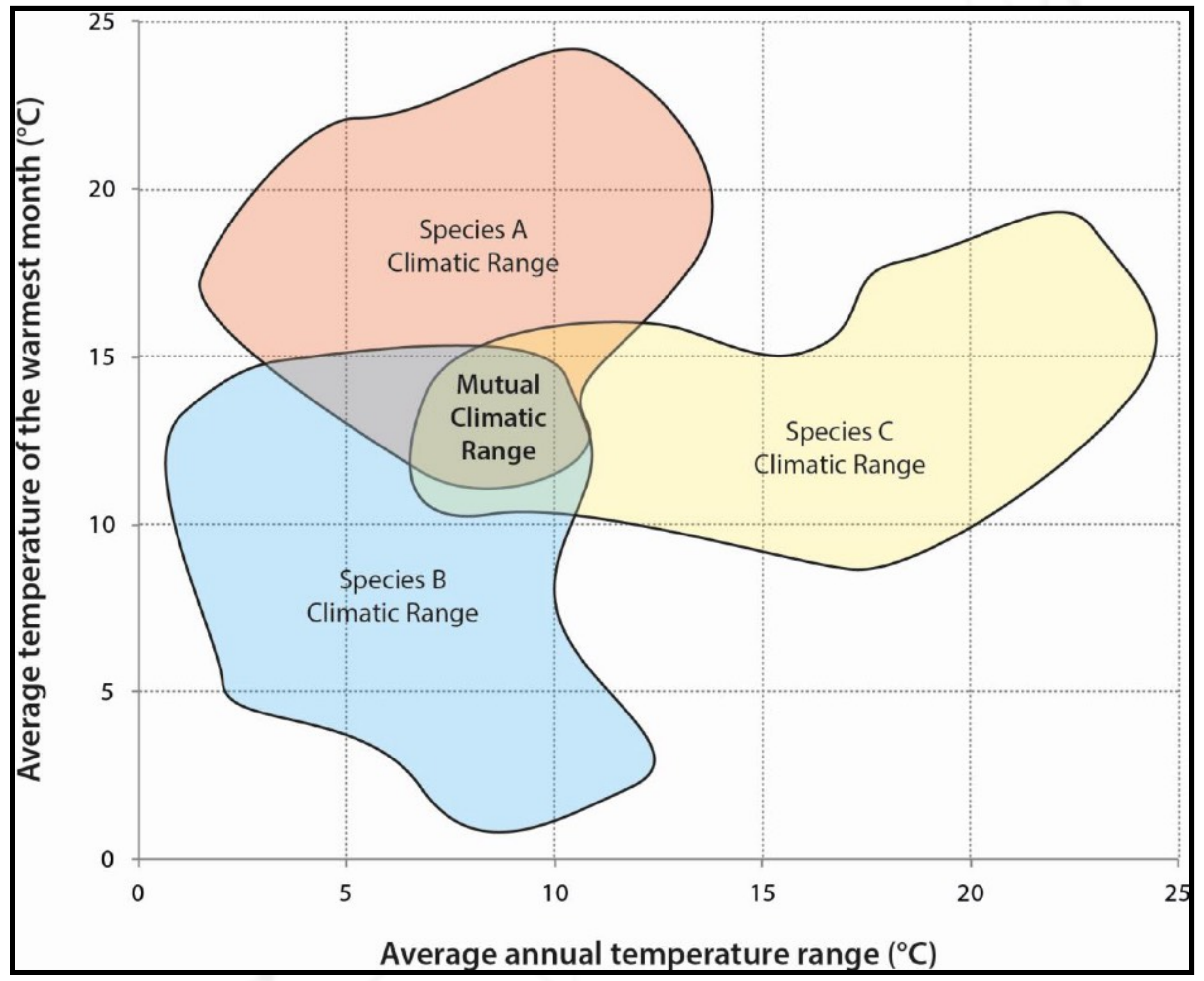
The 'regression' approach (WA / WA-PLS)

7. TWO WAYS OF RECONSTRUCTING CLIMATE FROM POLLEN (2/2)

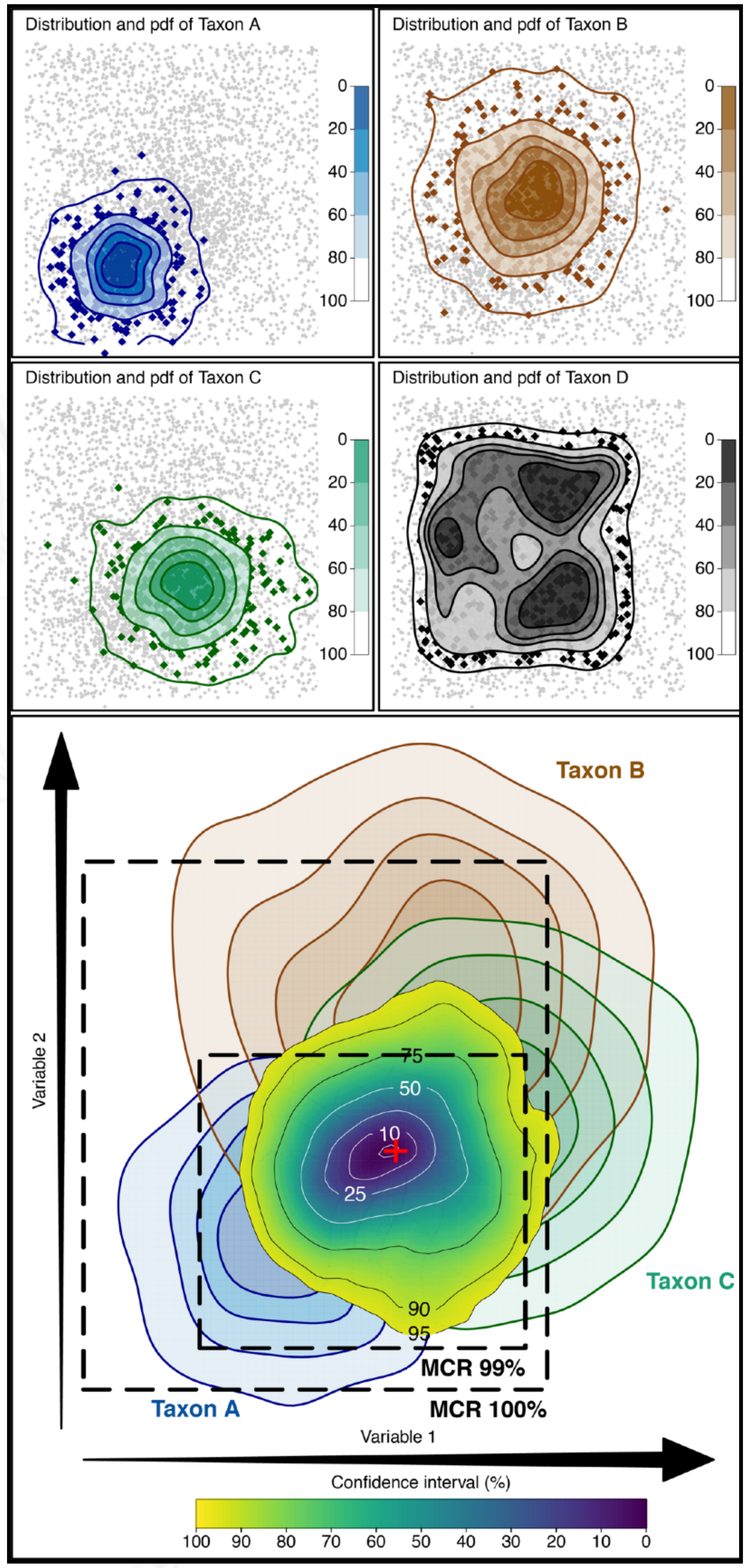
The "indicator species" approach







The Mutual Climate Range approach

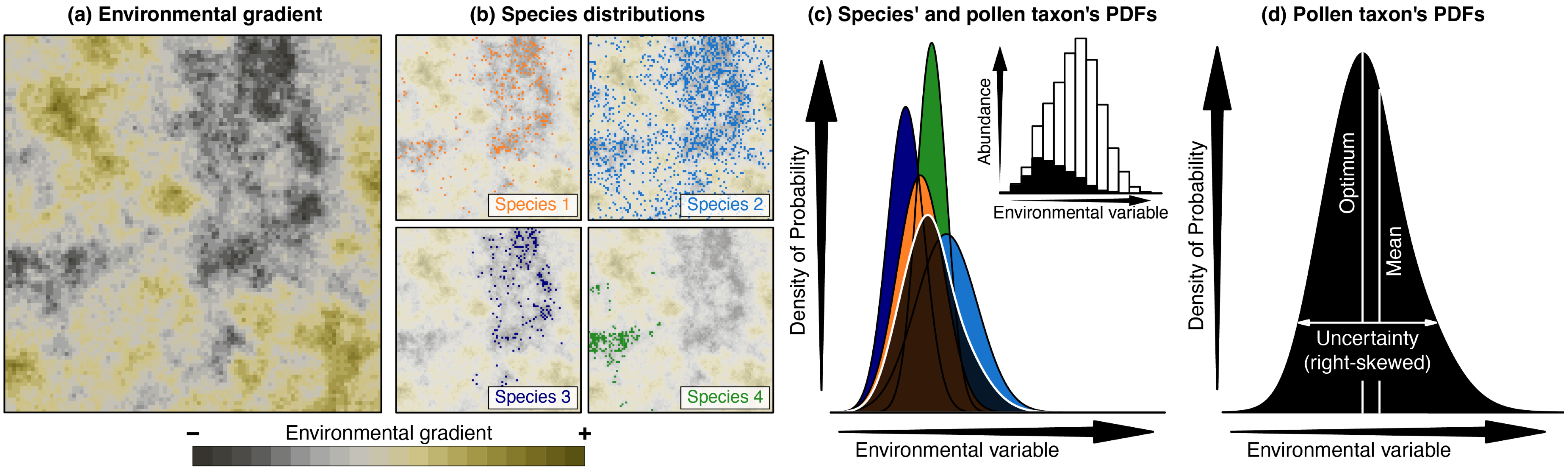


The Probability Density Function approach

TABLE OF CONTENT

How does it work?

10. THE BIG PICTURE

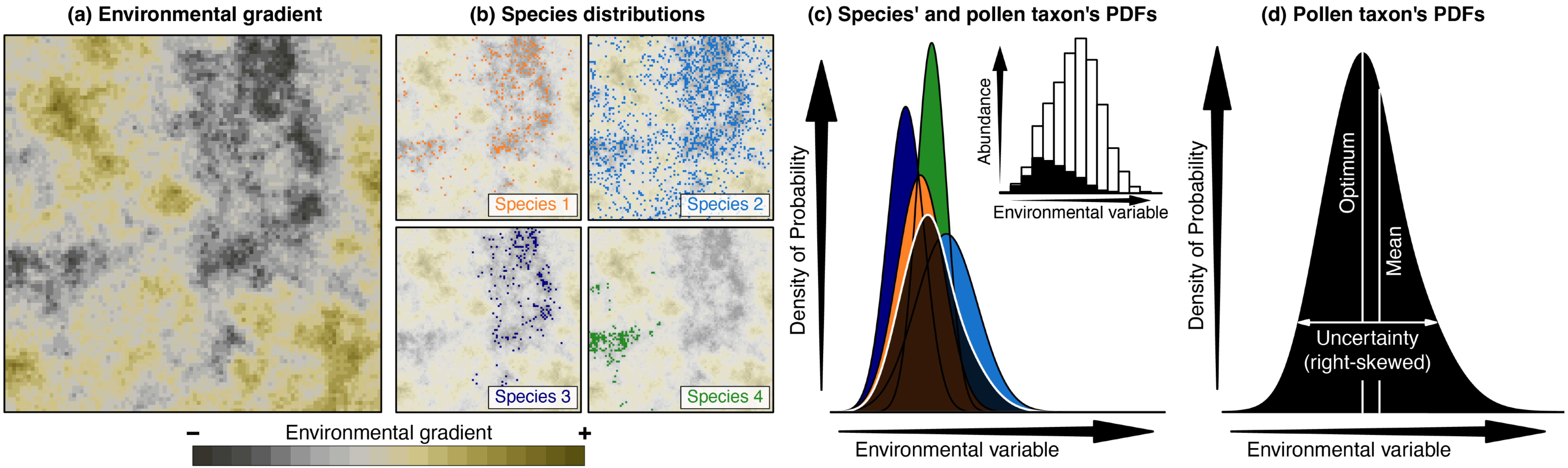


From the plants biogeographical information (A, B) to probabilistic climate signals (C, D).

Chevalier et al. (2014) Climate of the Past
Chevalier (2019) Global and Planetary Changes
Chevalier (2022) Climate of the Past

10. THE BIG PICTURE

1. Data acquisition: Where does the plant grow?



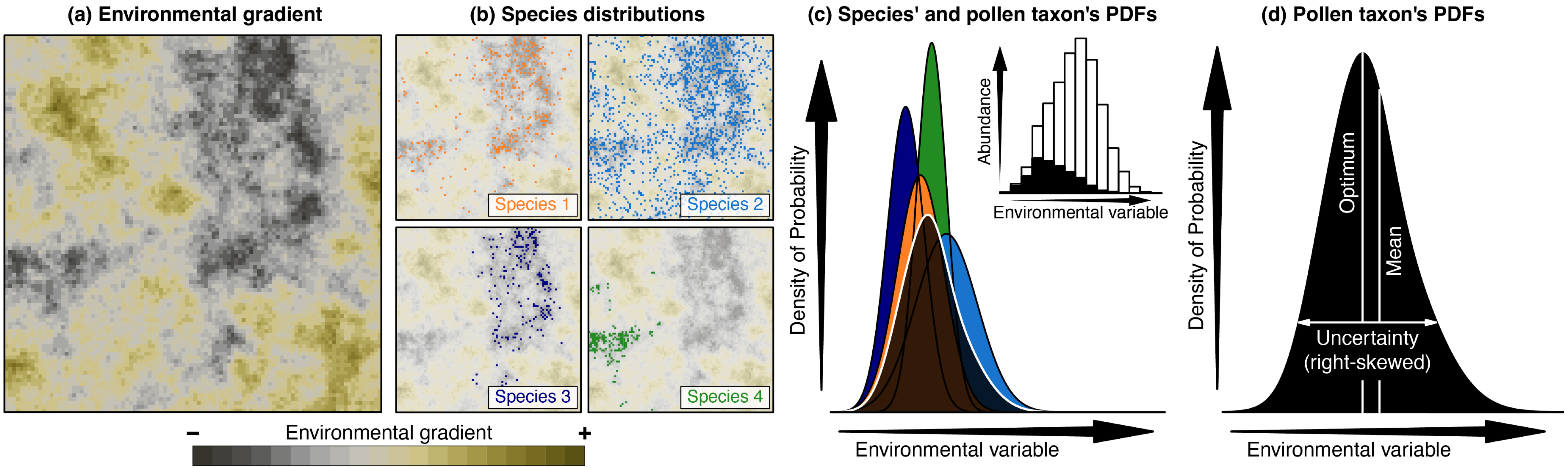
From the plants biogeographical information (A, B) to probabilistic climate signals (C, D).

Chevalier et al. (2014) Climate of the Past
Chevalier (2019) Global and Planetary Changes
Chevalier (2022) Climate of the Past

10. THE BIG PICTURE

1. Data acquisition: Where does the plant grow?

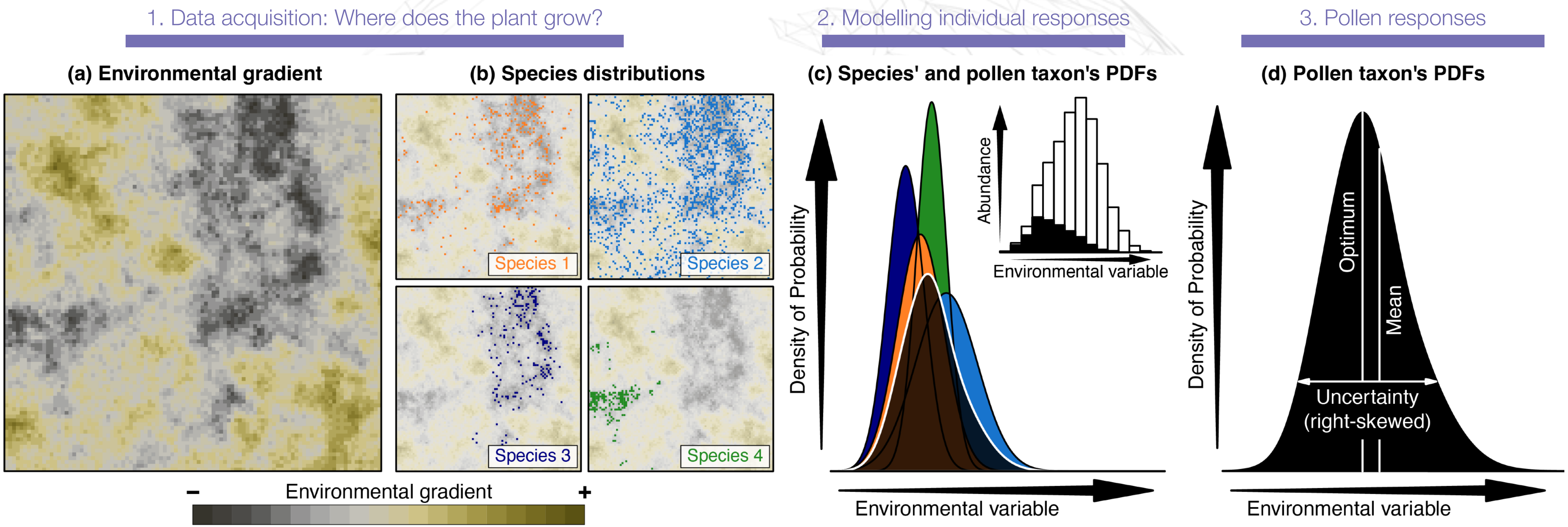
2. Modelling individual responses



From the plants biogeographical information (A, B) to probabilistic climate signals (C, D).

Chevalier et al. (2014) Climate of the Past
Chevalier (2019) Global and Planetary Changes
Chevalier (2022) Climate of the Past

10. THE BIG PICTURE



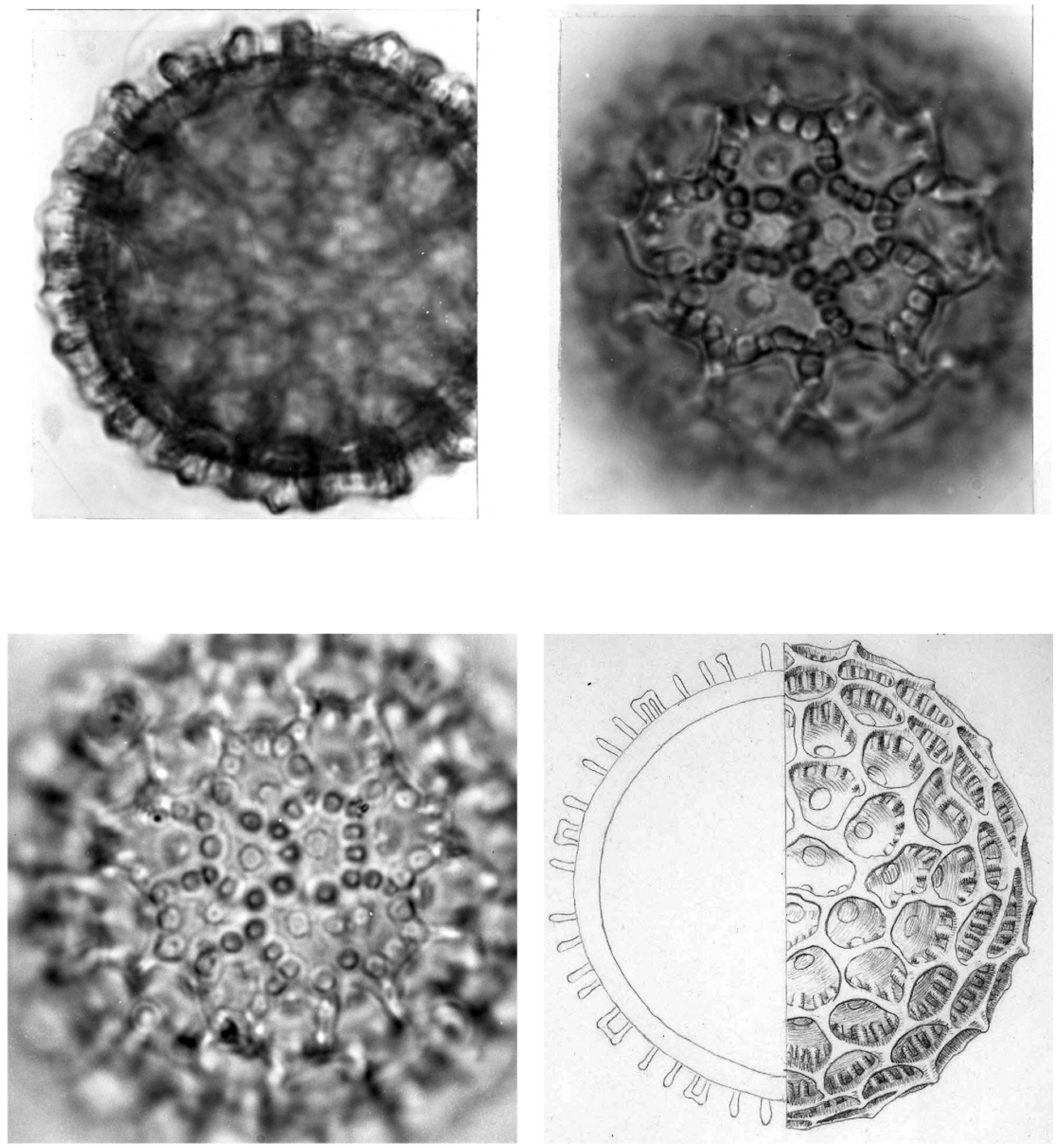
From the plants biogeographical information (A, B) to probabilistic climate signals (C, D).

Chevalier et al. (2014) Climate of the Past
Chevalier (2019) Global and Planetary Changes
Chevalier (2022) Climate of the Past

11. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

THE POLLEN TAXONOMIC RESOLUTION UNCERTAINTY

How to estimate the response of a pollen taxon?

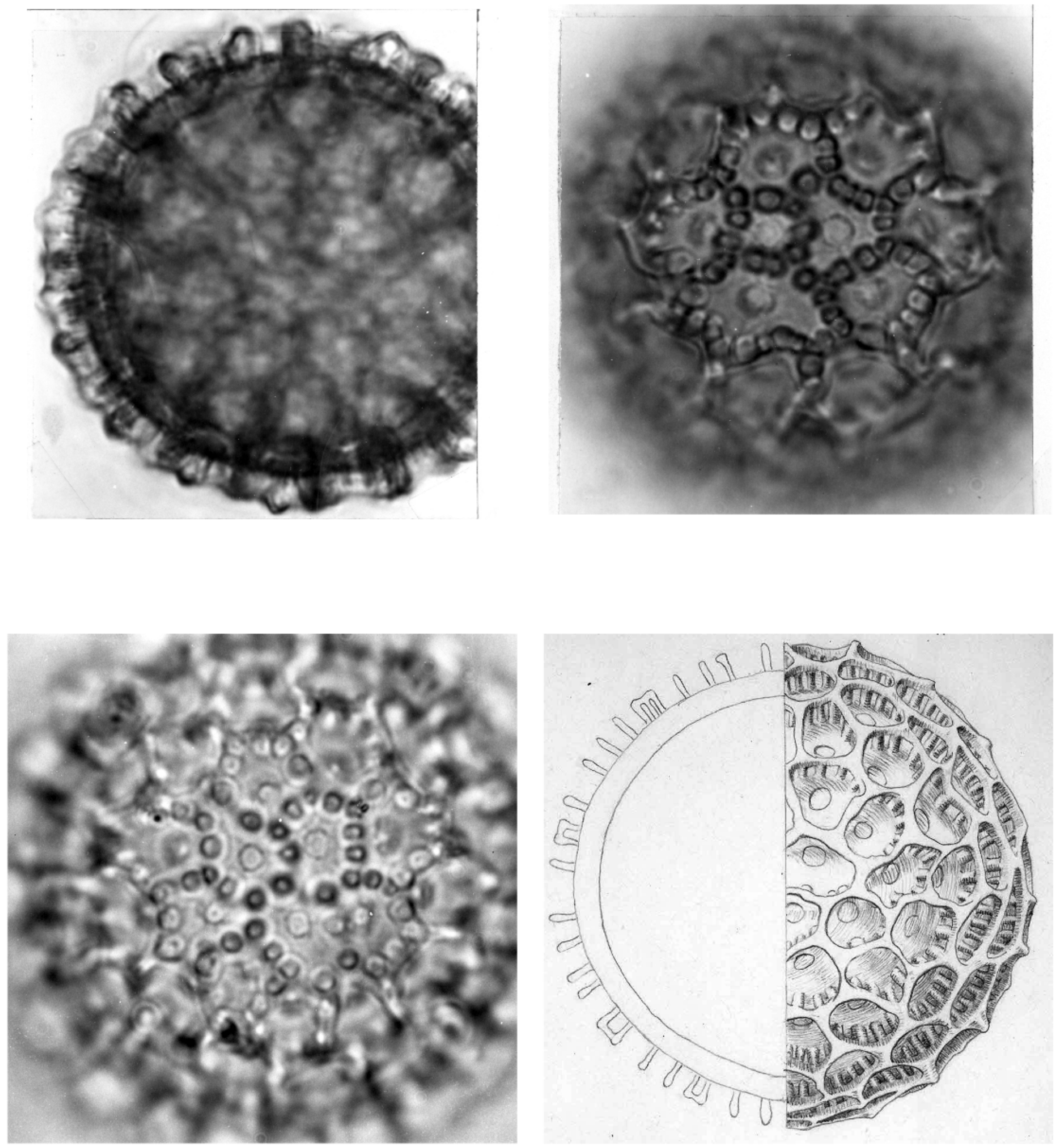


Zygophyllaceae
Tribulus terrestris (45 µm)

12. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

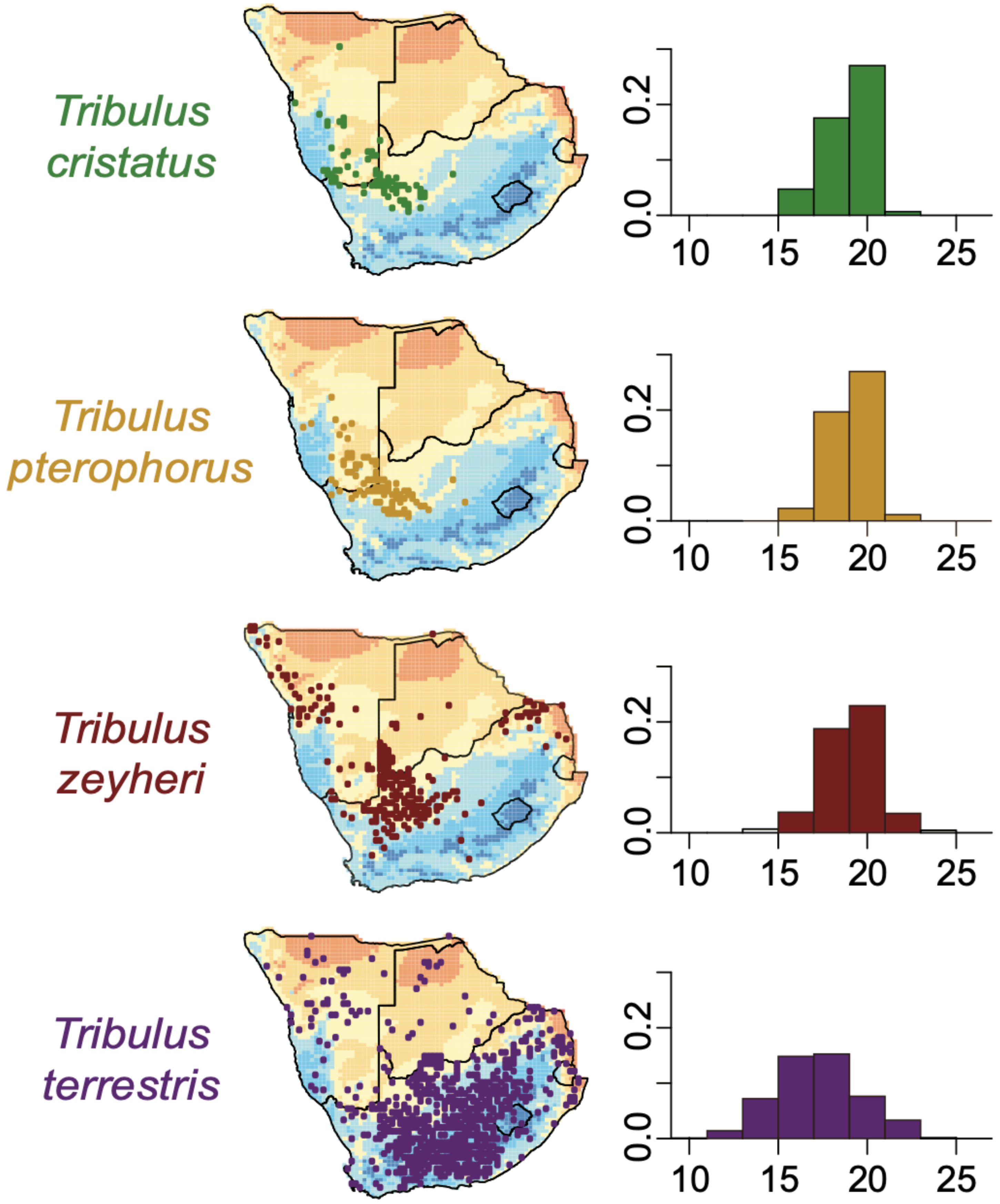
THE POLLEN TAXONOMIC RESOLUTION UNCERTAINTY

How to estimate the response of a pollen taxon?



Zygophyllaceae
Tribulus terrestris (45 µm)

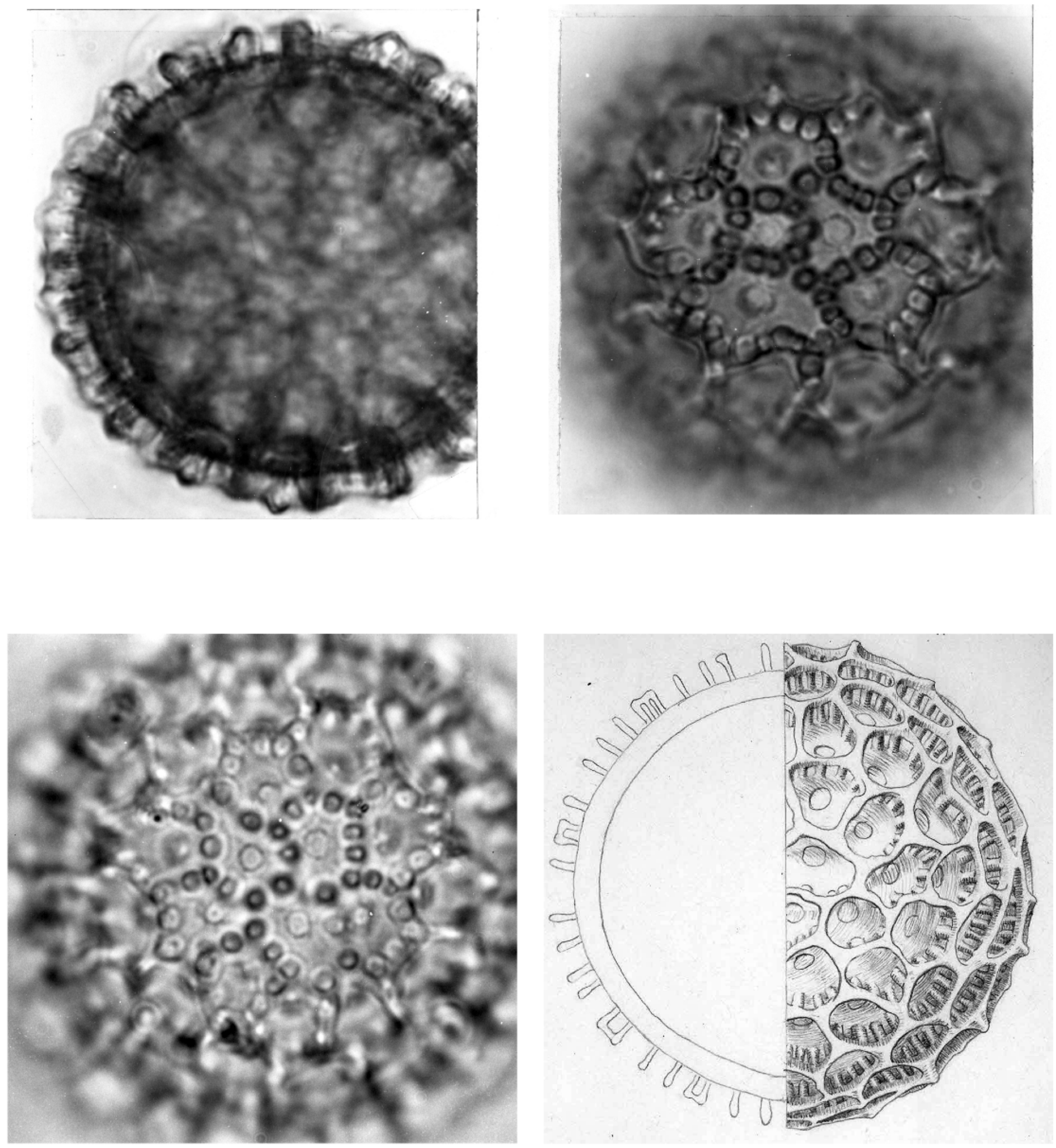
4 species in southern Africa
Here, represented against temperature.



12. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

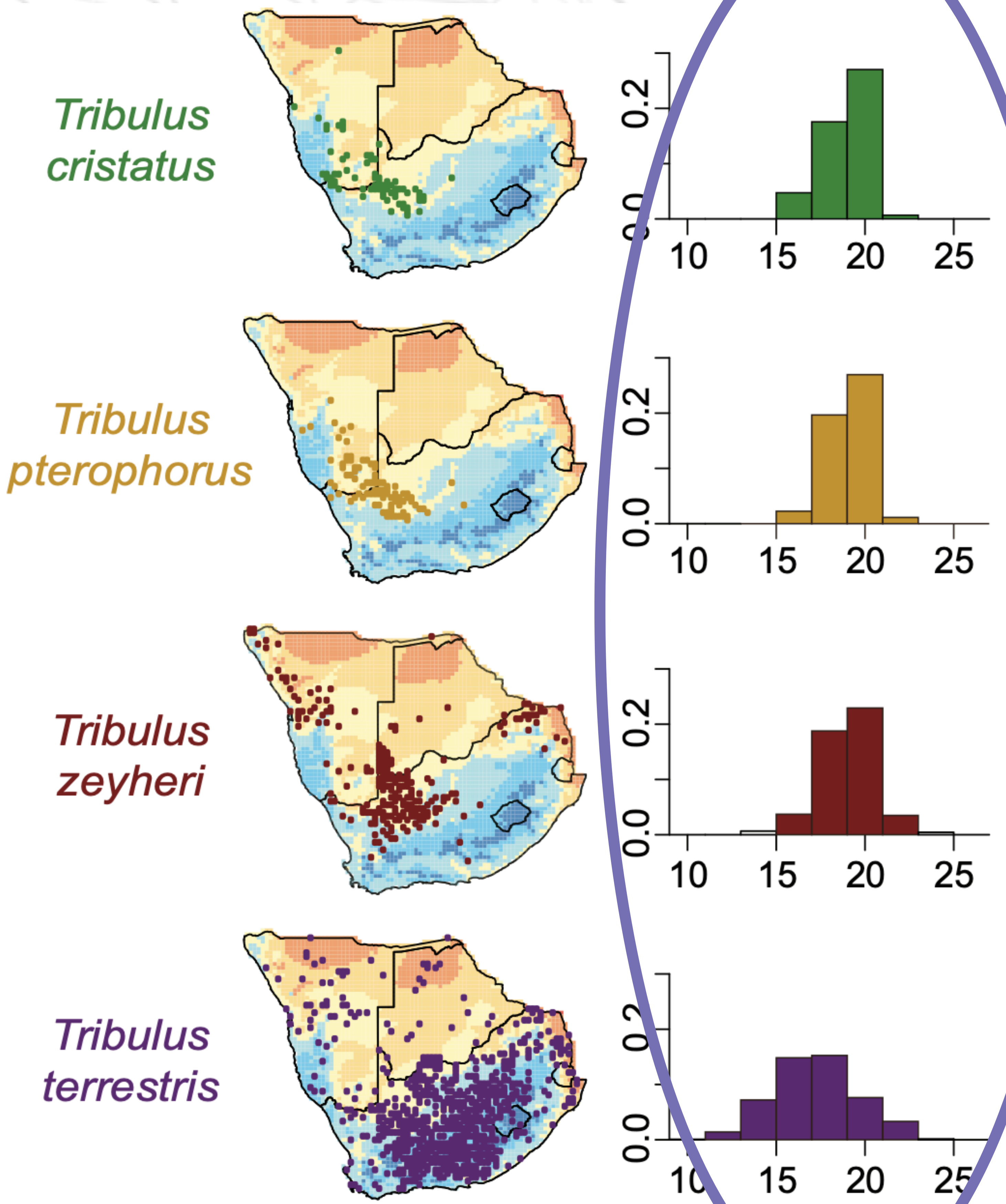
THE POLLEN TAXONOMIC RESOLUTION UNCERTAINTY

How to estimate the response of a pollen taxon?



Zygophyllaceae
Tribulus terrestris (45 µm)

4 species in southern Africa
Here, represented against temperature.

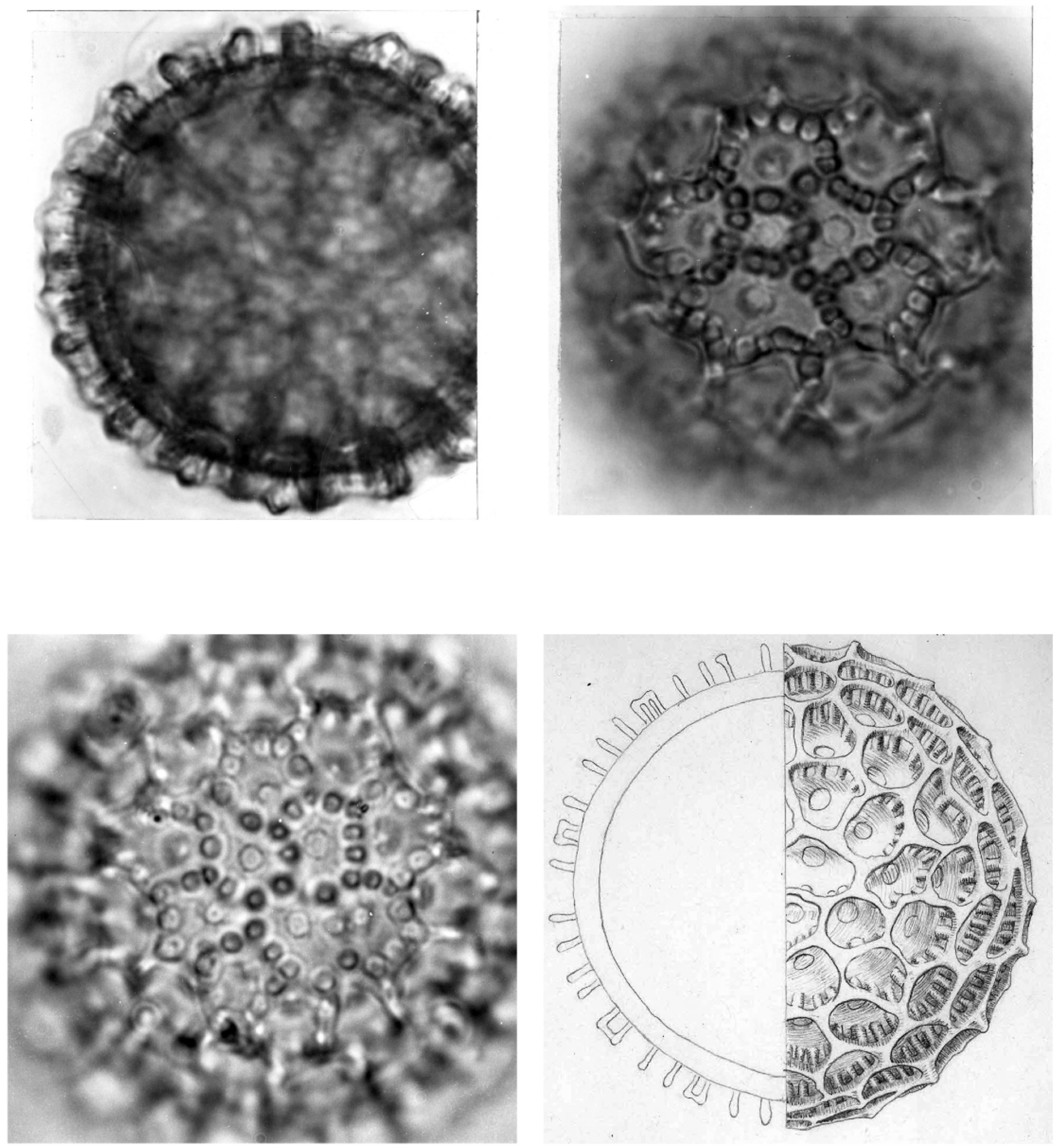


How to model these observations?

13. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

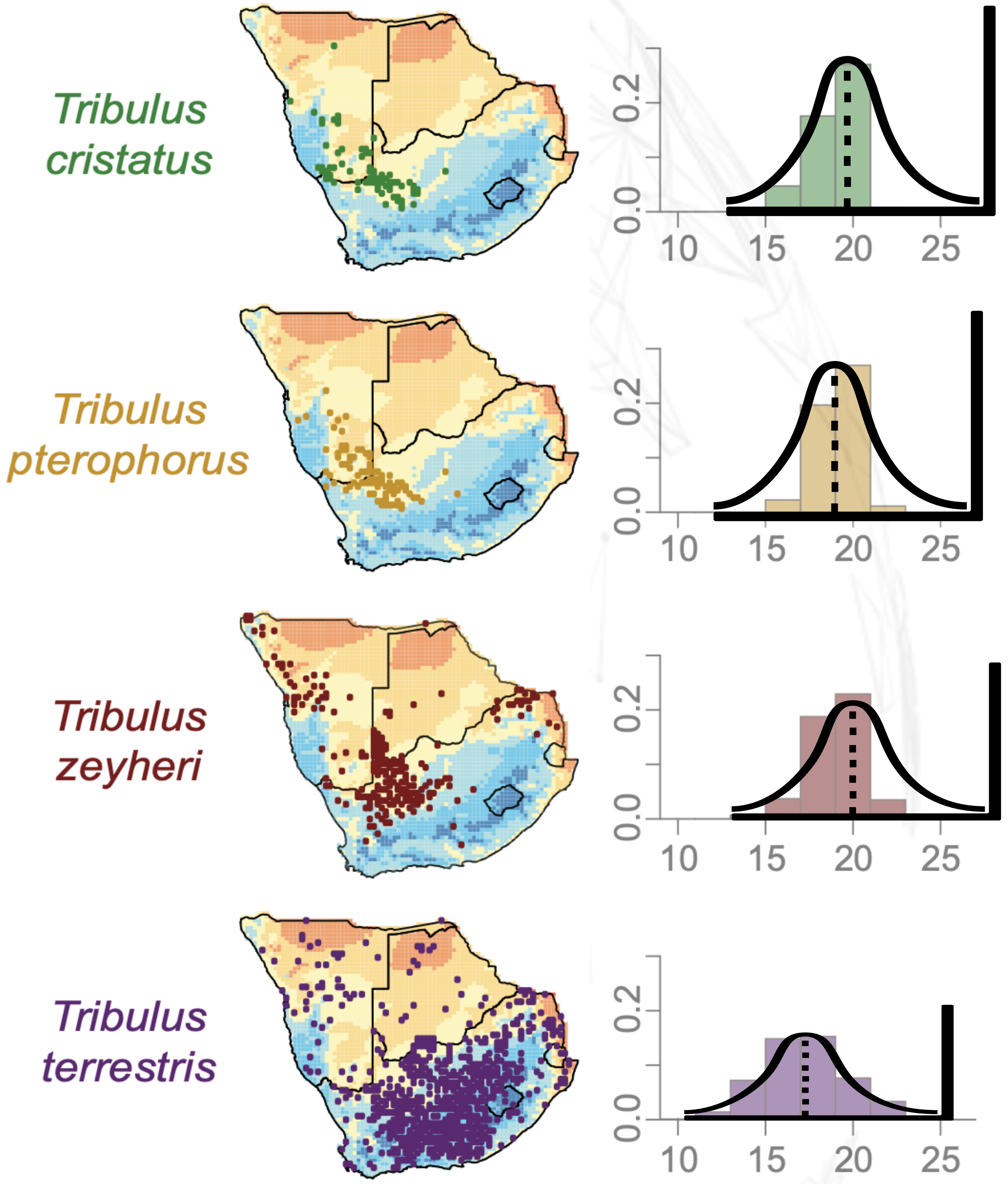
THE POLLEN TAXONOMIC RESOLUTION UNCERTAINTY

How to estimate the response of a pollen taxon?



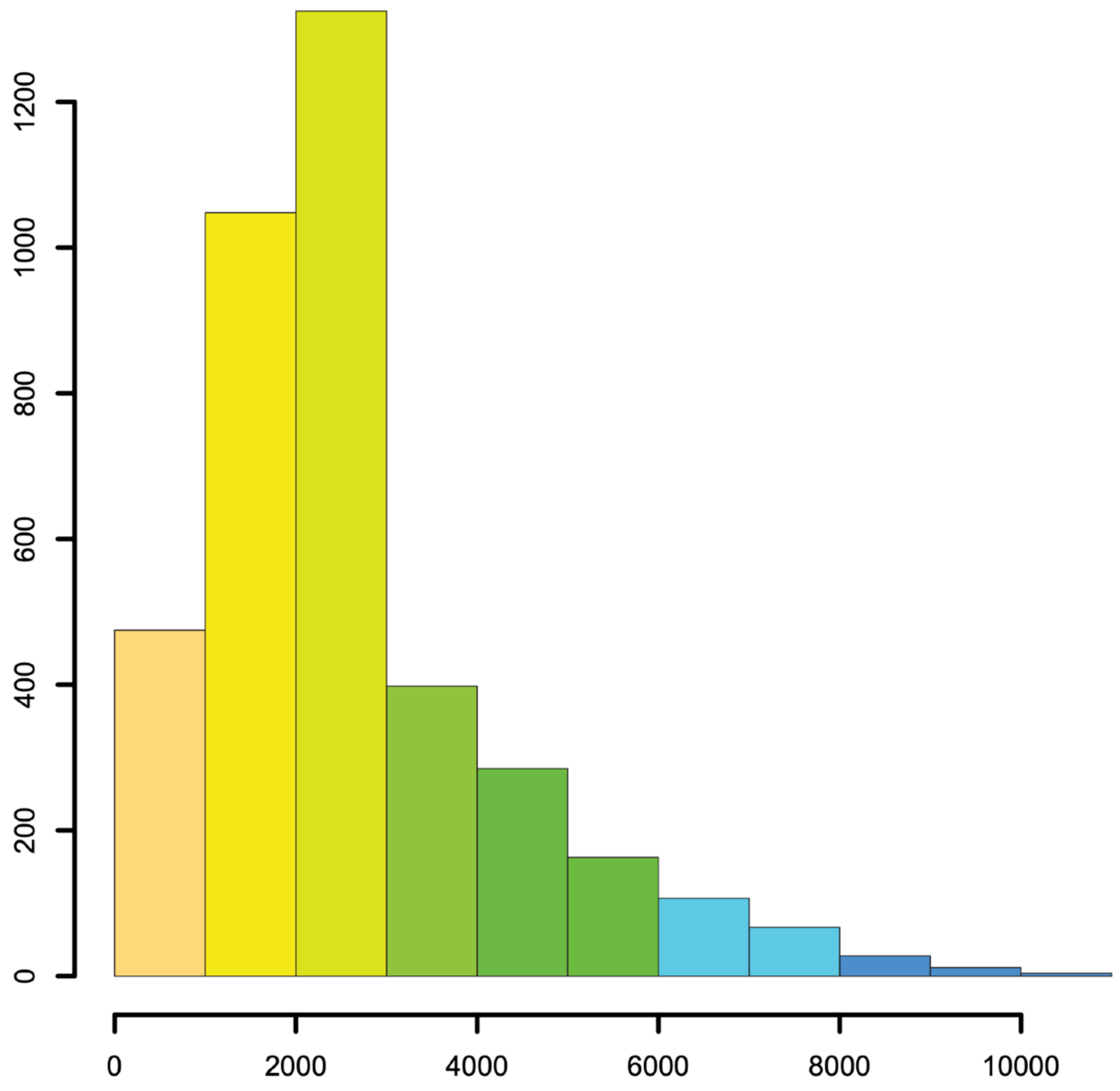
Zygophyllaceae
Tribulus terrestris (45 µm)

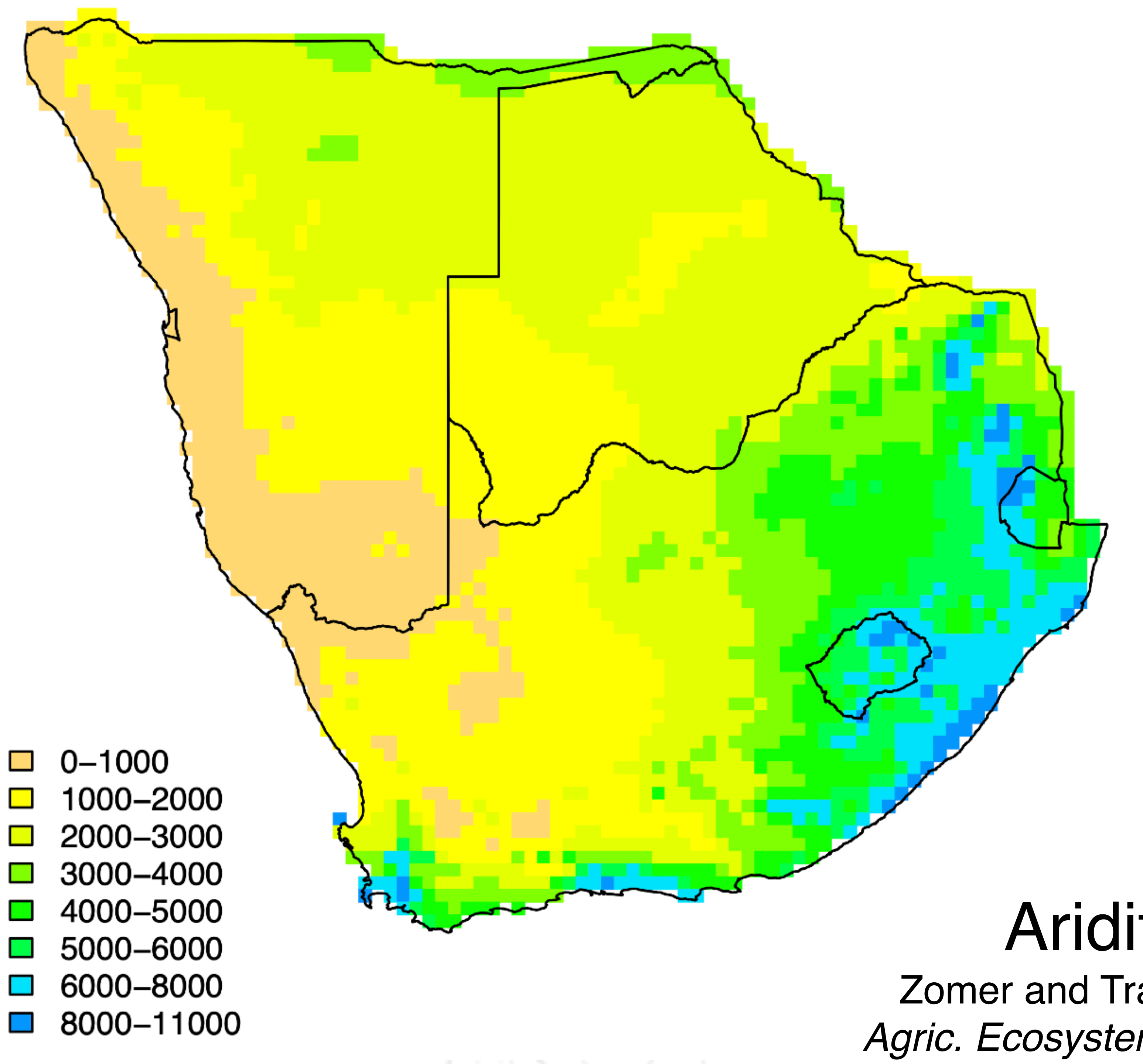
4 species in southern Africa
Here, represented against temperature.



How to model these observations?

14. Accounting for Unbalanced Climate Space



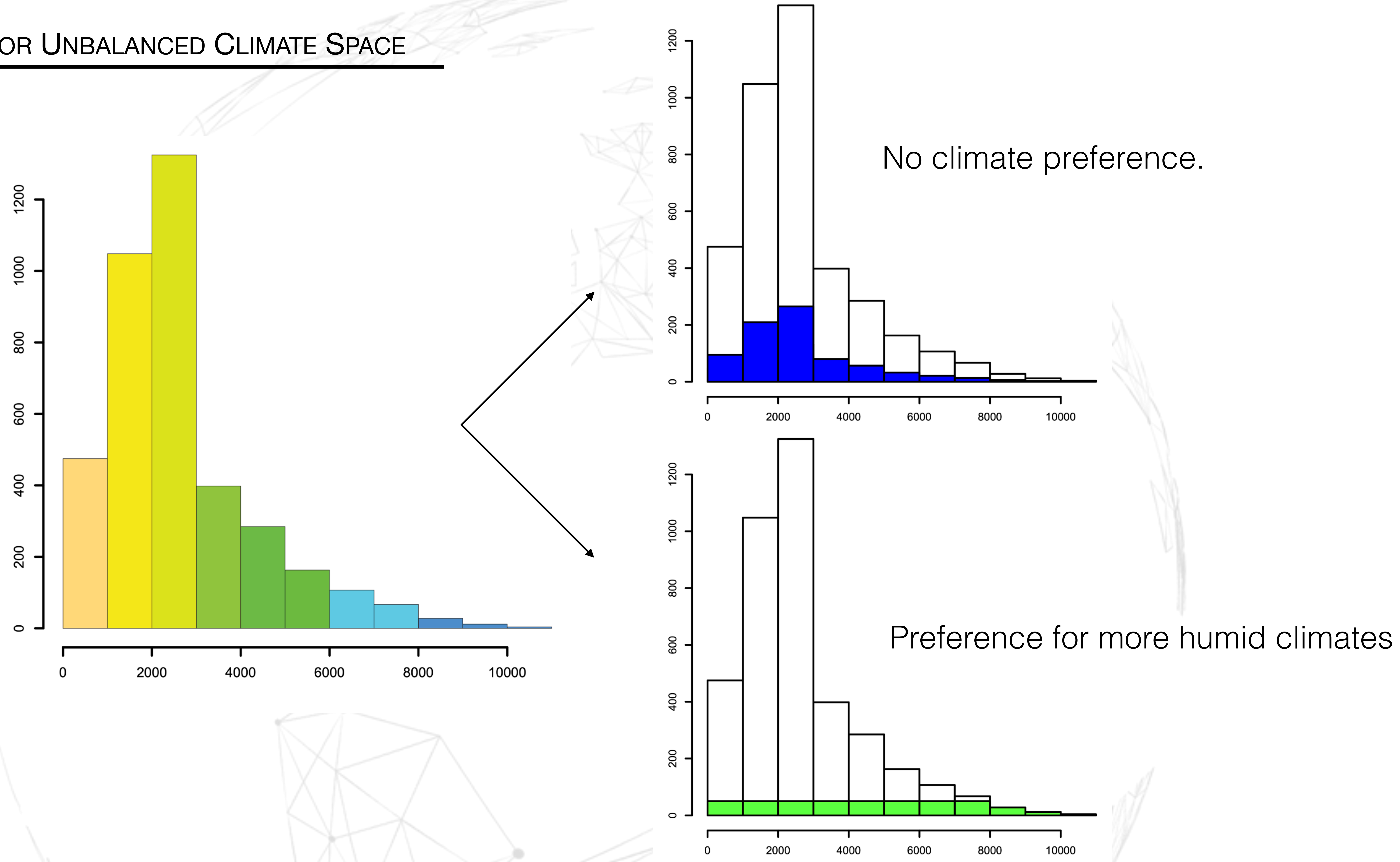


Aridity Index

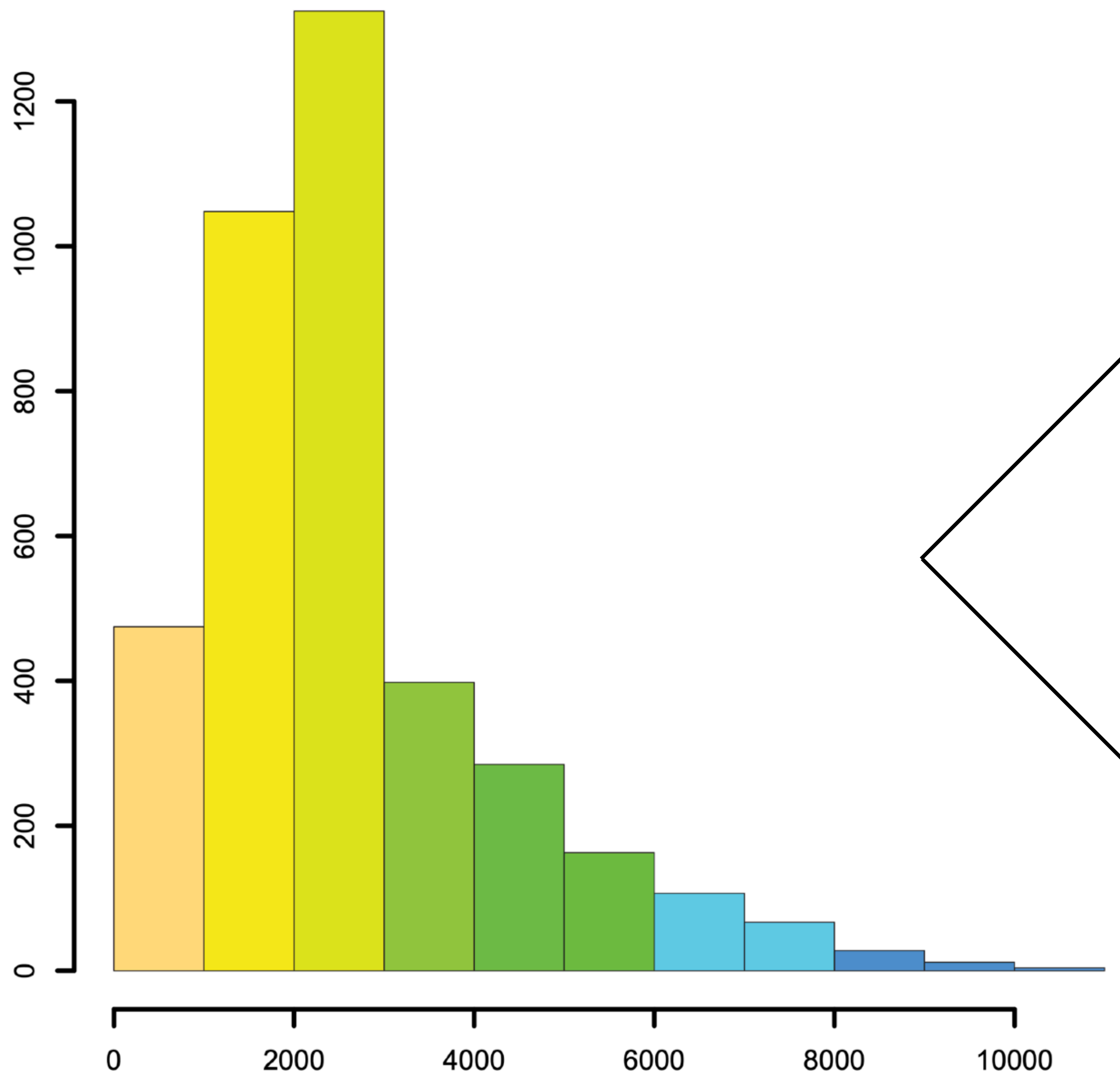
Zomer and Trabucco, 2008

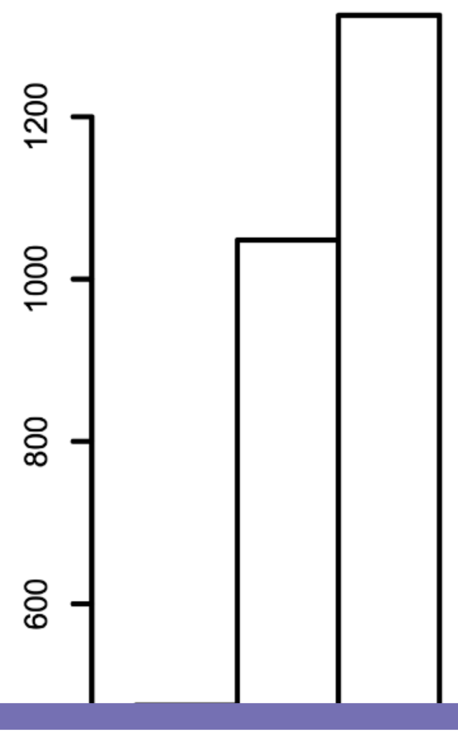
Agric. Ecosystems and Envir.

14. ACCOUNTING FOR UNBALANCED CLIMATE SPACE



14. ACCOUNTING FOR UNBALANCED CLIMATE SPACE



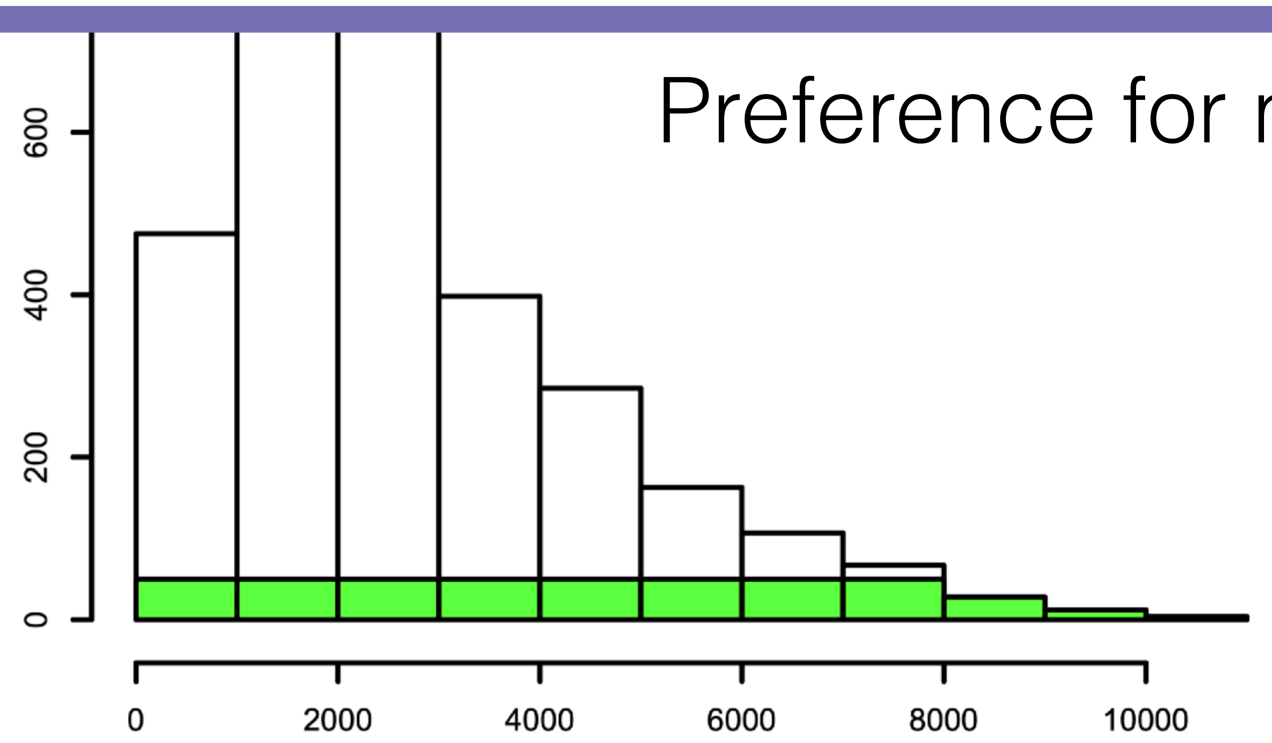


No climate preference.

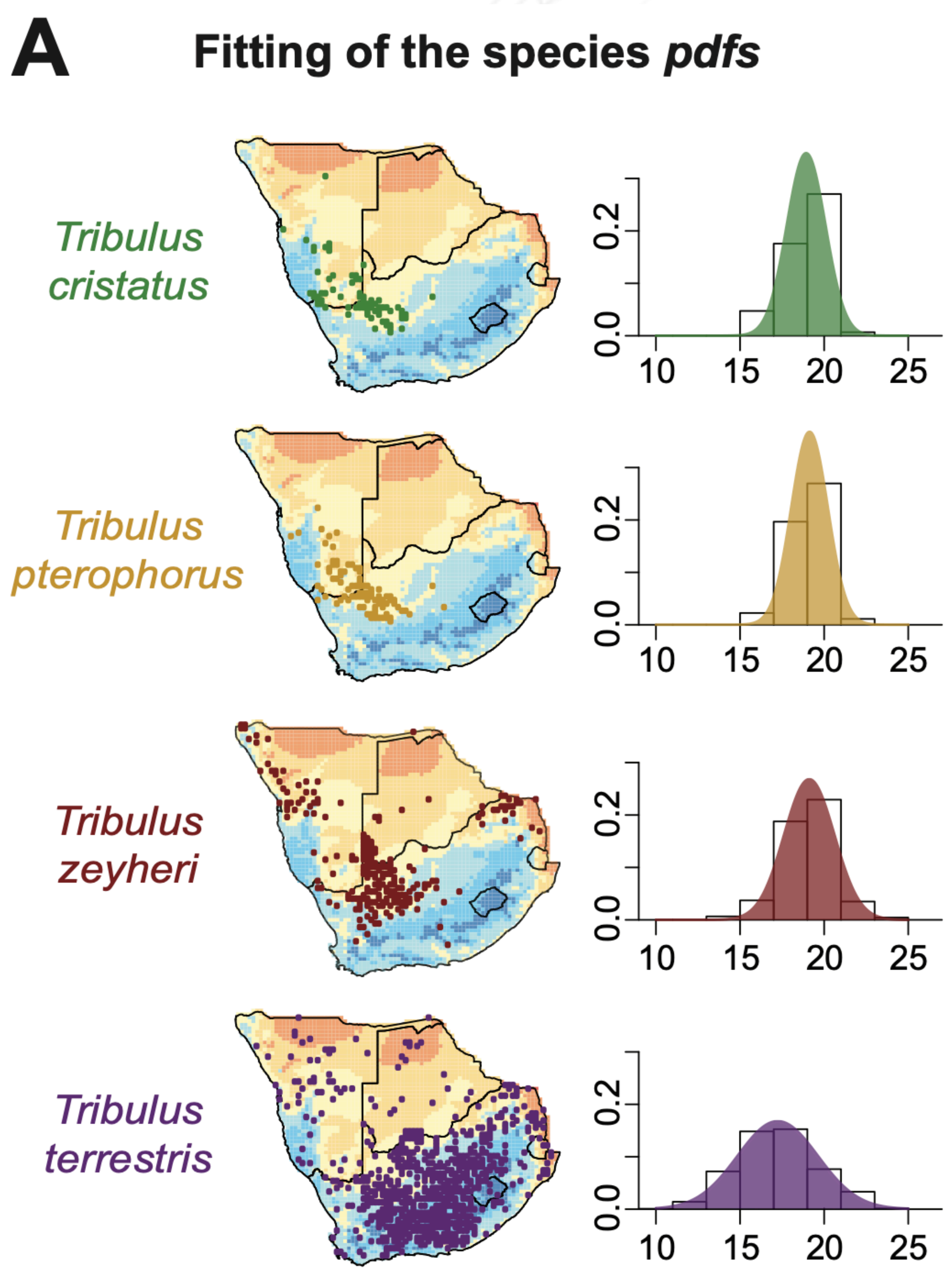
$$\bar{x}_{sp} = \frac{1}{\sum k_i} \sum_{i=1}^{N} k_i x_i$$

$$s^2_{x,sp} = \frac{1}{\sum k_i} \sum_{i=1}^{N} k_i \left(x_i - \bar{x}_{sp}\right)^2$$

Preference for more humid climates



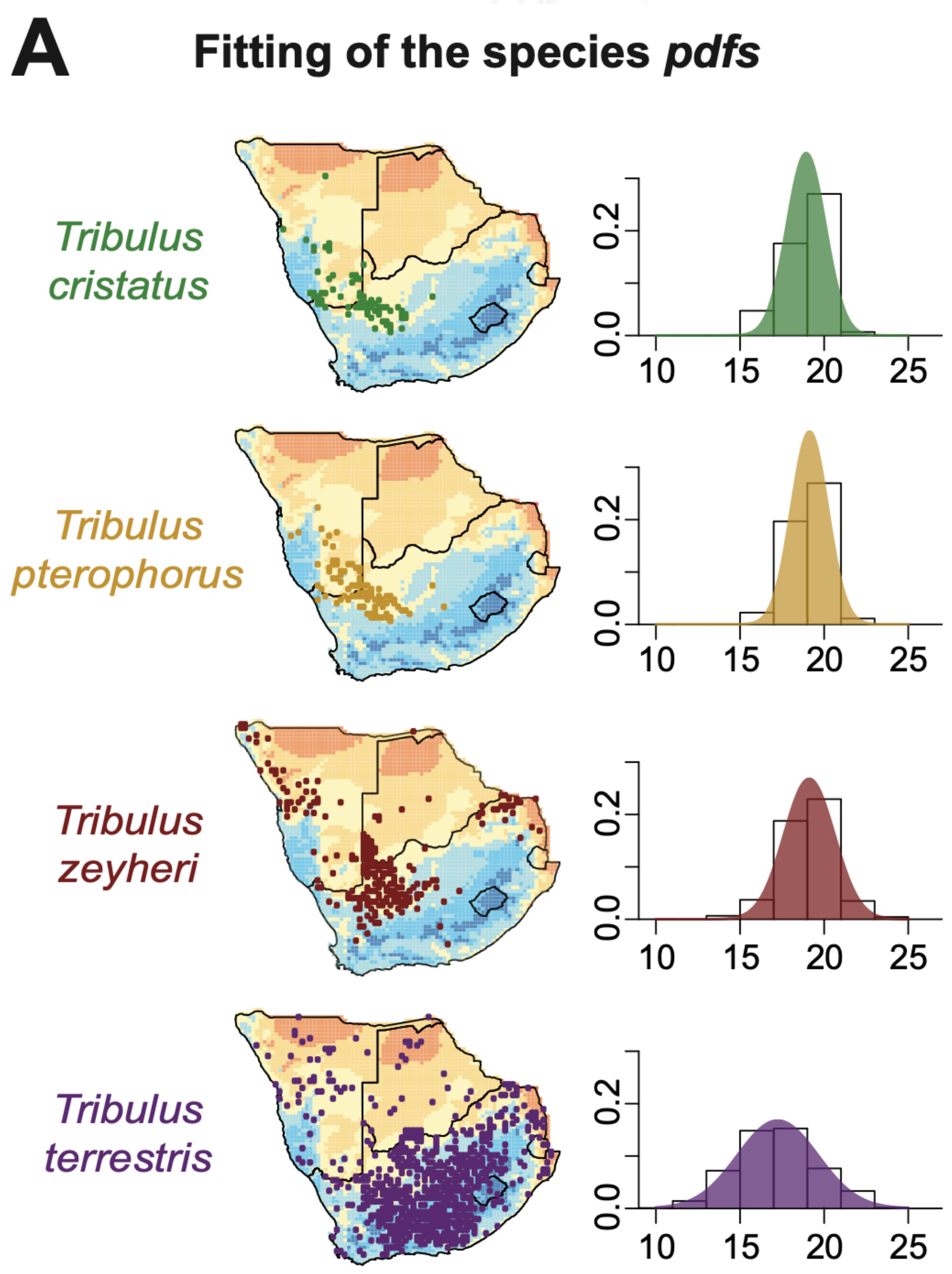
15. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS



4 species in southern Africa
Here, represented against temperature.

- One *pdf* for each species and variable ***pdf$_{sp}$(v)***
- Shape of the ***pdf$_{sp}$(v)*** defined *a priori*:
  - normal
  - log-normal (right-skewed and excludes negative values)
- ~20 occurrences minimum

15. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

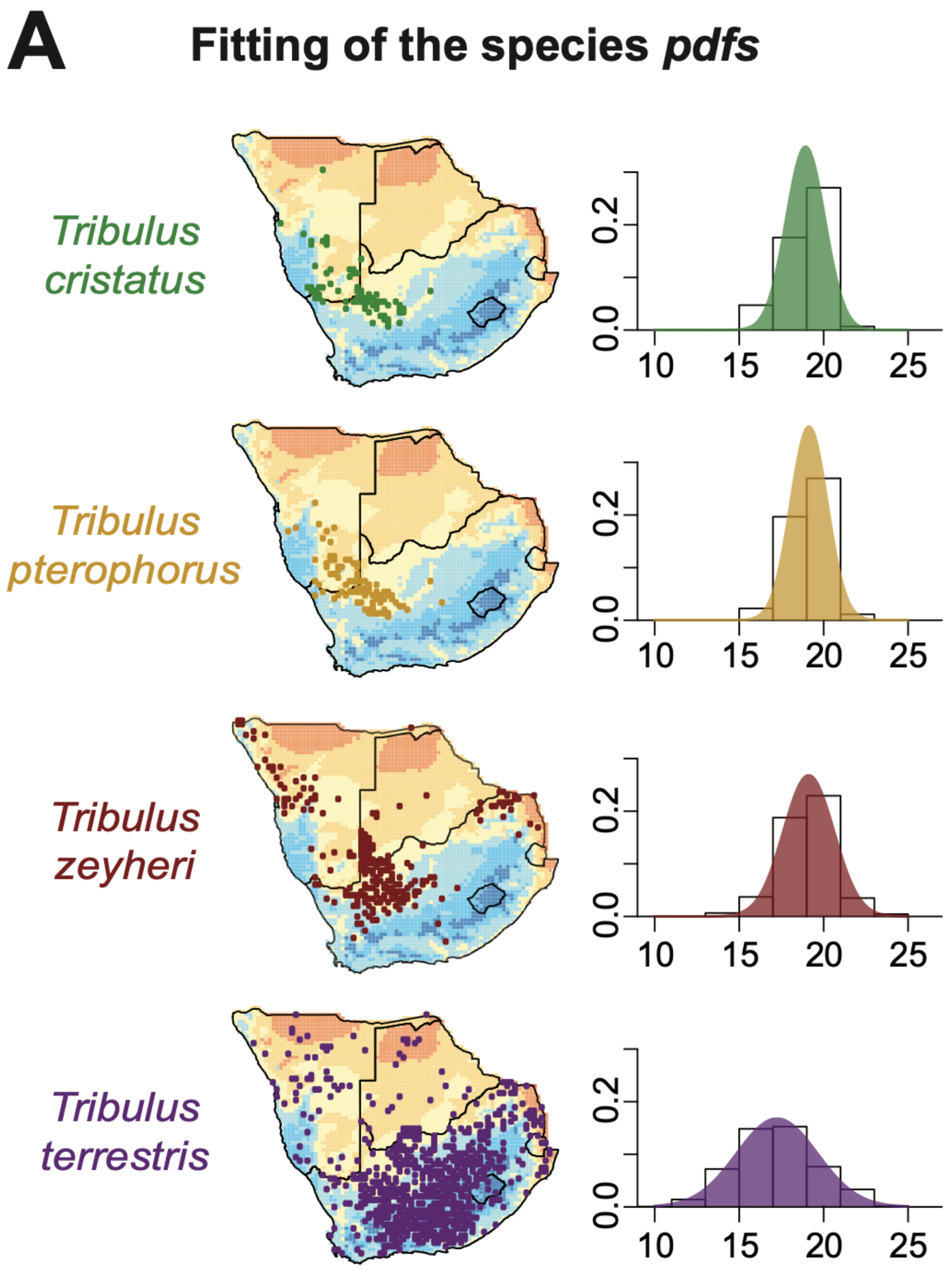


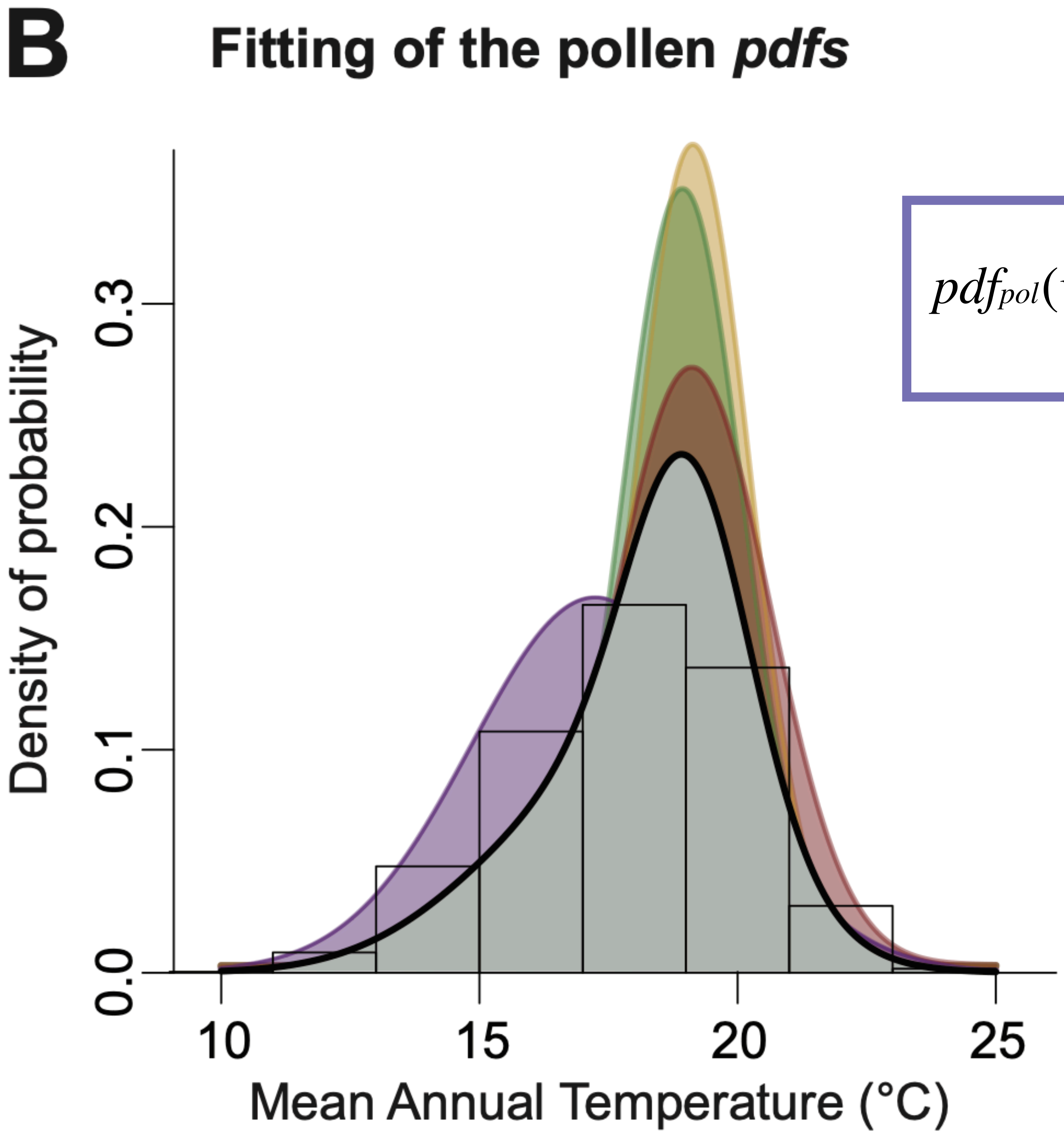
4 species in southern Africa
Here, represented against temperature.

- One *pdf* for each species and variable ***pdf$_{sp}$(v)***
- Shape of the ***pdf$_{sp}$(v)*** defined *a priori*:
  - normal
  - log-normal (right-skewed and excludes negative values)
- ~20 occurrences minimum

This is why CREST is applicable in many regions.
It does not require too many datapoints.

16. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS





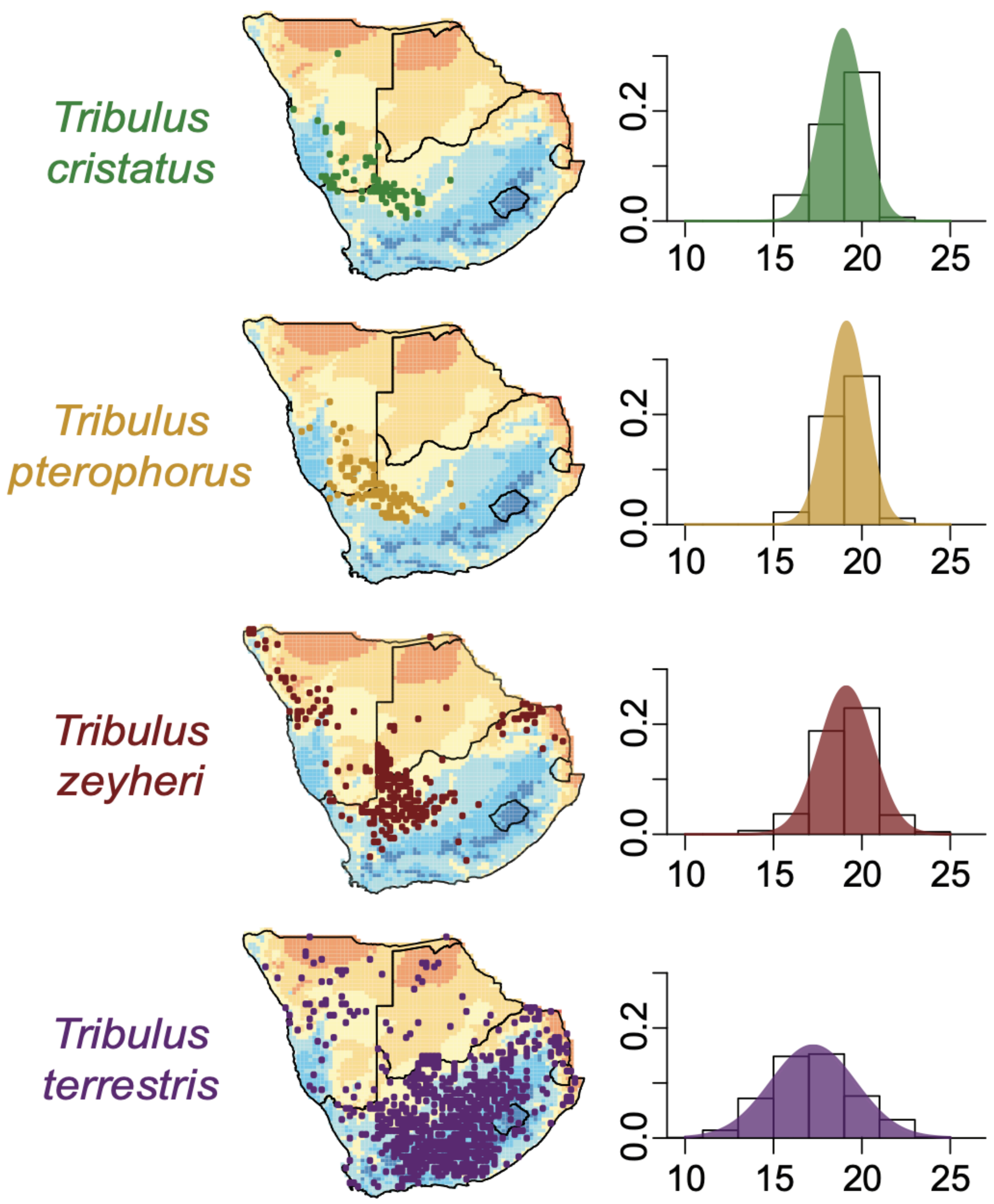
$$pdf_{pol}(v) = \frac{1}{\sum_{spi} \sqrt{n_{spi}}} \sum_{sp1}^{spN} \sqrt{n_{spi}}\, pdf_{spi}(v)$$

This represents the full 'breath' of the pollen taxon.

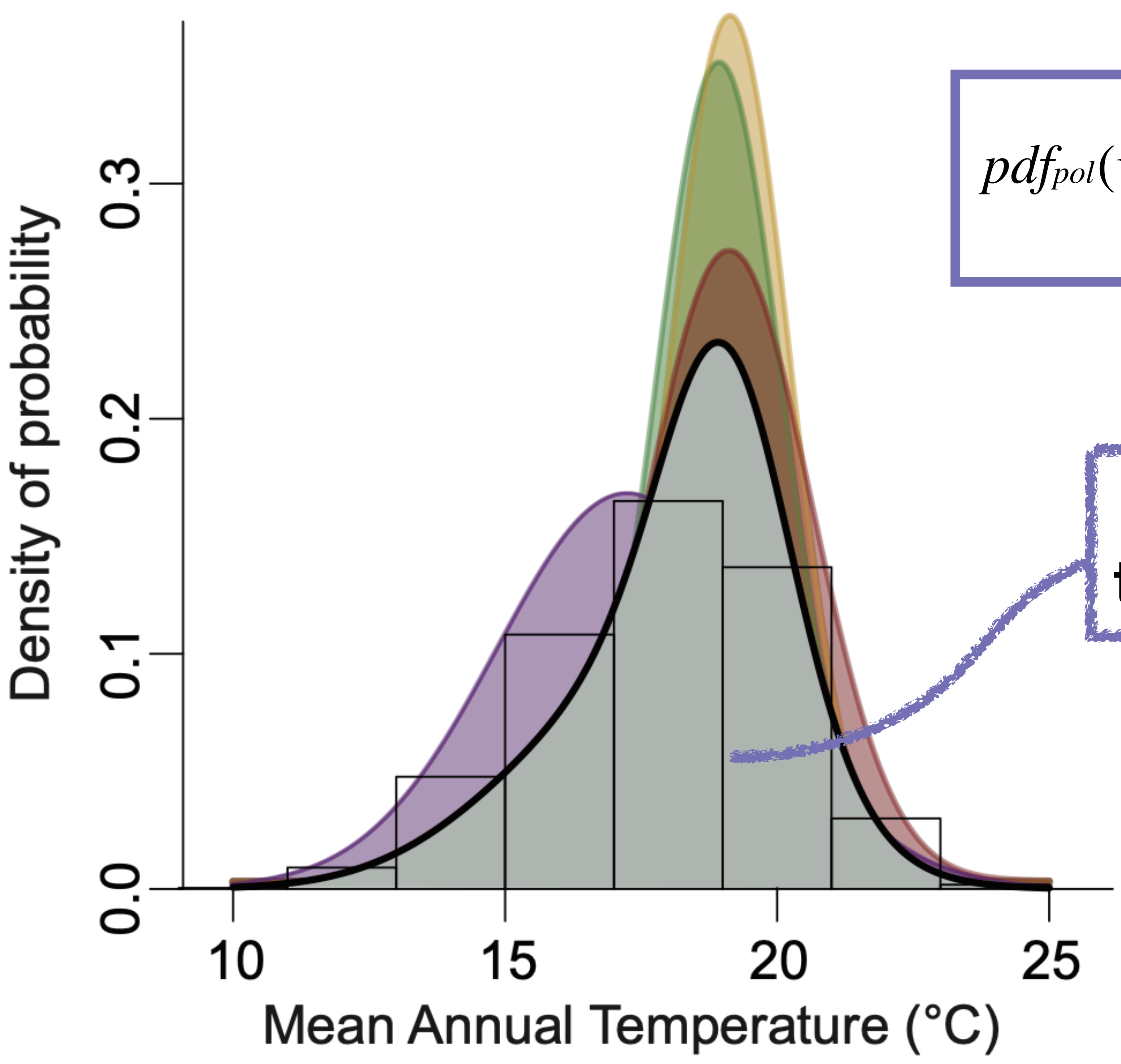
Note: its shape can be irregular, multimodal, skewed, etc.

16. INFERRING A CLIMATE RESPONSE FROM PLANT DISTRIBUTIONS

A **Fitting of the species *pdfs***



B **Fitting of the pollen *pdfs***



$$pdf_{pol}(v) = \frac{1}{\sum_{spi}\sqrt{n_{spi}}} \sum_{sp1}^{spN} \sqrt{n_{spi}}\, pdf_{spi}(v)$$

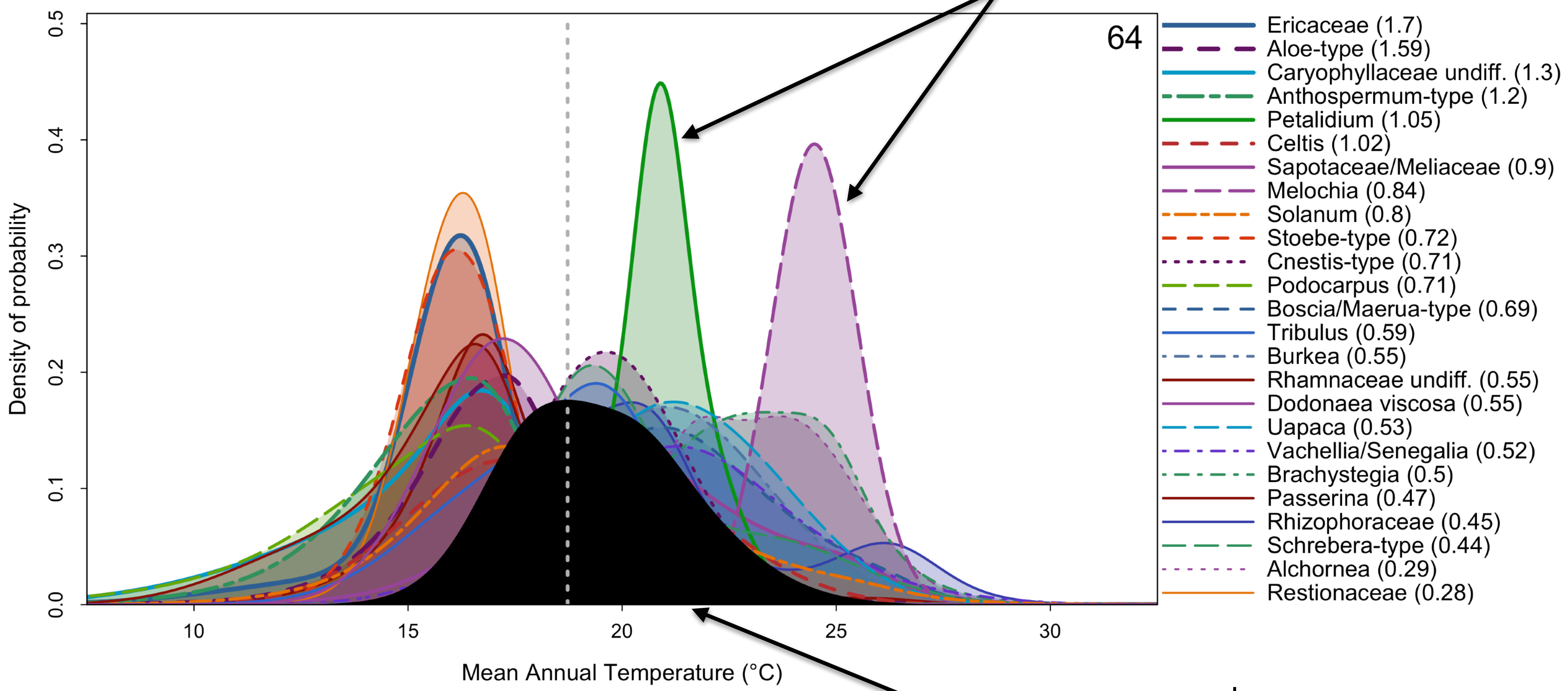
This is the unit used to reconstruct climate

This represents the full ‘breath’ of the pollen taxon.

Note: its shape can be irregular, multimodal, skewed, etc.

17. RECONSTRUCTING CLIMATE

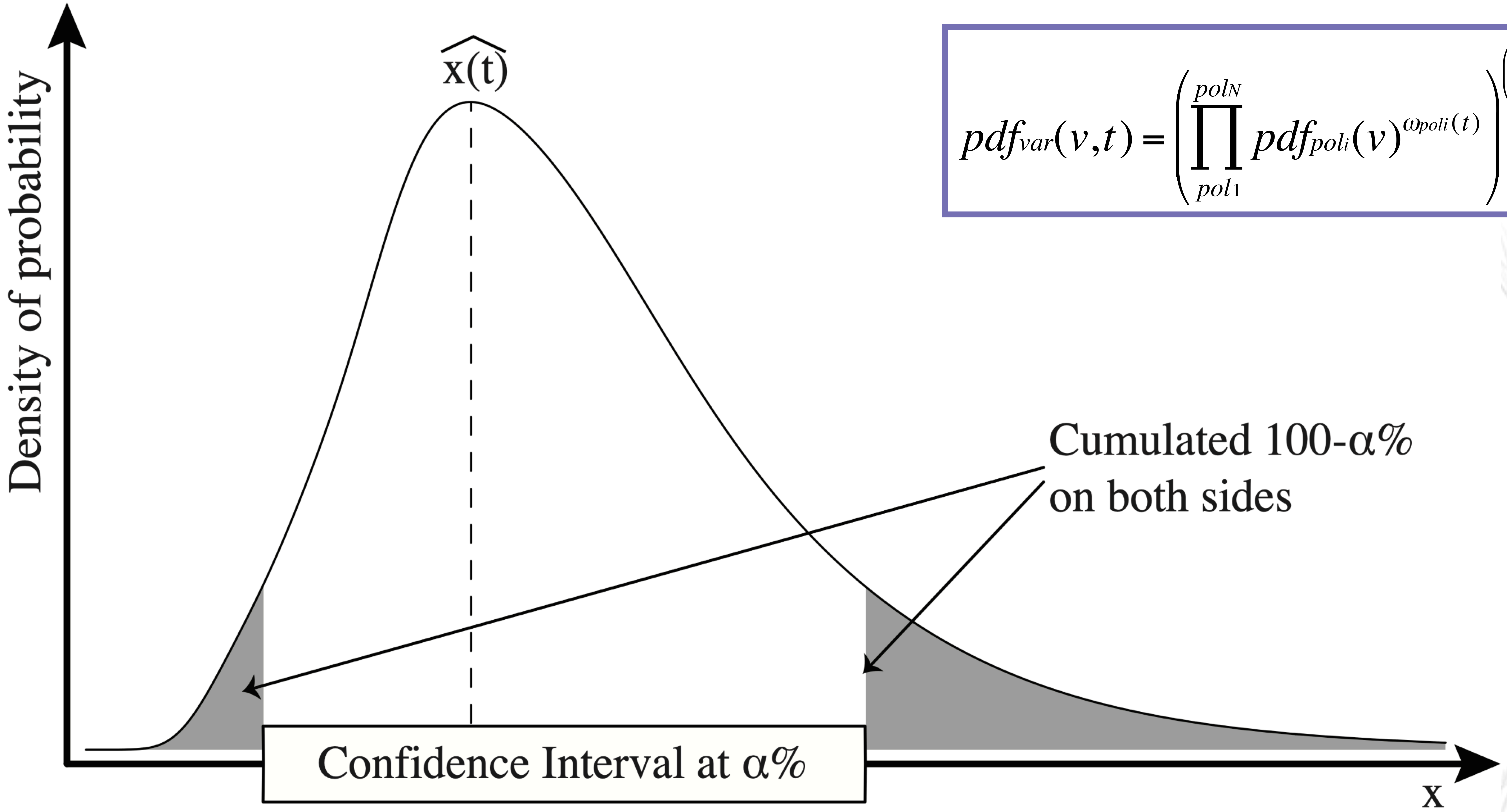
Pollen responses (defined from the previous slide)



Climate reconstruction

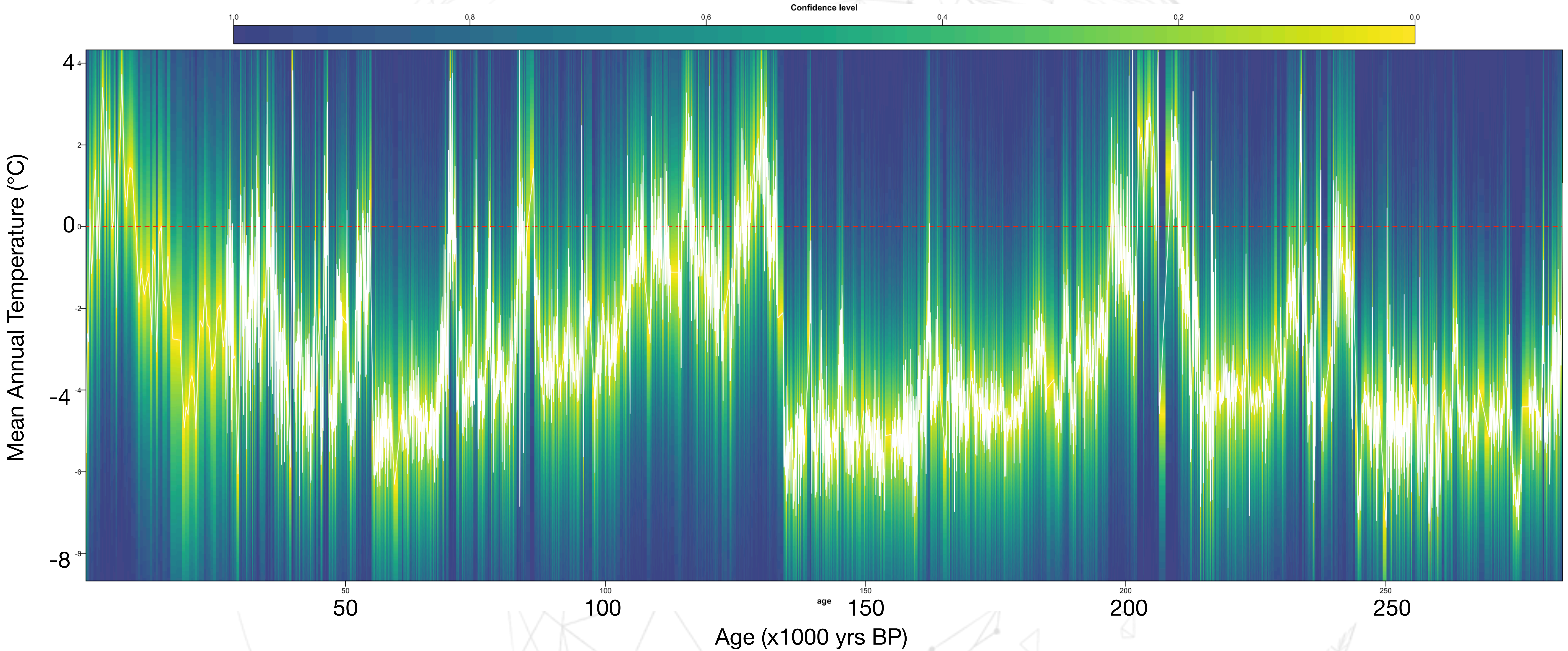
Most likely value ~18°C
95% range: ~15-25°C (note the asymmetry)

18. A WEIGHTED ENSEMBLE OF PLAUSIBLE CLIMATE VALUES

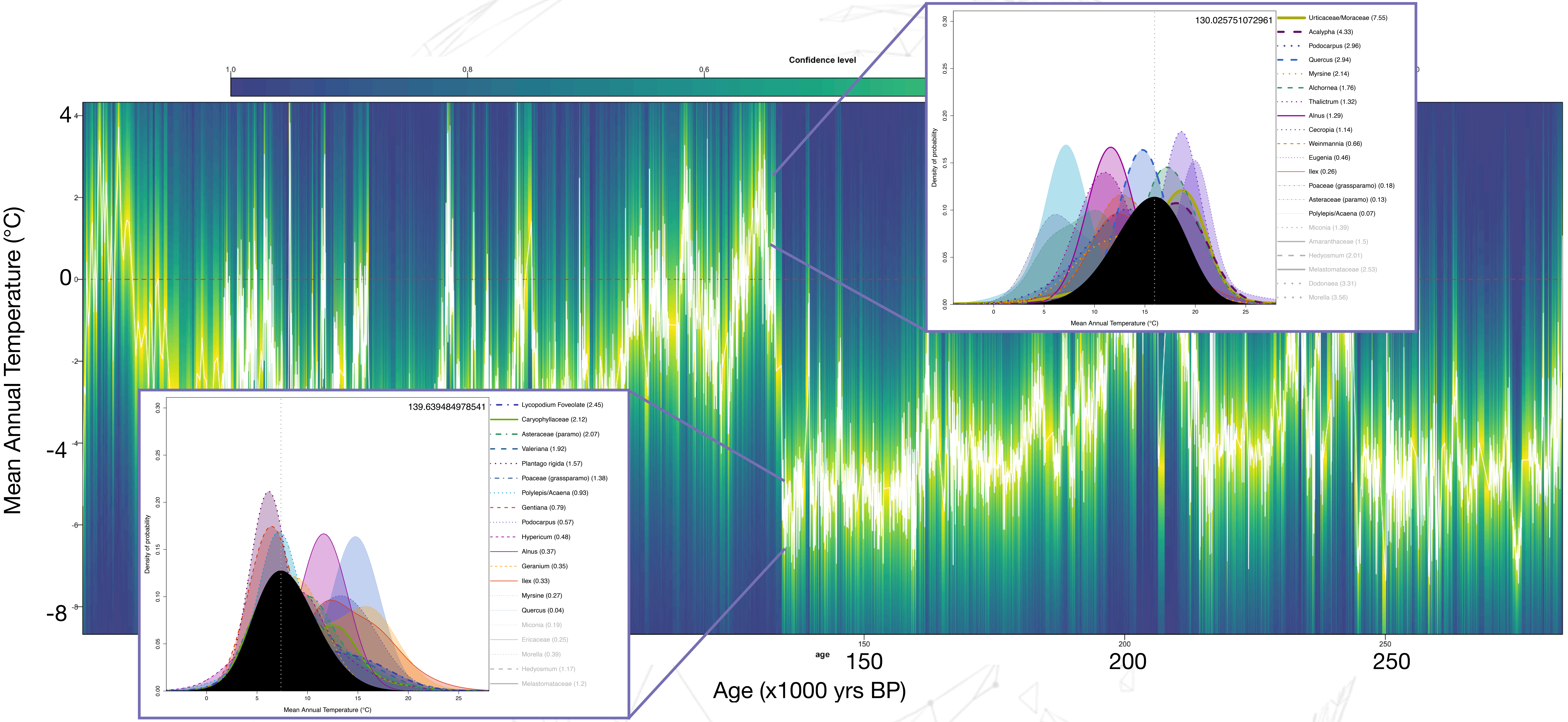


$$pdf_{var}(v,t) = \left(\prod_{pol_1}^{pol_N} pdf_{pol_i}(v)^{\omega_{pol_i}(t)} \right)^{\left(\sum_{pol} \omega_{pol_i}(t)\right)^{-1}}$$

19. A Weighted Ensemble of Plausible Climate Values



19. A Weighted Ensemble of Plausible Climate Values



20. Take Home Messages

- We want to account for all possible climate values, not just the most likely ones.
- The taxon-climate relationships take the species composition into account.
- Modern observations are debiased for the 'unbalanced' climate distributions.
- The uncertainties are not comparable with other methods (different modelling objectives, hence different uncertainties).
- Generating numbers is easy. Generating meaningful numbers is not (next part and the rest of today…).

20. TAKE HOME MESSAGES

- We want to account for all possible climate values not just the most likely ones.
- The taxon-climate relationships take the species composition into account.
- Modern observations are debiased for the ‘unbalanced’ climate distributions.
- The uncertainties are not comparable with other methods (different modelling objectives, hence different uncertainties).
- **Generating numbers is easy. Generating meaningful numbers is not (next part and the rest of today…).**

How to generate meaningful reconstructions?

Most concepts are also transposable to other reconstruction methods.

22. Where Things Get Complicated: How to Produce Meaningful Reconstructions?

Rule #1: The data define which variable(s) you can reconstruct. You do not get to decide.

There can be 1, 2, and occasionally more. But it can also be 0!
AND it may change over time (otherwise it would be too easy).

Think about the climate you study, annual vs seasonal aspects, and consider contrasting the dominating season (e.g. 'Winter rainfall') with the non-dominating season (e.g. Can changes in summer rainfall amounts, even if small(er) in amplitude, make a bigger difference for the studied vegetation? —> reduction of the annual drought stress)

23. WHERE THINGS GET COMPLICATED: HOW TO PRODUCE MEANINGFUL RECONSTRUCTIONS?

Rule #2: it is an ITERATIVE process.

Rule #2: it is an ITERATIVE process.

Rule #2: it is an ITERATIVE process.

It is a high-dimensionality problem (*i.e.* many variables/aspects to consider).

So, do not aim to get it right in the first place: be curious, try different modelling options, look at the data from different angles, and try to observe the data without any preconceived ideas.

24. WHERE THINGS GET COMPLICATED: HOW TO PRODUCE MEANINGFUL RECONSTRUCTIONS?

Rule #3: Be critical at each step of the analysis.

This is not a point-and-click analysis where you input data and get results (e.g. linear model). You will have to investigate the data to extract the information (e.g. why is my LGM reconstruction warm? Which taxa might be driving that unusual signal? Does this make sense? (go back to the ecology of your taxa!)

25. THE MAIN QUESTIONS YOU WILL HAVE TO ANSWER

- **Is this calibration zone/climate space big enough to analyse my data?**
- **Is this taxon that 'warms' my LGM actually sensitive to temperature?**
- **Is my reconstruction robust?**

This is where the experience (i.e. the accumulation of MANY iterations) comes into action.
But the package proposes different outputs to facilitate the process.

Table of Content

27. IMPORTANT DISTINCTION

No prior coding knowledge is required to use the package. You can use the package even if you know little about coding.

28. WHERE TO START?

https://www.manuelchevalier.com/crestr/

crestr 1.3.0.9000 | How it works | Get started | Prepare your data | Calibration data | Reference | Community | Vignettes | News

crestr An R package to perform probabilistic palaeoclimate reconstructions from palaeoecological datasets

The goal of `crestr` is to produce probabilistic reconstructions of past climate change from fossil assemblage data (Chevalier, 2022). The approach is based on the estimation of responses of studies bio-proxy studied to climate parameters using probability density functions (*PDFs*; see Chevalier *et al.* (2014) and Chevalier (2019)). The theory underpinning this package is explained in section *A bit of theory* and is illustrated with an application based on pseudo-data in section *Get Started*. The different vignettes present different aspects of the structure of the package and the data it contains, along with applications based on real data.

> *NOTE:* If you notice any bug, or if you would like to see some specific functions implemented, you can contact me at chevalier.manuel@gmail.com.

Installation

The package is currently unavailable from CRAN. However, you can install the latest development version (1.3.0) from GitHub with:

```
if(!require(devtools)) install.packages("devtools")
devtools::install_github("mchevalier2/crestr")
```

Links
View on CRAN
Browse source code
Report a bug

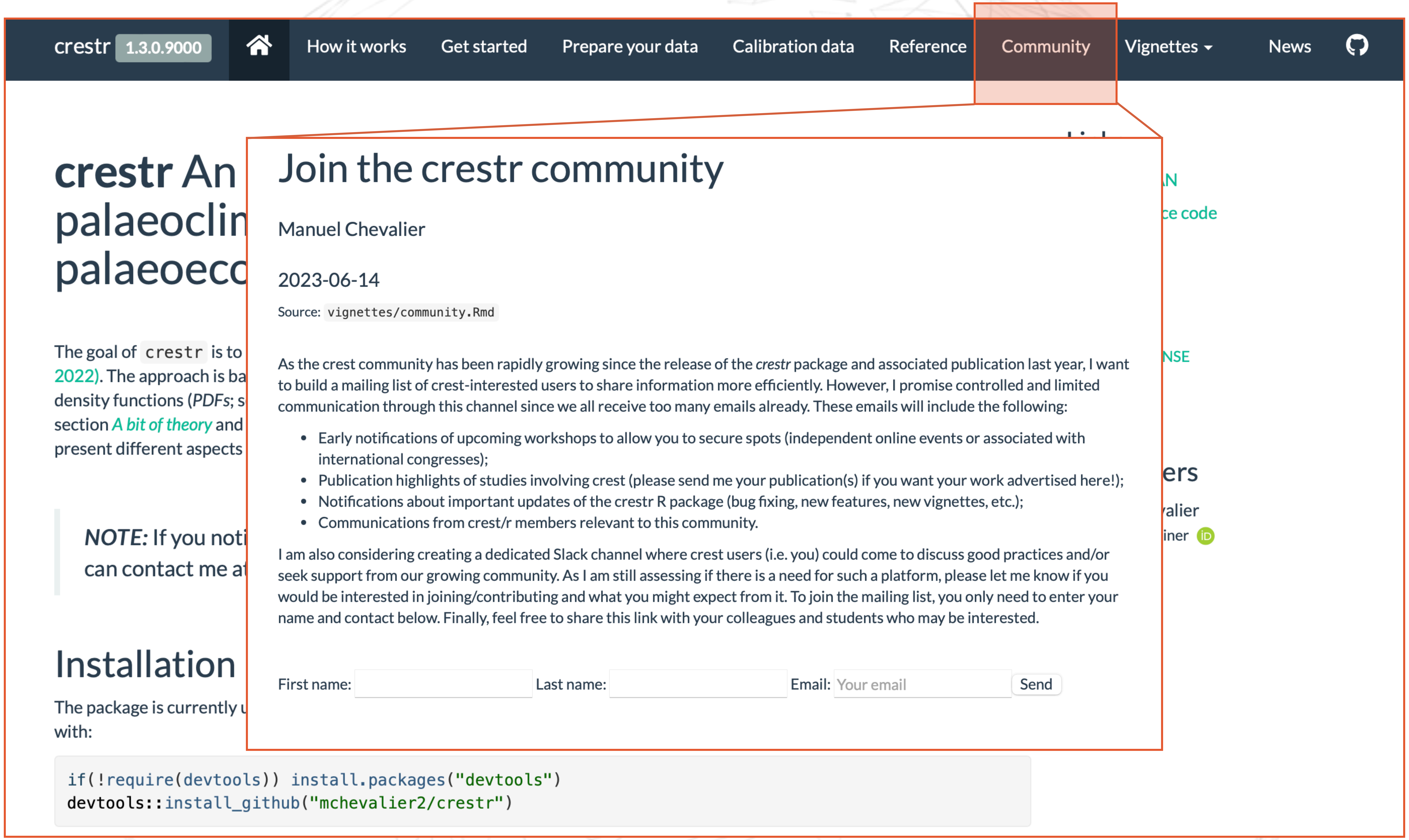
License
Full license
MIT + file LICENSE

Citation
Citing crestr

Developers
Manuel Chevalier
Author, maintainer

28. WHERE TO START?

https://www.manuelchevalier.com/crestr/


crestr 1.3.0.9000 | How it works | Get started | Prepare your data | Calibration data | Reference | Community | Vignettes | News

Join the crestr community

Manuel Chevalier

2023-06-14

Source: `vignettes/community.Rmd`

As the crest community has been rapidly growing since the release of the *crestr* package and associated publication last year, I want to build a mailing list of crest-interested users to share information more efficiently. However, I promise controlled and limited communication through this channel since we all receive too many emails already. These emails will include the following:

- Early notifications of upcoming workshops to allow you to secure spots (independent online events or associated with international congresses);
- Publication highlights of studies involving crest (please send me your publication(s) if you want your work advertised here!);
- Notifications about important updates of the crestr R package (bug fixing, new features, new vignettes, etc.);
- Communications from crest/r members relevant to this community.

I am also considering creating a dedicated Slack channel where crest users (i.e. you) could come to discuss good practices and/or seek support from our growing community. As I am still assessing if there is a need for such a platform, please let me know if you would be interested in joining/contributing and what you might expect from it. To join the mailing list, you only need to enter your name and contact below. Finally, feel free to share this link with your colleagues and students who may be interested.

First name: Last name: Email: Your email Send

crestr An palaeoclim palaeoeco

The goal of `crestr` is to 2022). The approach is ba density functions (*PDFs*; s section *A bit of theory* and present different aspects

NOTE: If you noti can contact me at

Installation

The package is currently u with:

```
if(!require(devtools)) install.packages("devtools")
devtools::install_github("mchevalier2/crestr")
```

The calibration dataset

30. A PACKAGE FOR MANY PROXIES!

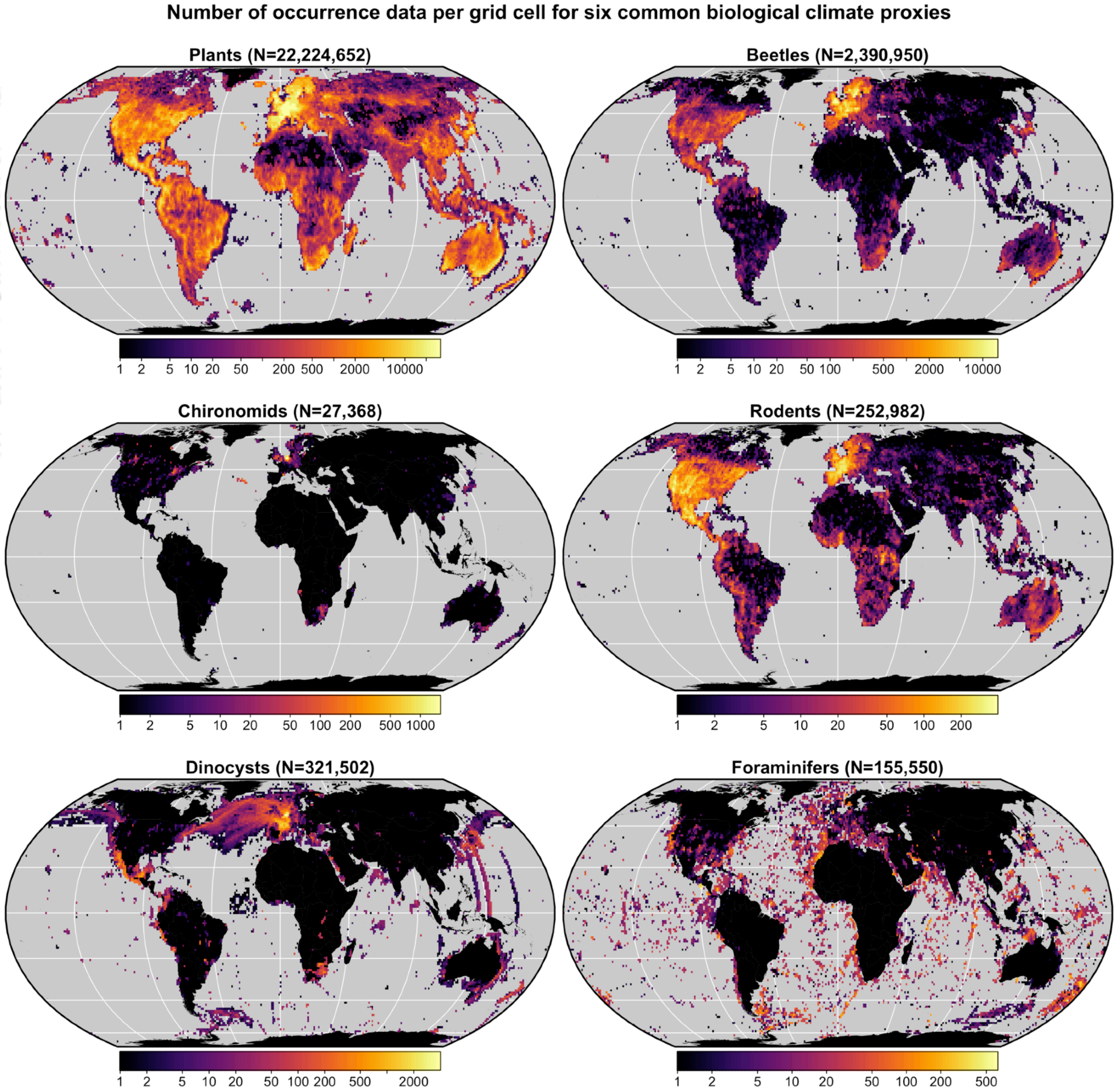
A 'global' calibration dataset for 6 palaeoecological proxies

- 4 terrestrial proxies (plants, rodents, chironomids, beetles)
- 2 marine proxies (forams, diatoms)

GBIF Global Biodiversity Information Facility

www.gbif.org

The package can be used without any additional ressource (except your fossil data to analyse, of course)



31. THE CLIMATE VARIABLES AVAILABLE

Terrestrial variables

Code	Full name	Source
bio1	Mean annual temp. (°C)	Fick and Hijmans (2017)
bio2	Mean diurnal range (°C)	Fick and Hijmans (2017)
bio3	Isothermality (×100) (unitless)	Fick and Hijmans (2017)
bio4	Temp. seasonality (standard deviation ×100) (°C)	Fick and Hijmans (2017)
bio5	Max temp. of the warmest month (°C)	Fick and Hijmans (2017)
bio6	Min temp. of the coldest month (°C)	Fick and Hijmans (2017)
bio7	Temp. annual range (°C)	Fick and Hijmans (2017)
bio8	Mean temp. of the wettest quarter (°C)	Fick and Hijmans (2017)
bio9	Mean temp. of the driest quarter (°C)	Fick and Hijmans (2017)
bio10	Mean temp. of the warmest quarter (°C)	Fick and Hijmans (2017)
bio11	Mean temp. of the coldest quarter (°C)	Fick and Hijmans (2017)
bio12	Annual precip. (mm)	Fick and Hijmans (2017)
bio13	Precip. of the wettest month (mm)	Fick and Hijmans (2017)
bio14	Precip. of the driest month (mm)	Fick and Hijmans (2017)
bio15	Precip. seasonality (coefficient of variation) (mm)	Fick and Hijmans (2017)
bio16	Precip. of the wettest quarter (mm)	Fick and Hijmans (2017)
bio17	Precip. of the driest quarter (mm)	Fick and Hijmans (2017)
bio18	Precip. of the warmest quarter (mm)	Fick and Hijmans (2017)
bio19	Precip. of the coldest quarter (mm)	Fick and Hijmans (2017)
ai	Aridity index (unitless)	Zomer et al. (2008)

Oceanic variables

Code	Full name	Source
sst_ann	Mean annual SST (°C)	Locarnini et al. (2018)
sst_jfm	Mean winter SST (°C)	Locarnini et al. (2018)
sst_amj	Mean spring SST (°C)	Locarnini et al. (2018)
sst_jas	Mean summer SST (°C)	Locarnini et al. (2018)
sst_ond	Mean fall SST (°C)	Locarnini et al. (2018)
sss_ann	Mean annual SSS (PSU)	Zweng et al. (2018)
sss_jfm	Mean winter SSS (PSU)	Zweng et al. (2018)
sss_amj	Mean spring SSS (PSU)	Zweng et al. (2018)
sss_jas	Mean summer SSS (PSU)	Zweng et al. (2018)
sss_ond	Mean fall SSS (PSU)	Zweng et al. (2018)
icec_ann	Mean annual sea ice concentration (%)	Reynolds et al. (2007)
icec_jfm	Mean winter sea ice concentration (%)	Reynolds et al. (2007)
icec_amj	Mean spring sea ice concentration (%)	Reynolds et al. (2007)
icec_jas	Mean summer ice concentration (%)	Reynolds et al. (2007)
icec_ond	Mean fall sea ice concentration (%)	Reynolds et al. (2007)
diss_oxy	Dissolved oxygen concentration (μmol L^{-1})	Garcia et al. (2018a)
nitrate	Nitrate concentration (μmol L^{-1})	Garcia et al. (2018b)
phosphate	Phosphate concentration (μmol L^{-1})	Garcia et al. (2018b)
silicate	Silicate concentration (μmol L^{-1})	Garcia et al. (2018b)

32. A PACKAGE FOR MANY PROXIES!

We can also discuss the possibility of adding another proxy or environmental variable to the calibration data.

Come talk to me at the break/end!

33. A PACKAGE FOR MANY PROXIES!

But you do not have to use 'my' data.
You can also use your own calibration data.

Structure of the data in R

35. THE OBJECT IN THE R ENVIRONMENT

The package is based on calls to functions and the repeated use/update of one single nested (or rather nasty?) R object.

recon.obj <— crestr.function (recon.obj, …)

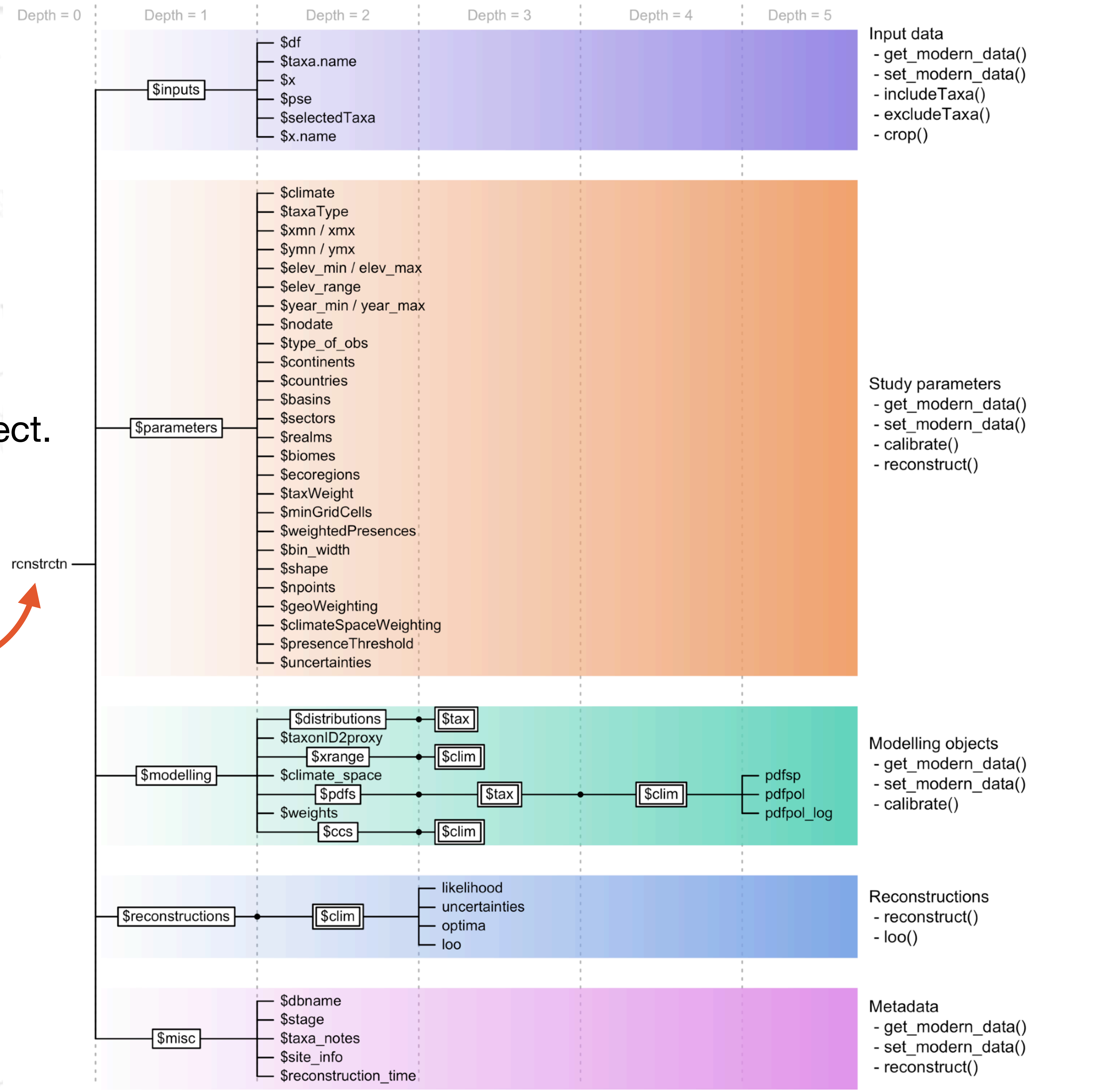
The mechanics of how everything is made is hidden to the users.

35. THE OBJECT IN THE R ENVIRONMENT

The package is based on calls to functions and the repeated use/update of one single nested (or rather nasty?) R object.

recon.obj <— crestr.function (recon.obj, …)

The mechanics of how everything is made is hidden to the users.



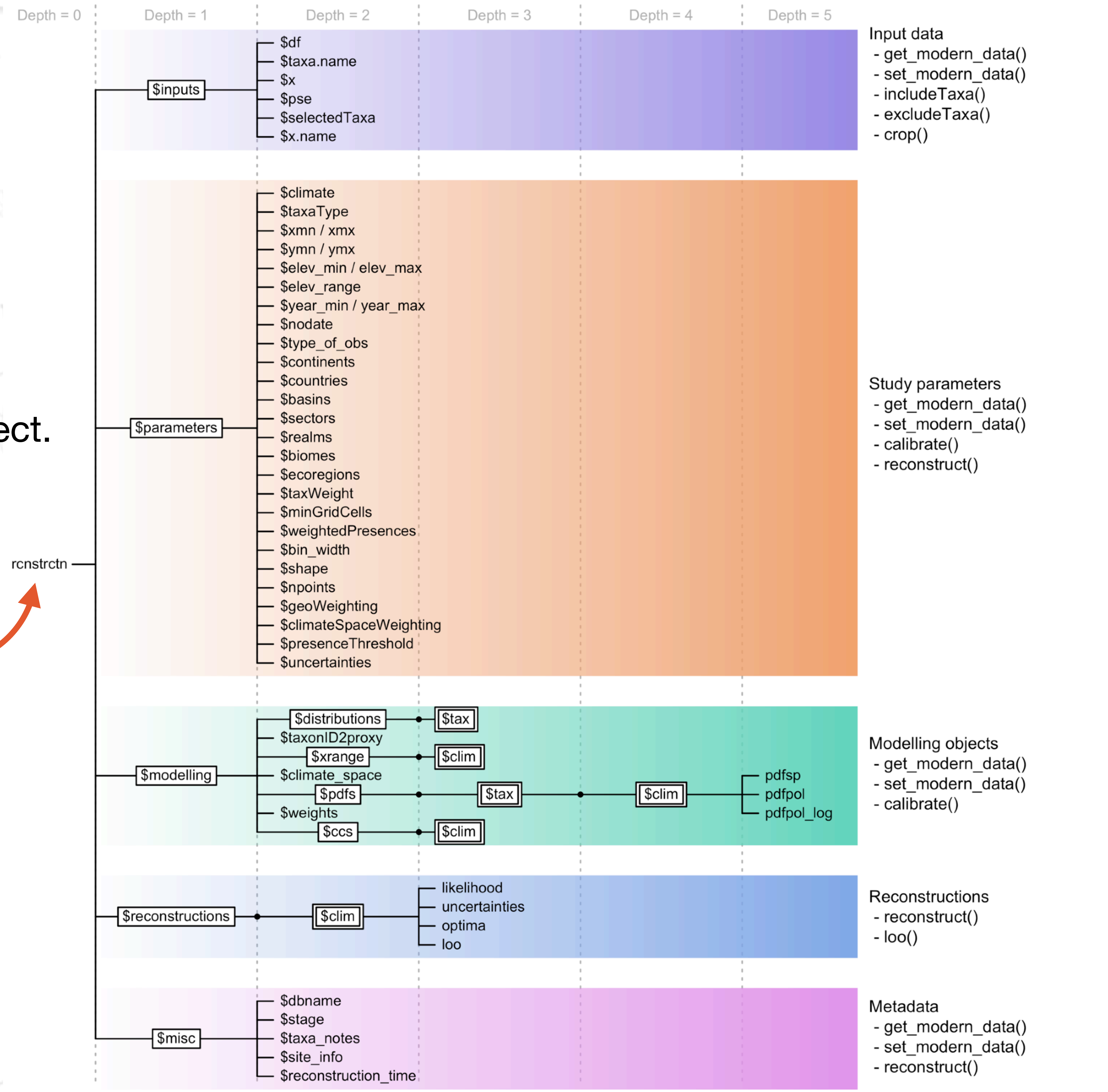
35. THE OBJECT IN THE R ENVIRONMENT

The package is based on calls to functions and the repeated use/update of one single nested (or rather nasty?) R object.

recon.obj <— crestr.function (recon.obj, ...)

The mechanics of how everything is made is hidden to the users.

First initialised by crest.get_modern_data() or crest.set_modern_Data()

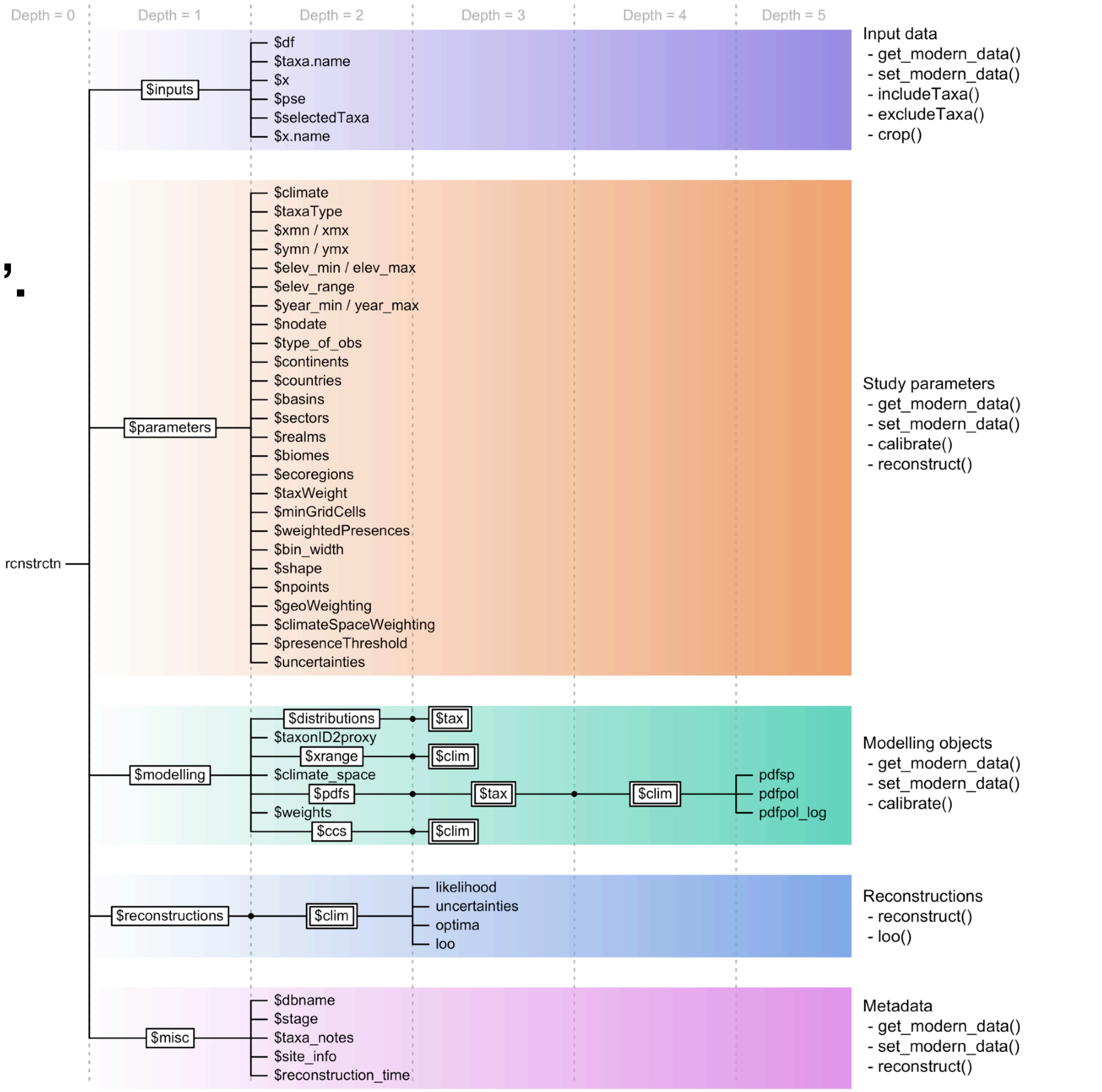


36. THE OBJECT IN THE R ENVIRONMENT

You will never have to modify this object 'by hand'.

You can, but I recommend using the dedicated functions that will be introduced in the last part.

recon.obj <— crestr.function (recon.obj, …)



36. THE OBJECT IN THE R ENVIRONMENT

You will never have to modify this object 'by hand'.

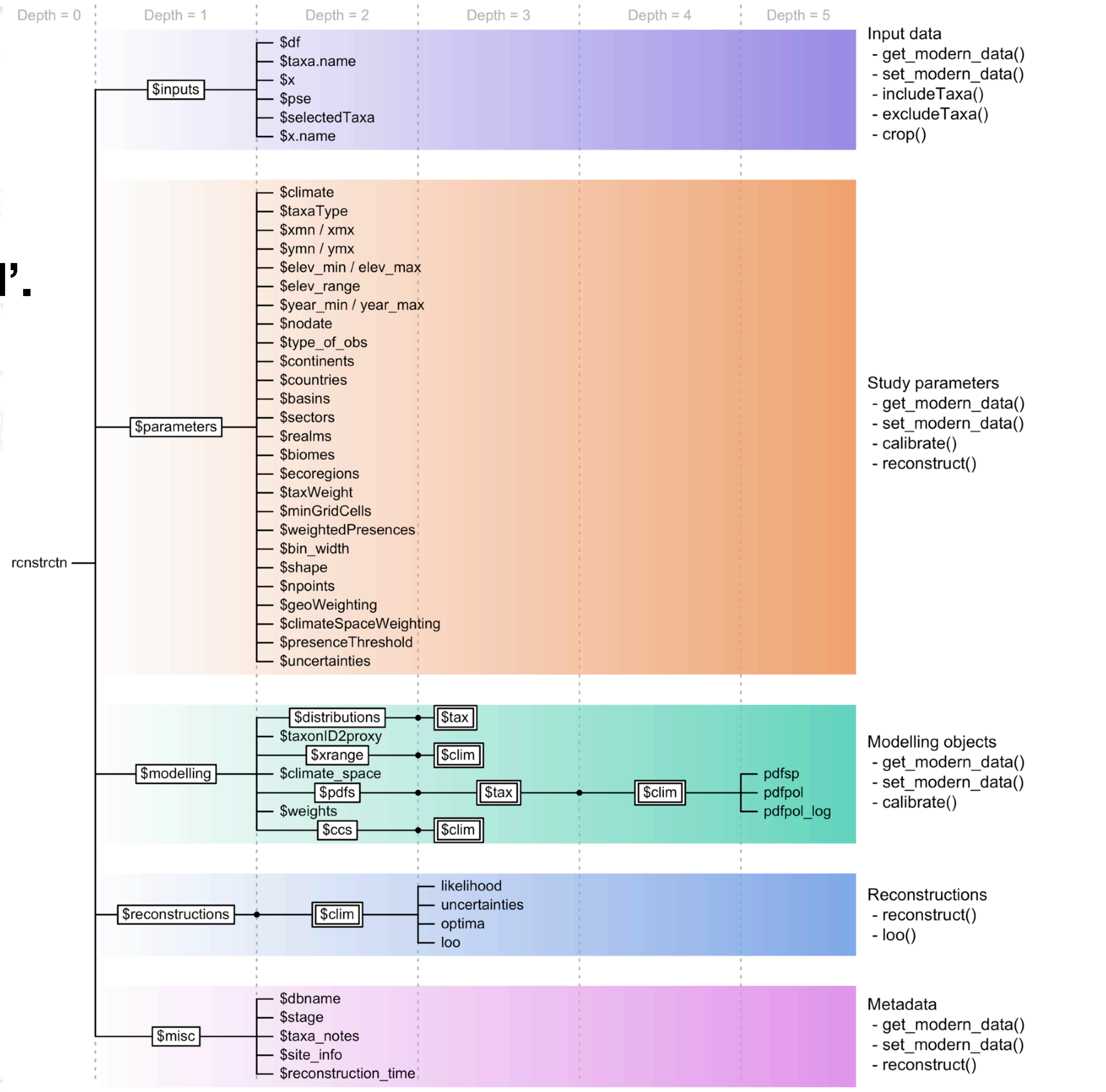
You can, but I recommend using the dedicated functions that will be introduced in the last part.

recon.obj <— crestr.function (recon.obj, …)

The updated object

What you want to do

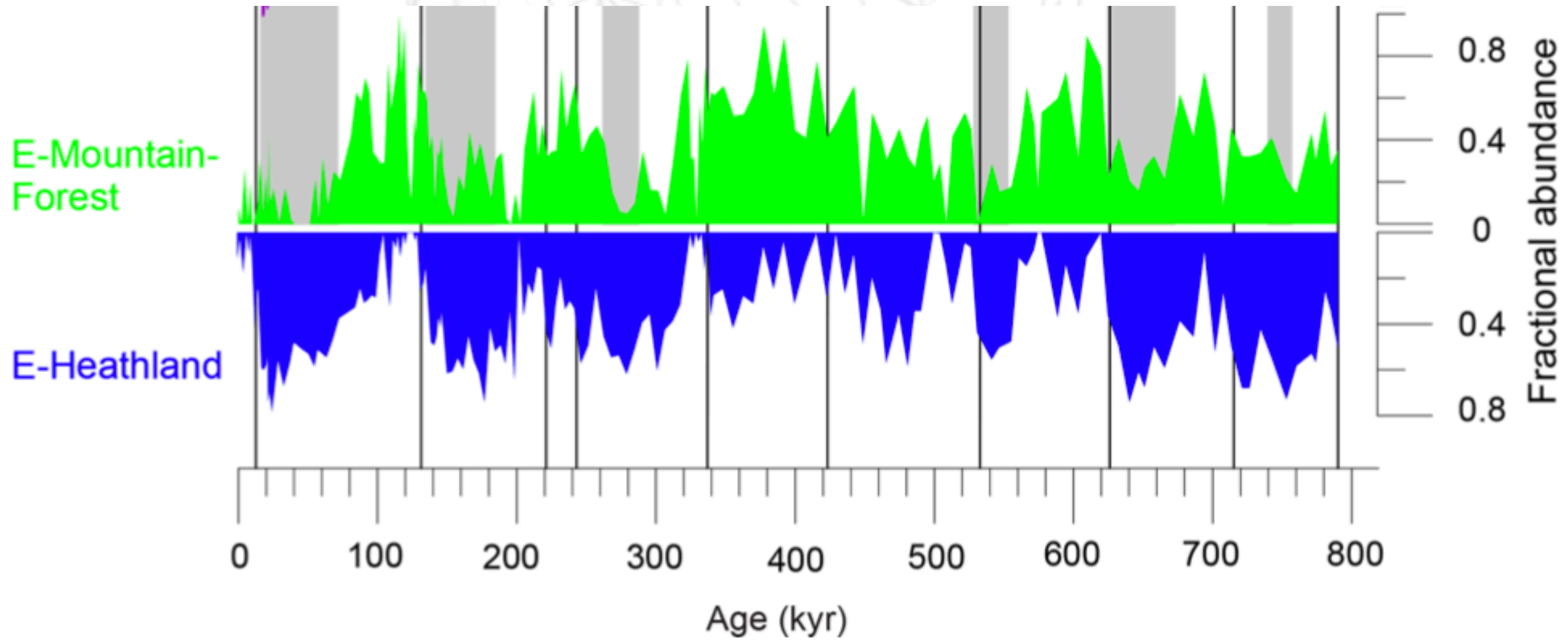
The current object

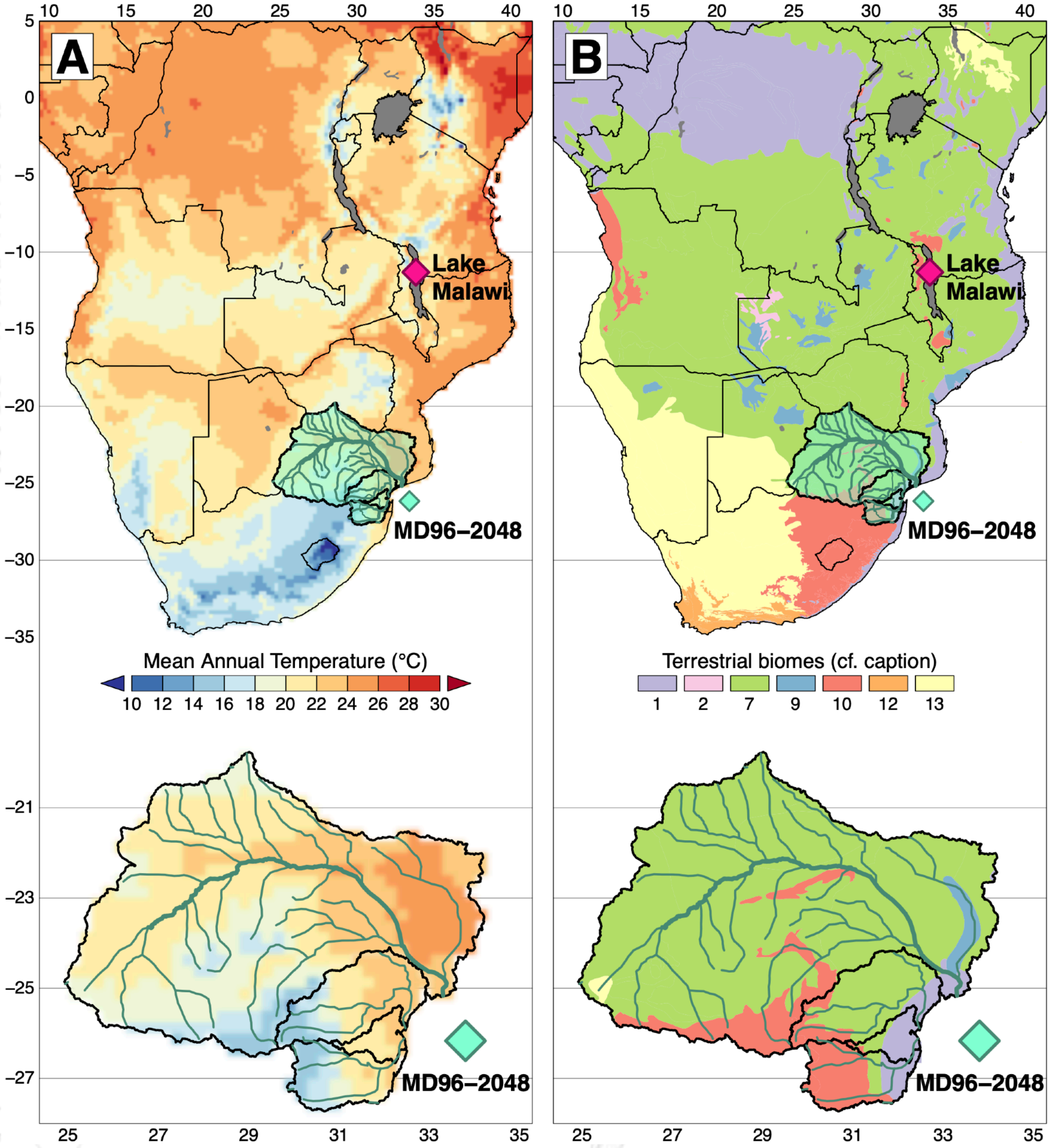


Preparing your data for crestr

38. Pollen Record from Marine Core MD96-2048

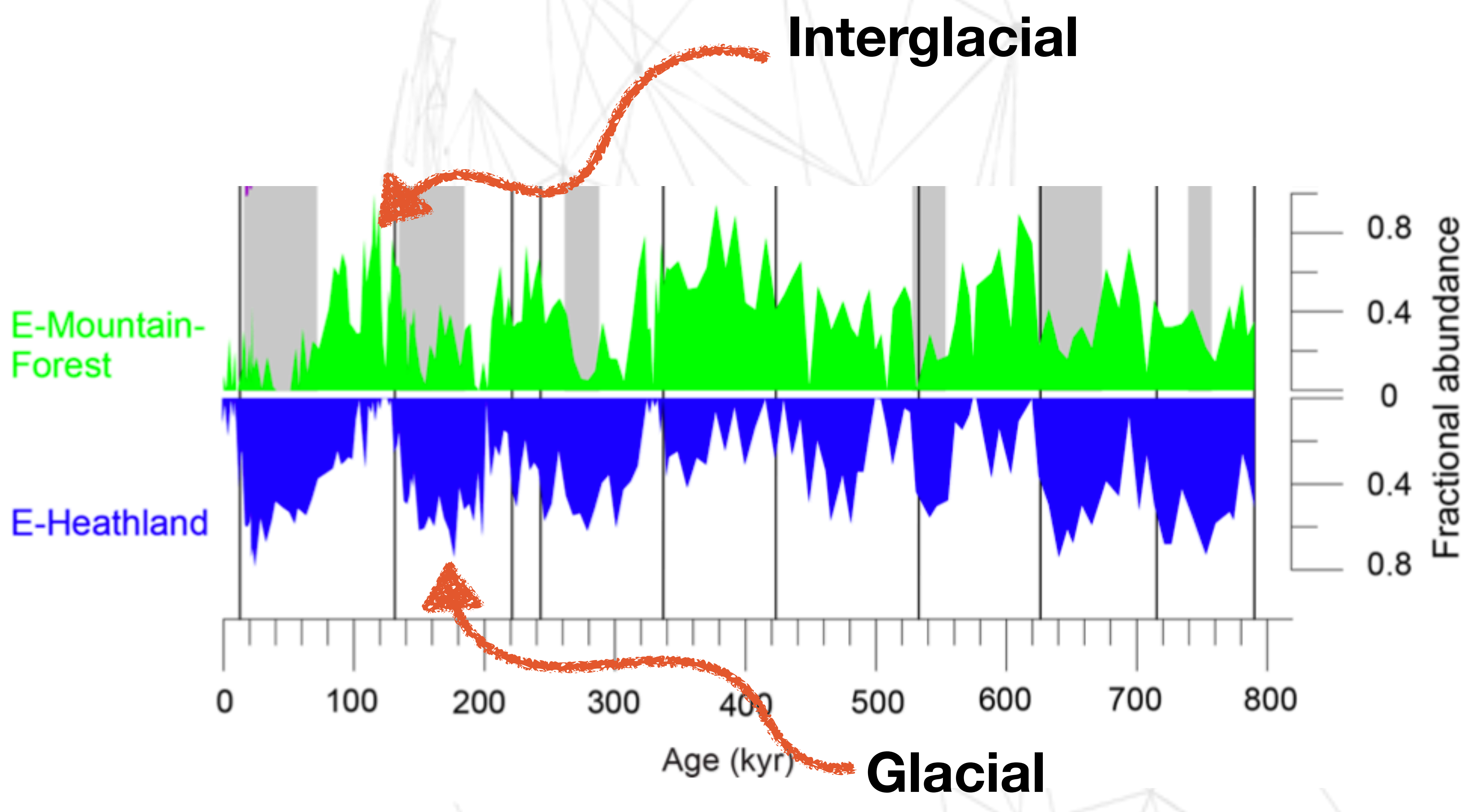
- 790,000 years long
- 181 pollen samples
- 150 terrestrial taxa
- Marine core, hence huge catchment
- Strong glacial-interglacial temperature signal

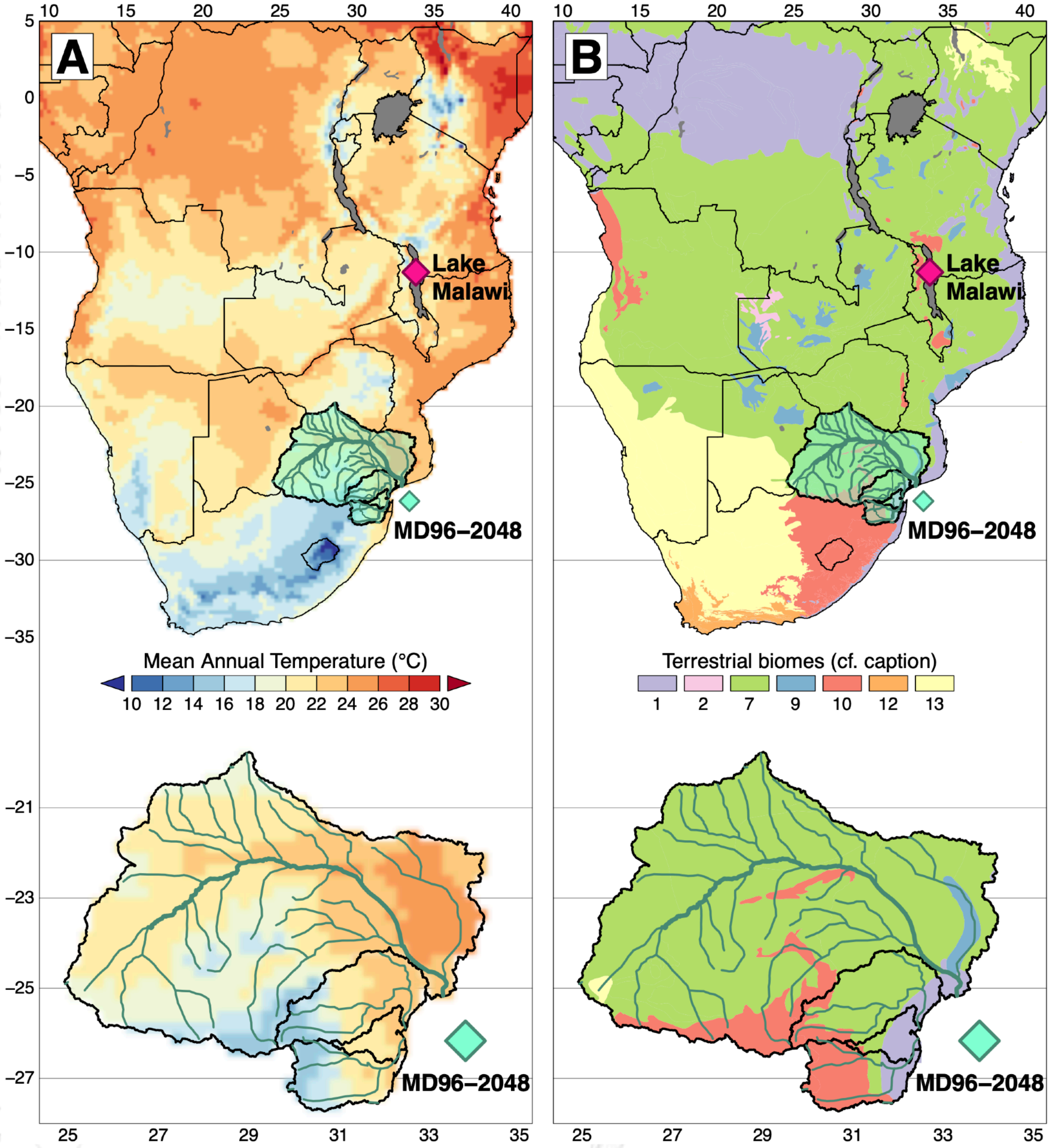




38. POLLEN RECORD FROM MARINE CORE MD96-2048

- 790,000 years long
- 181 pollen samples
- 150 terrestrial taxa
- Marine core, hence huge catchment
- Strong glacial-interglacial temperature signal





39. NATURE OF THE INPUT FILES

Possible input files

- A species-proxy equivalency table (necessary to use the GBIF / Worldclim data)
- The fossil data to inform the reconstructions
- Alternative calibration data if you do not want to use the GBIF / Worldclim data

40. STRUCTURE OF THE INPUT FILES (1/3)

The input fossil data

A data frame with:

- The first column contains an age, a depth, a sample ID (only of one of these)
- Then, one column per taxon (only the taxa to be used in the reconstruction)
- The first row contains the taxa names
- The data can be many things (percentages, 1/0s, weights)

	Age	Vachellia/Senegalia	Acalypha	Acanthaceae undiff.	Adenia	Afraegle	Afzelia
1	0	1.0101010	0	0.5050505	0	0	0
2	2	0.0000000	0	0.0000000	0	0	0
3	4	0.0000000	0	0.0000000	0	0	0
4	7	1.0000000	0	0.0000000	0	0	0
5	9	0.0000000	0	0.0000000	0	0	0
6	9	0.0000000	0	0.0000000	0	0	0
7	11	0.0000000	0	0.0000000	0	0	0
8	13	0.0000000	0	0.0000000	0	0	0
9	14	0.0000000	0	0.0000000	0	0	0
10	15	0.4504505	0	0.0000000	0	0	0

41. STRUCTURE OF THE INPUT FILES (2/3)

The 'Proxy-Species' equivalency (PSE) table

(Tedious, but critical)

A data frame with five columns:

- Level: the level of identification (Family, Genus, Species)
- Family, Genus, Species
- ProxyName

Level	Family	Genus	Species	*ProxyName*	Taxonomic resolution
1	Asteraceae			Asteraceae undiff.	Family
2	Asteraceae	*Stoebe*		*Stoebe*-type	Subfamily
2	Asteraceae	*Elytropappus*		*Stoebe*-type	Subfamily
2	Asteraceae	*Artemisia*		*Artemisia*	Genus level
3	Arecaceae	*Elaeis*	*Elaeis guineensis*	*Elaeis guineensis*	Species
4				Triletes spores	To be excluded

```
createPSE( colnames(input_data) )
```

42. Structure of the Input Files (3/3)

The 'Distributions' table (alternative calibration dataset)

Species name	Taxon name	Longitude	Latitude	Weight*	clim_1	...	clim_n
Stoebe plumosa	*Stoebe*-type	18.875	−34.375	20	15.8	...	711
Elytropappus rhinocerotis	*Stoebe*-type	18.375	−33.625	32	16.9	...	477
...	...	...	...	...	...	...	...
Elaeis guineensis	*Elaeis guineensis*	−4.375	10.875	4	27.4	...	1020

TABLE OF CONTENT

44. OVER-LUNCH HOMEWORK: THINK ABOUT

What are the characteristics of a good calibration dataset?

How to determine if a taxon is a good climate indicator?

TABLE OF CONTENT

General structure of an analysis

47. BASIC FRAMEWORK

Formatting the fossil data	Previous slides
Extracting the modern calibration dataset	`crest.get_modern_data(...)`
Fitting the proxy-climate relationships	`crest.calibrate(...)`
Reconstructing climate	`crest.reconstruct(...)`

BASIC FRAMEWORK

Formatting the fossil data — Previous slides

Extracting the modern calibration dataset — `crest.get_modern_data(...)`

+ diagnostic tools

Fitting the proxy-climate relationships — `crest.calibrate(...)`

+ diagnostic tools

Reconstructing climate — `crest.reconstruct(...)`

+ diagnostic tools

47. BASIC FRAMEWORK

ITERATIONS

Formatting the fossil data — Previous slides

Extracting the modern calibration dataset — `crest.get_modern_data(...)`

+ diagnostic tools

Fitting the proxy-climate relationships — `crest.calibrate(...)`

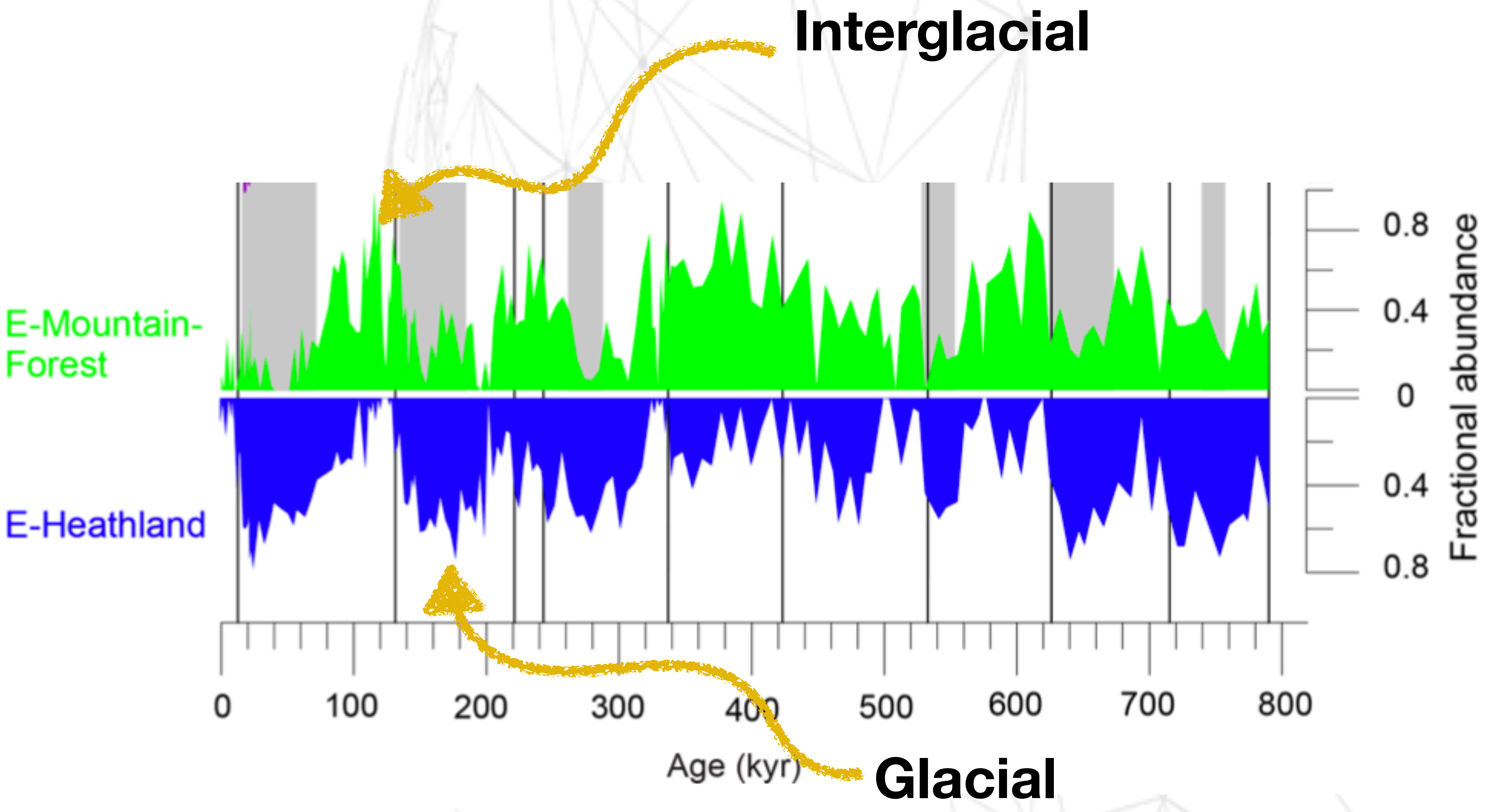
+ diagnostic tools

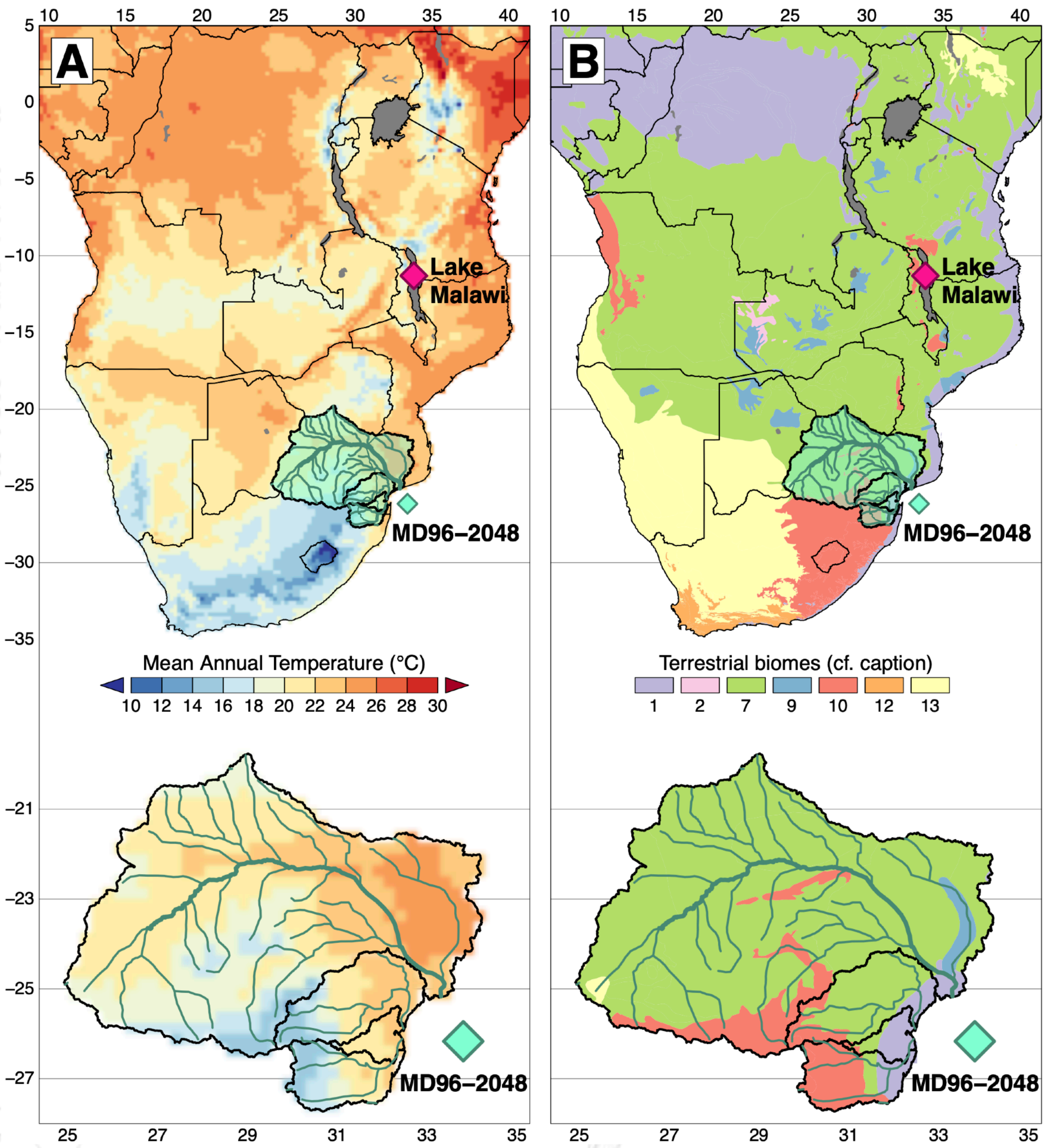
Reconstructing climate — `crest.reconstruct(...)`

+ diagnostic tools

48. Pollen Record from Marine Core MD96-2048

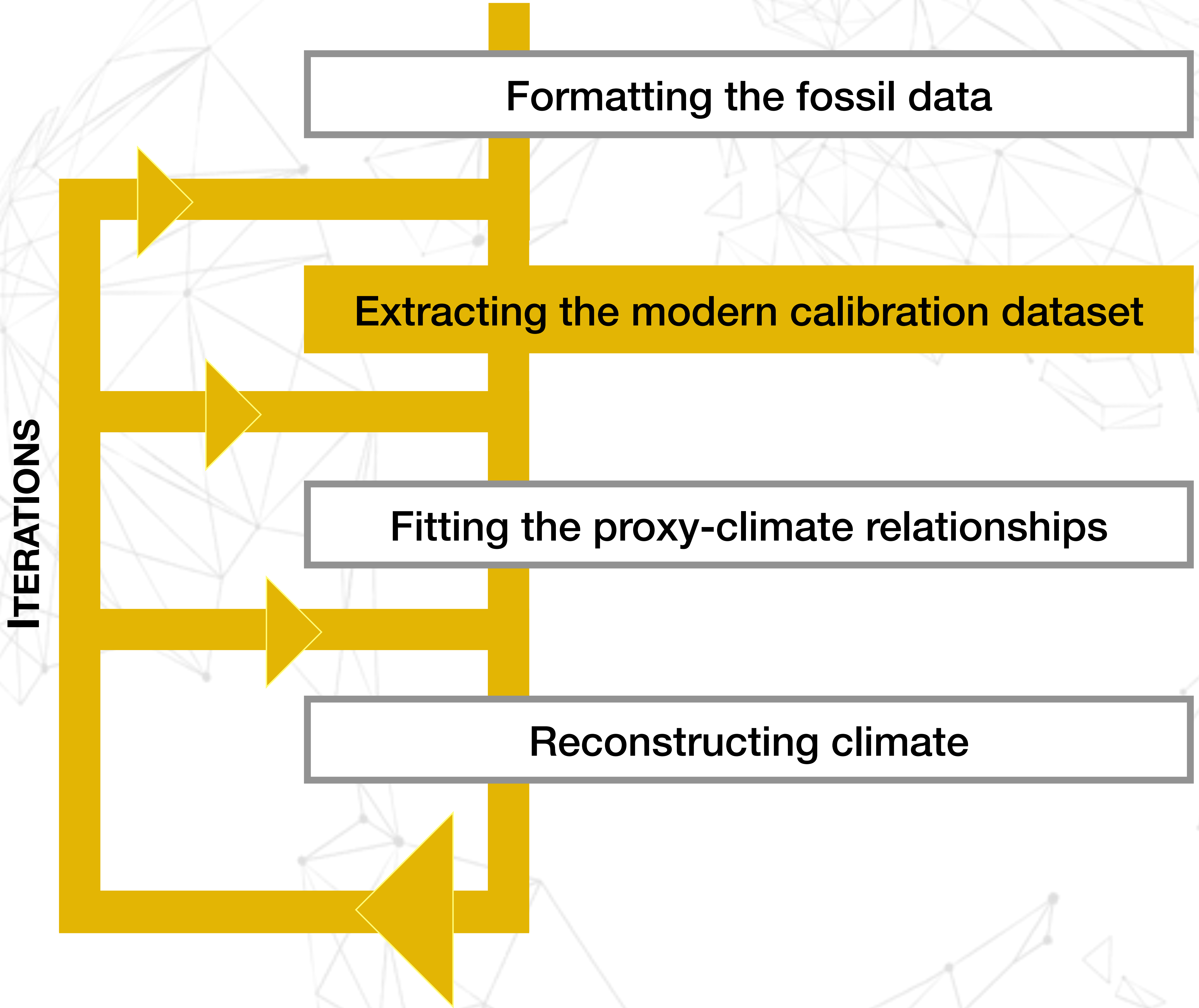
- 790,000 years long
- 181 pollen samples
- 150 terrestrial taxa
- Marine core, hence huge catchment
- Strong glacial-interglacial temperature signal





Let’s code! (Copy paste ;-))

50. BASIC FRAMEWORK



Previous slides

`crest.get_modern_data(...)`

+ diagnostic tools

`crest.calibrate(...)`

+ diagnostic tools

`crest.reconstruct(...)`

+ diagnostic tools

51. DEFINING A CALIBRATION DATASET

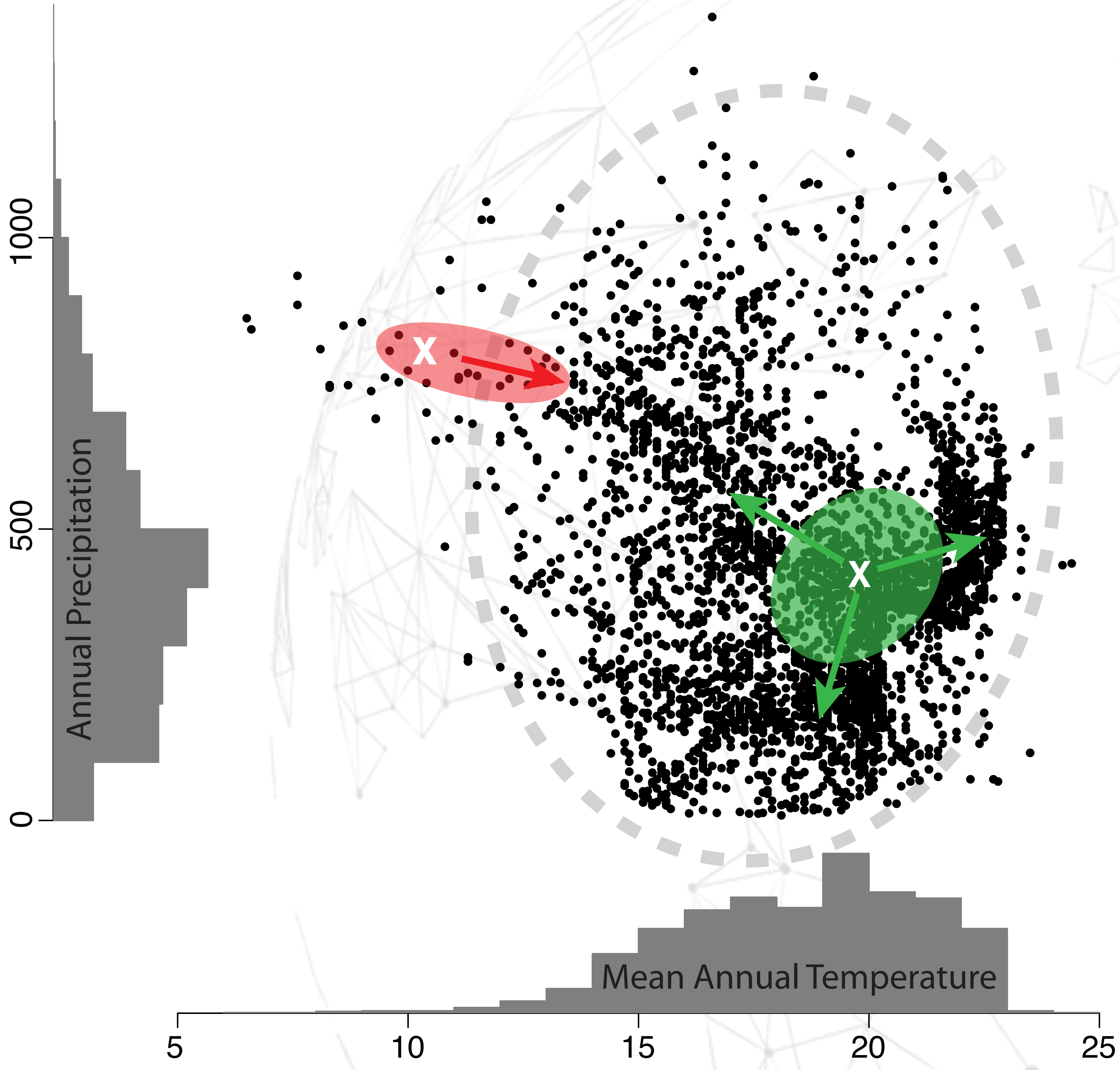
What are the characteristics of a good calibration dataset?

51 DEFINING A CALIBRATION DATASET

What are the characteristics of a good calibration dataset?

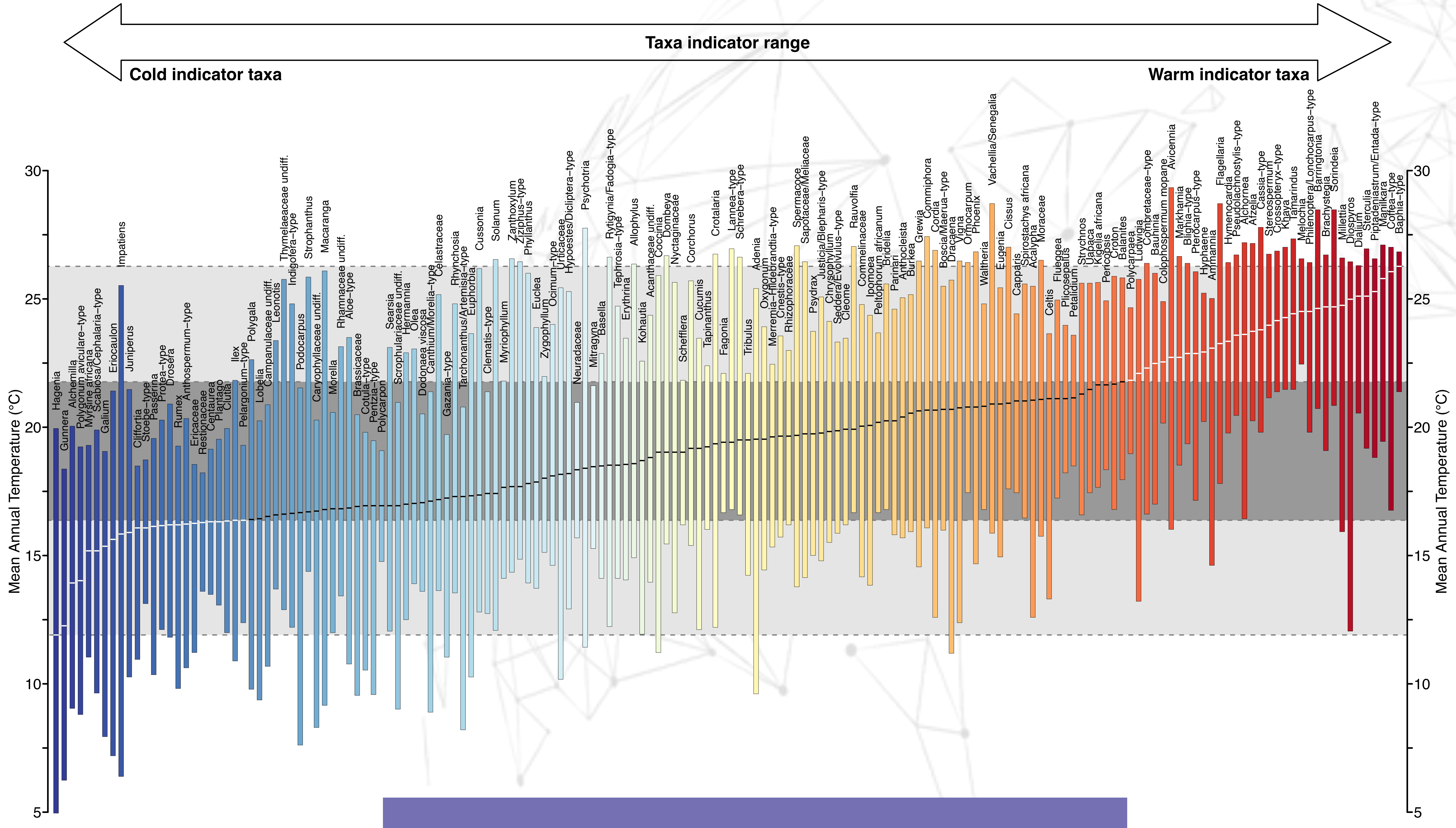
- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

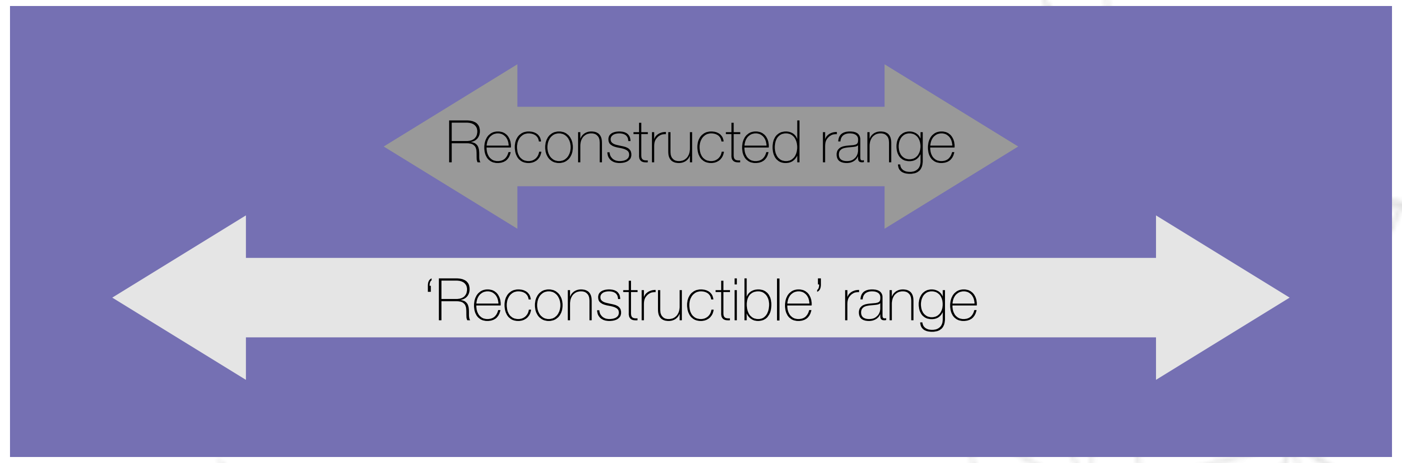
52. DEFINING A CALIBRATION DATASET



- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

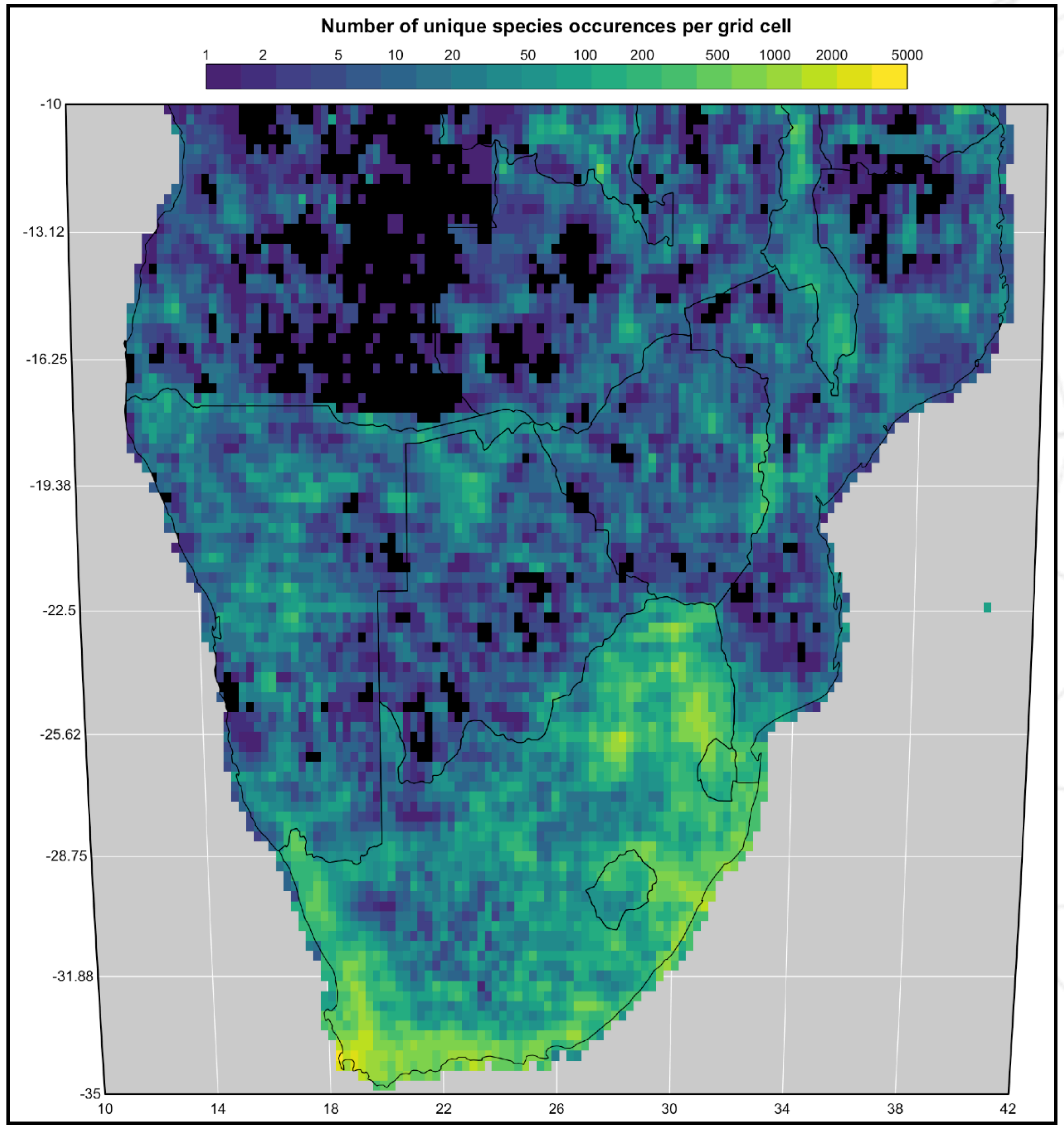
53. DEFINING A CALIBRATION DATASET





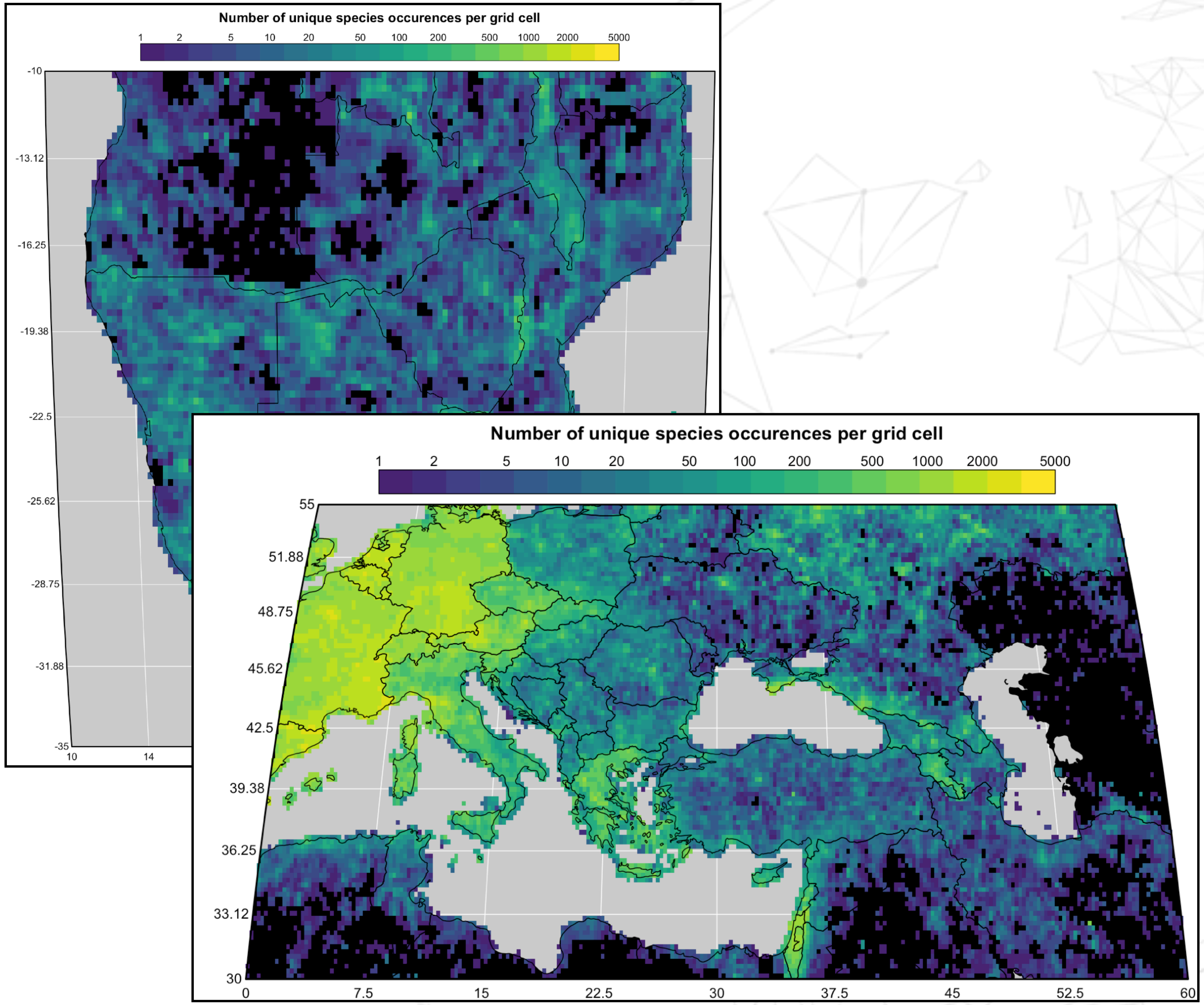
- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

54. Defining a Calibration Dataset



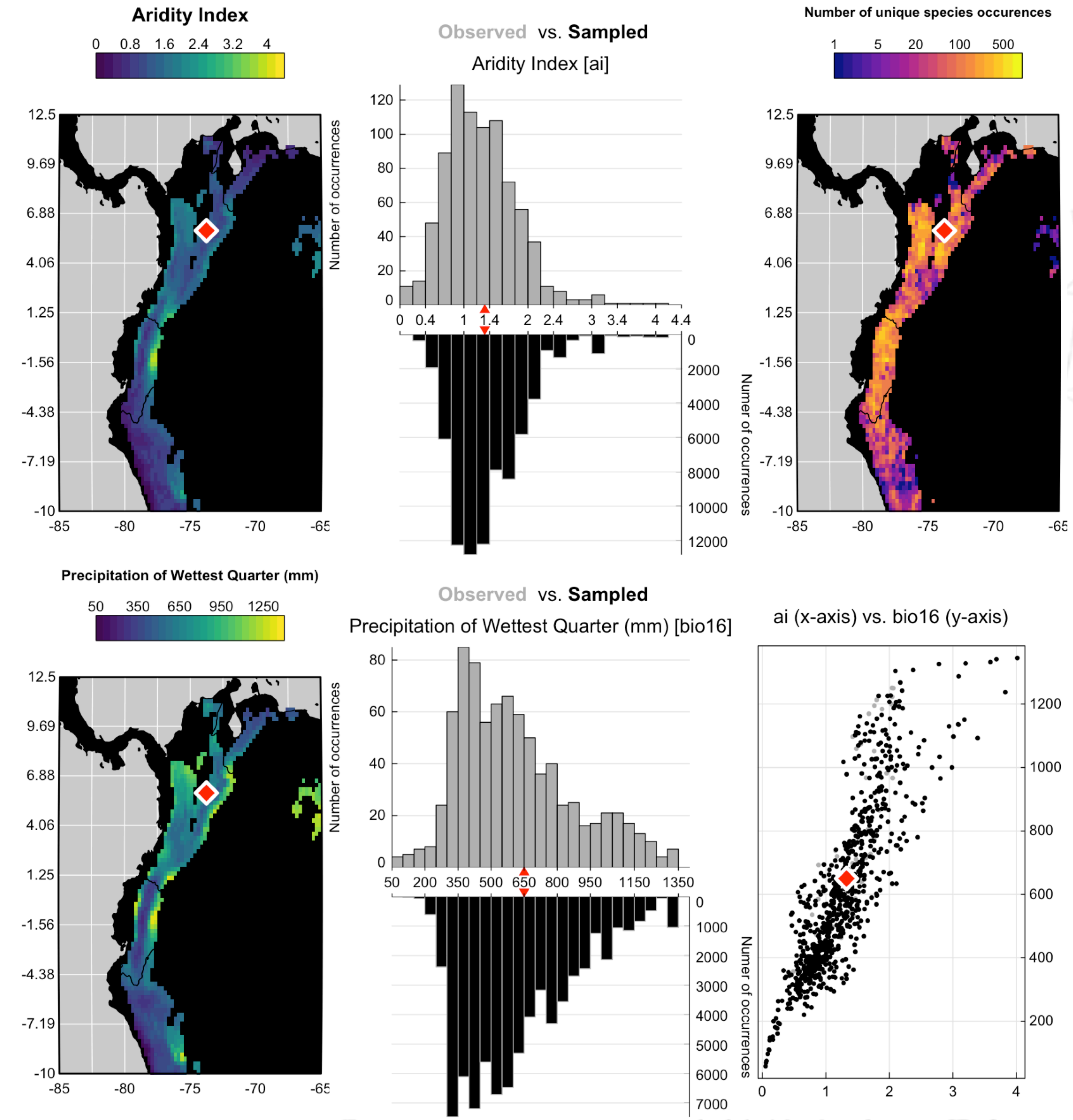
- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

54. DEFINING A CALIBRATION DATASET



- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

55. DEFINING A CALIBRATION DATASET



- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

56. DEFINING A CALIBRATION DATASET

What are the characteristics of a good calibration dataset?

- All the pollen taxa observed in the fossil record are well represented.
- The site(s) to reconstruct is(are) not on the edge of the climate space.
- The climate to reconstruct (unknown) is well represented. [can be checked with the reconstructions]
- There is no large sampling bias in the vegetation/climate spaces.
- If reconstructing more than one variable, ensuring the absence of correlation.
- The selected variables drove the observed fossil data.

57. EXTRACTING THE CALIBRATION DATASET

Input data:

- The pollen data 'MD96-2048'
- The PSE file 'PSE'

Parameters:

- Reconstructing mean annual temperature (bio1) and aridity index (ai)
- East and southern Africa (Angola and Congo excluded due to limited data)
- No subdivision by biome/ecoregion
- 20 distinct occurrences to estimate a pdf

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
              'Lesotho', 'eSwatini', 'Botswana',
              'Mozambique', 'Zimbabwe', 'Zambia',
              'Malawi', 'Tanzania', 'Namibia',
              'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = "gbif4crest_02",
        verbose = TRUE
)
```

57. EXTRACTING THE CALIBRATION DATASET

Input data:

- The pollen data 'MD96-2048'
- The PSE file 'PSE'

Parameters:

- Reconstructing mean annual temperature (bio1) and aridity index (ai)
- East and southern Africa (Angola and Congo excluded due to limited data)
- No subdivision by biome/ecoregion
- 20 distinct occurrences to estimate a pdf

***crestr* accesses the online calibration data**

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
              'Lesotho', 'eSwatini', 'Botswana',
              'Mozambique', 'Zimbabwe', 'Zambia',
              'Malawi', 'Tanzania', 'Namibia',
              'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = "gbif4crest_02",
        verbose = TRUE
)
```

58. EXTRACTING THE CALIBRATION DATASET

Input data:
- The pollen data ‘MD96-2048’
- The PSE file ‘PSE’

Parameters:
- Reconstructing mean annual temperature (bio1) and aridity index (ai)
- East and southern Africa (Angola and Congo excluded due to limited data)
- No subdivision by biome/ecoregion
- 20 distinct occurrences to estimate a pdf

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
              'Lesotho', 'eSwatini', 'Botswana',
              'Mozambique', 'Zimbabwe', 'Zambia',
              'Malawi', 'Tanzania', 'Namibia',
              'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = “gbif4crest_02”,
        verbose = TRUE
)
```

***crestr* is a chatty package.**
Read the warnings!

59. EXTRACTING THE CALIBRATION DATASET

Always check the notes about the proxy-species associations!!

`PSE_log(rcnstrctn)`

- "One or more taxa were are not in the proxy-species equivalence table and have been ignored."
- "One or more taxa were are not in the selectedTaxa table. They have been added but are not selected for any variable."
- "One or more taxa were not associated with species."
- "The percentages of one or more taxa were always 0 and have been removed accordingly."
- "One or more taxa were are not recorded in the data file."
- "The classification of one or more taxa into species was not successful."
- "For one or more taxa, no species remained associated with the proxy name at the end of the classification."
- "No data were available within the study area for one or more taxa."
- "An insufficient amount of calibration data points was available within the study area for one or more taxa. Consider reducing 'minGridCells'."

60. ACCESSING ONLINE DATA IS SLOW

You can get a local copy of the calibration data to accelerate the whole process.

1. You can create a local snapshot of the online database using dbSubset().

```
dbSubset(
        taxaType = 1,
        xmn = -15, xmx = 55,
        ymn = -35, ymx = 35
)
```

2. You can download a local copy of the whole database using dbDownload().

```
dbDownload(
      filename = "gbif4crest.zip",
      lite = TRUE
)
```

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
              'Lesotho', 'eSwatini', 'Botswana',
              'Mozambique', 'Zimbabwe', 'Zambia',
              'Malawi', 'Tanzania', 'Namibia',
              'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = "/path/sqlite.file",
        verbose = TRUE
)
```

61. Accessing online data is SLOW

```
dbSubset(
        taxaType = 1,
        xmn = -15, xmx = 55,
        ymn = -35, ymx = 35
)
```

```
dbDownload(
        filename = "gbif4crest.zip",
        lite = TRUE
)
```

These two functions will create a .sqlite3 file that you can then use in every function with the dbname parameter.

You can even download the whole database and then use dbSubset() on it!

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
            'Lesotho', 'eSwatini', 'Botswana',
            'Mozambique', 'Zimbabwe', 'Zambia',
            'Malawi', 'Tanzania', 'Namibia',
            'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = "/path/sqlite.file",
        verbose = TRUE
)
```

61. Accessing online data is SLOW

```
dbSubset(
        taxaType = 1,
        xmn = -15, xmx = 55,
        ymn = -35, ymx = 35
)
```

```
dbDownload(
        filename = "gbif4crest.zip",
        lite = TRUE
)
```

```
rcnstrctn <- crest.get_modern_data(
        df = MD96_2048,
        pse = PSE,
        taxaType = 1,
        climate = c('bio1', 'ai'),
        xmn = NA, xmx = NA,
        ymn = -35, ymx = NA,
        continents = 'Africa',
        countries = c('South Africa', 'Kenya',
            'Lesotho', 'eSwatini', 'Botswana',
            'Mozambique', 'Zimbabwe', 'Zambia',
            'Malawi', 'Tanzania', 'Namibia',
            'Uganda', 'Rwanda', 'Burundi'),
        realms = NA,
        biomes = NA,
        ecoregions = NA,
        minGridCells = 20,
        site_info = c(34.0167, -26.1667),
        site_name = 'MD96-2048',
        dbname = "/path/sqlite.file",
        verbose = TRUE
)
```

These two functions will create a .sqlite3 file that you can then use in every function with the dbname parameter.

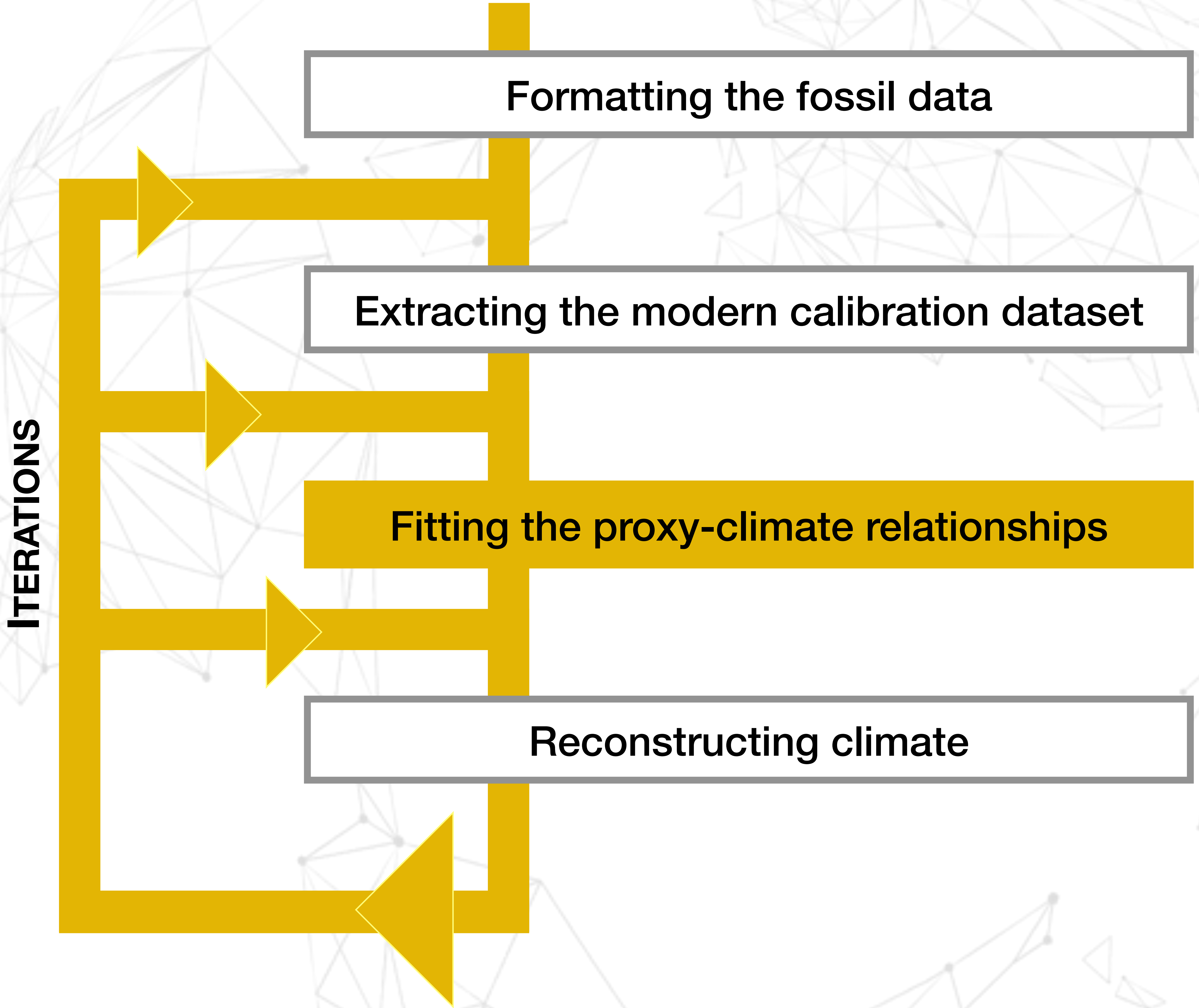
You can even download the whole database and then use dbSubset() on it!

62. ACCESSING ONLINE DATA IS SLOW AND ABOUT TO DISAPPEAR

Soon, the only available option will be that one:

```
dbDownload(
    filename = "gbif4crest.zip",
    lite = TRUE
)
```

63. Basic Framework



Previous slides

`crest.get_modern_data(...)`

+ diagnostic tools

`crest.calibrate(...)`

+ diagnostic tools

`crest.reconstruct(...)`

+ diagnostic tools

FITTING THE PROXY-CLIMATE RESPONSES

Input data:

- The crestObj created at the last stage

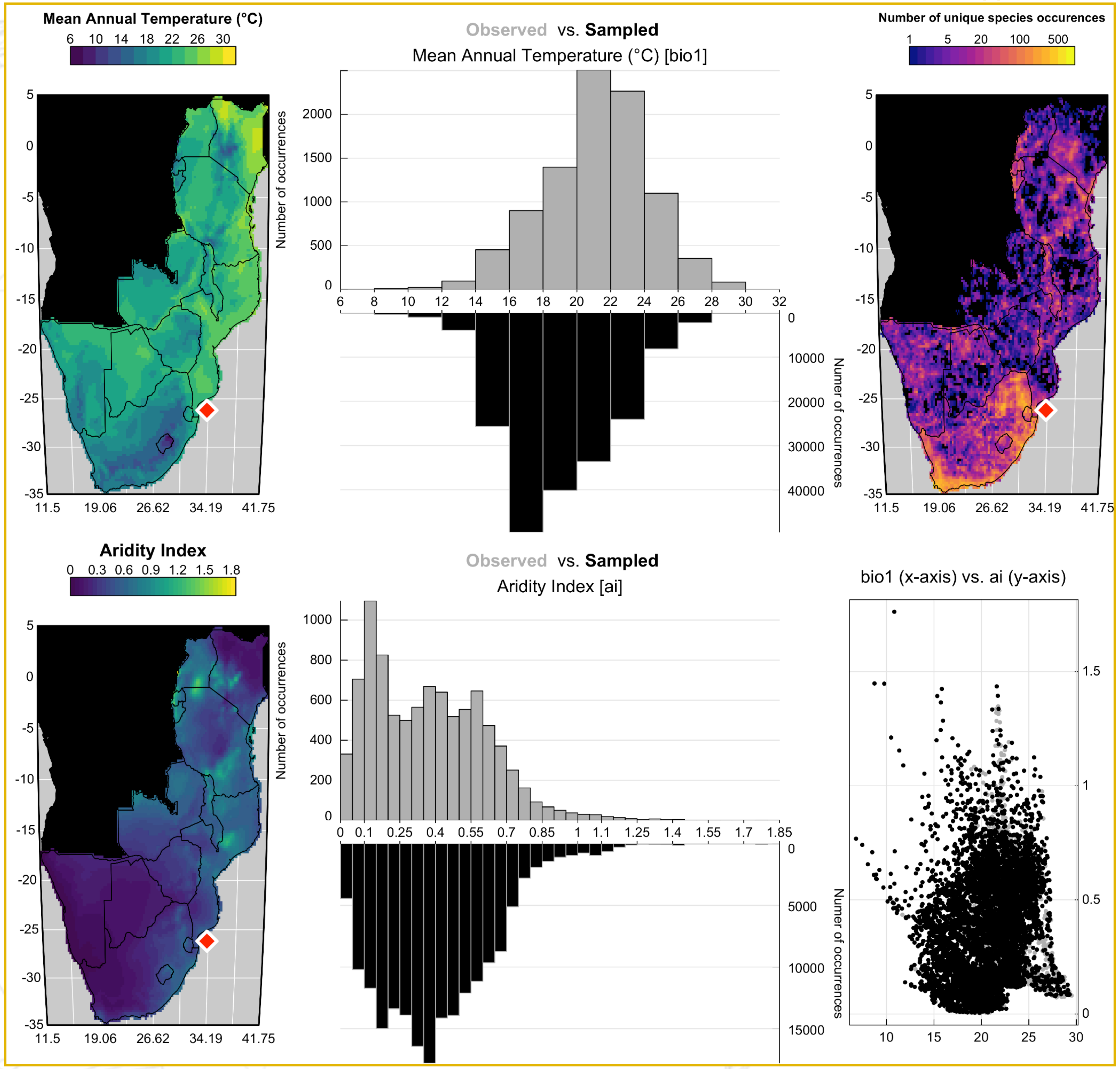
Parameters:

- The shape of the pdfs
- Selecting the balancing of the observed climate space
- If the previous is TRUE, the width of the climate bins
- The number of points for the pdfs

```
rcnstrctn <- crest.calibrate(
                 rcnstrctn,
                 shape = c('normal', 'lognormal'),
                 climateSpaceWeighting = TRUE,
                 bin_width = c(2, 0.05),
                 npoints = 500,
                 verbose = TRUE
)
```

65. BASIC FRAMEWORK

```
plot_climateSpace(rcnstrctn, save=TRUE,
                  filename='Figure 6.png',
                  as.png=TRUE, png.res=600,
                  width=6.9, height=6.6,
                  y0=0.4,
                  add_modern=TRUE
)
```



66 IDENTIFYING CLIMATE-SENSITIVE TAXA

How to determine if a taxon is a good climate indicator?

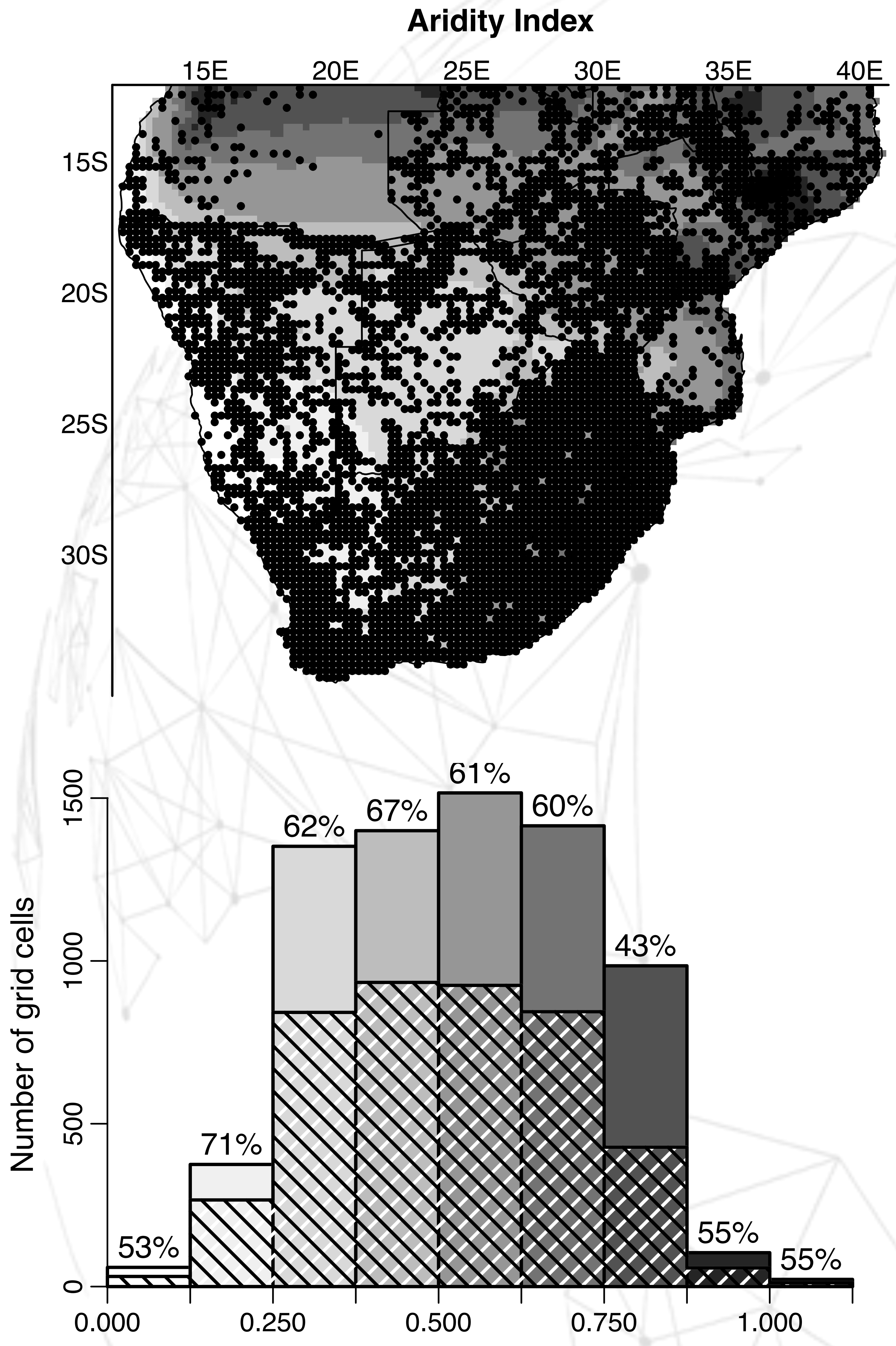
66 IDENTIFYING CLIMATE-SENSITIVE TAXA

How to determine if a taxon is a good climate indicator?

- A limited diversity. Too many composing species can mask climate signals.
- A homogeneous response of the composing species.
- A distribution constrained by one or two variable(s) of interest.

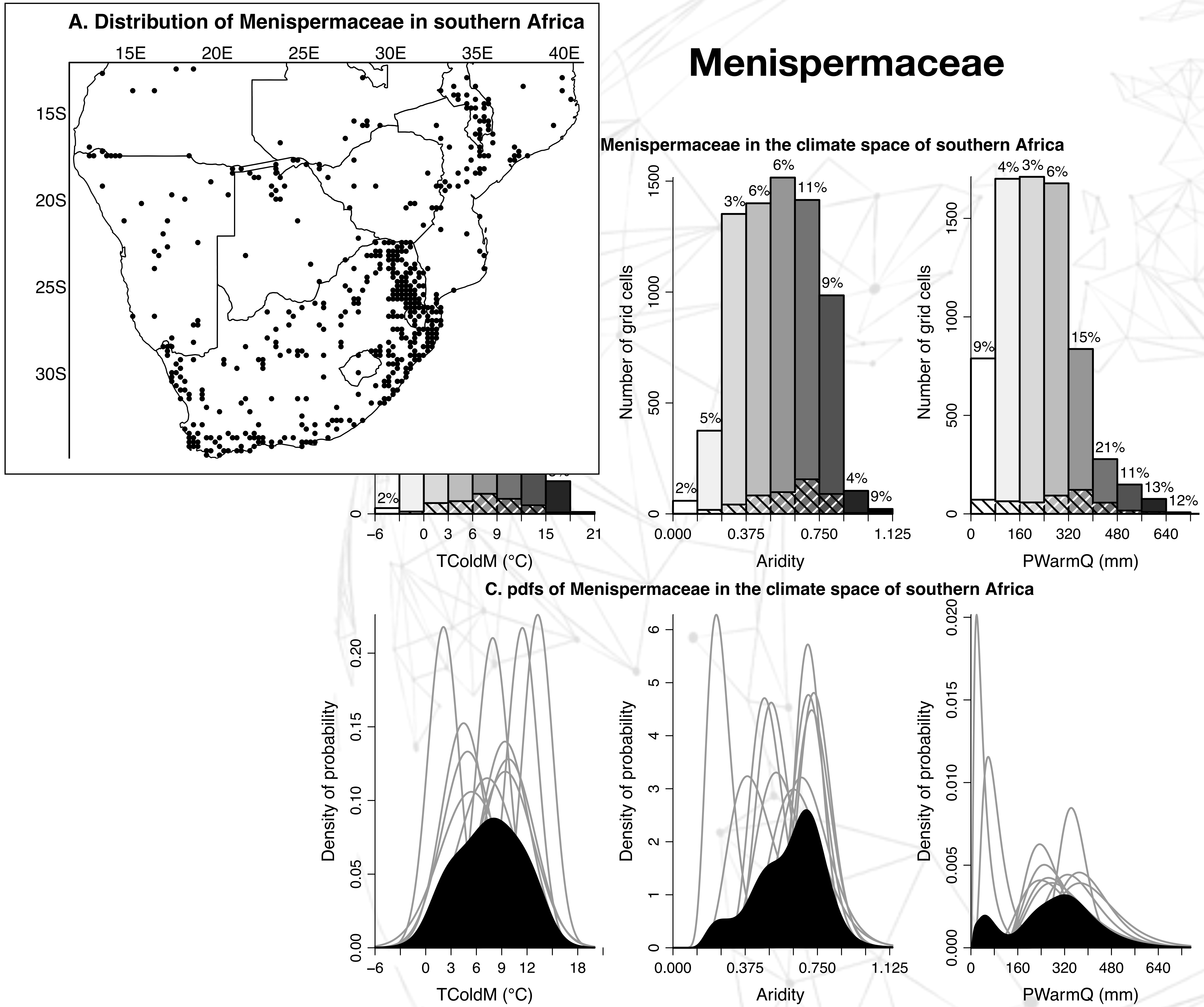
67 IDENTIFYING CLIMATE-SENSITIVE TAXA

Poaceae



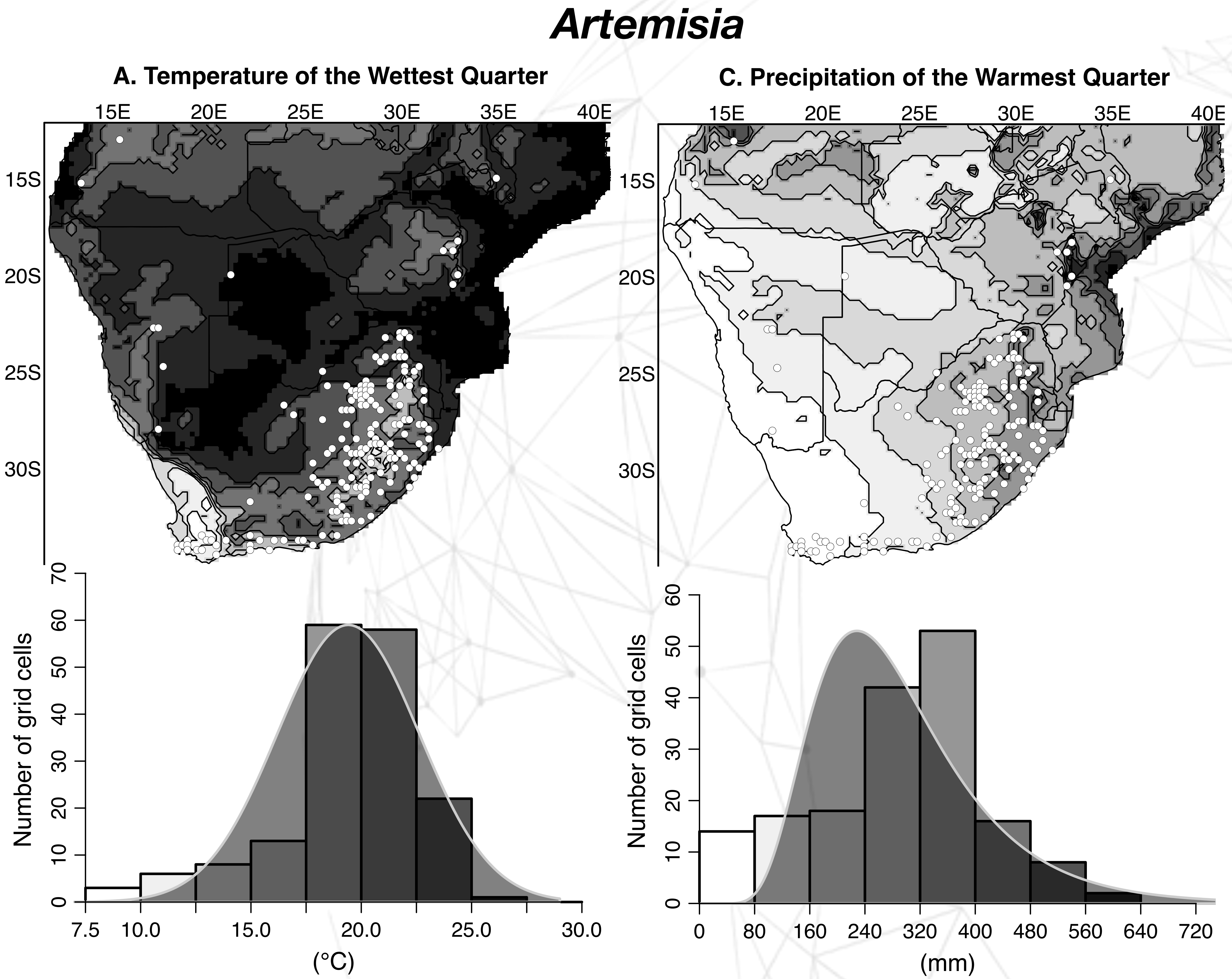
- A limited diversity. Too many composing species can mask climate signals.
- A homogeneous response of the composing species.
- A distribution constrained by one or two variable(s) of interest.

68. IDENTIFYING CLIMATE-SENSITIVE TAXA



- A limited diversity. Too many composing species can mask climate signals.
- A homogeneous response of the composing species.
- A distribution constrained by one or two variable(s) of interest.

69. IDENTIFYING CLIMATE-SENSITIVE TAXA



- A limited diversity. Too many composing species can mask climate signals.
- A homogeneous response of the composing species.
- A distribution constrained by one or two variable(s) of interest.

70. THE CREST POLLEN ATLAS OF SOUTHERN AFRICA

M. Chevalier, B. M. Chase, L. J. Quick, and L. Scott, 'An atlas of southern African pollen types and their climatic affinities', in Quaternary Vegetation Dynamics – The African Pollen Database, vol. 35, London: CRC Press, 2021, pp. 239–258. doi: 10.1201/9781003162766-15.

CHAPTER 15

An atlas of southern African pollen types and their climatic affinities

Manuel Chevalier[1]

Institute of Earth Surface Dynamics, Geopolis, University of Lausanne, Switzerland

Brian M. Chase[2]

Institut des Sciences de l'Evolution-Montpellier (ISEM), University of Montpellier, Centre National de la Recherche Scientifique (CNRS), EPHE, IRD, Montpellier, France.

Lynne J. Quick

African Centre for Coastal Palaeoscience, Nelson Mandela University, Port Elizabeth, South Africa

Louis Scott

Department of Plant Sciences, University of the Free State, Bloemfontein, South Africa

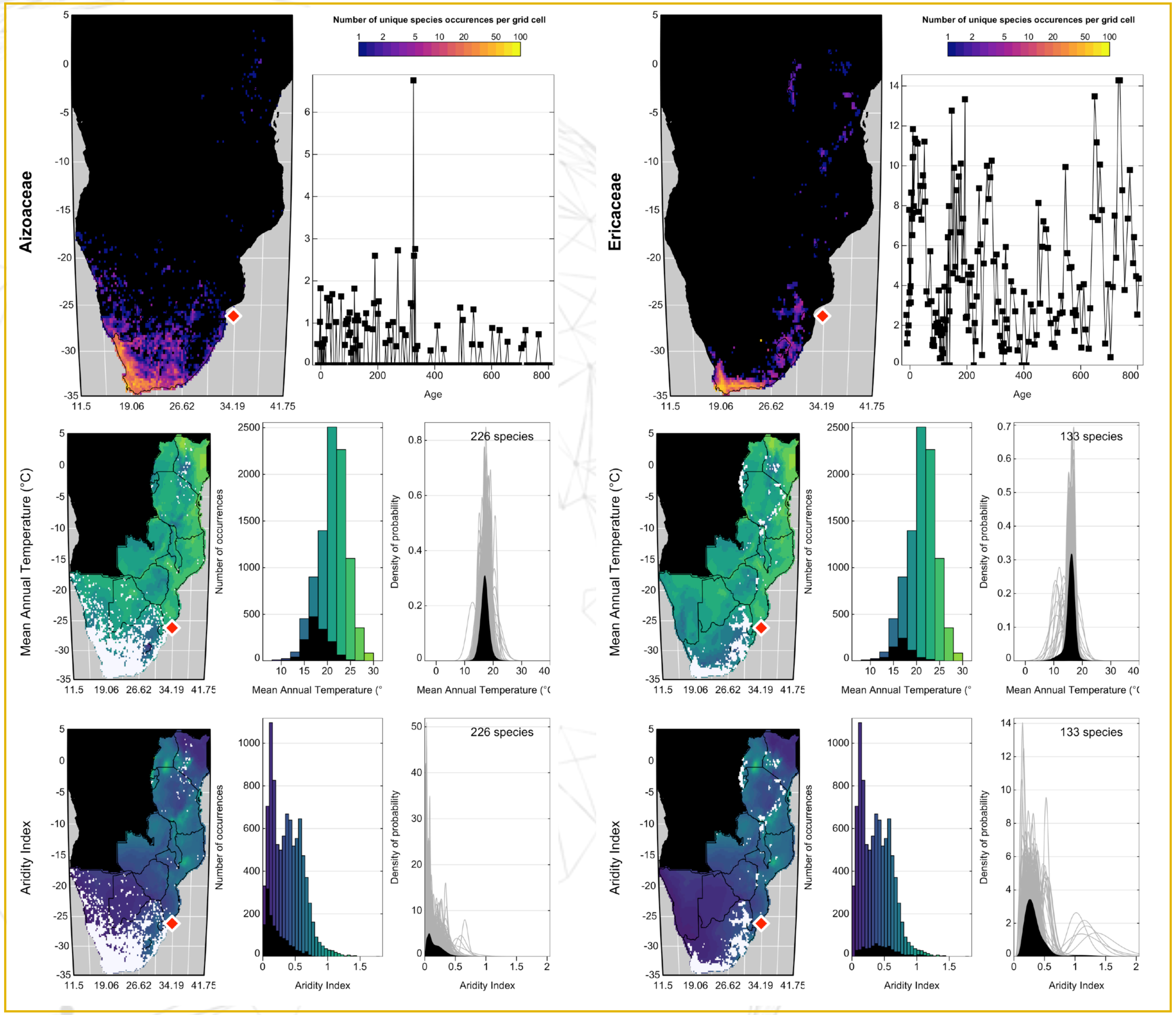
ABSTRACT: Interpretations of fossil pollen data are often limited to broad, qualitative assessments of past climatic and environmental conditions (e.g. colder *vs.* warmer, wetter *vs.* drier, open *vs.* closed landscape). These assessments can be particularly imprecise in regions such as southern Africa, where botanical biodiversity is high, and there exists an associated uncertainty regarding the climatic/environmental sensitivities of the plants contributing to a given pollen type. This atlas addresses this limitation by characterising the climate sensitivities of the 140 pollen morphotypes most often recorded in Late Quaternary palaeoecology studies in southern Africa, relying on their parent plant distributions as one of the basic factors that determine their presence. The atlas is designed as a suite of graphical diagnostic tools and photographs together with analyses of the modern geographical distribution of more than 22,000 plant species to identify their primary climatic sensitivities across southern Africa. Together, the elements included span the complete workflow from pollen identification through interpretation and climate reconstruction. The atlas can be accessed from https://doi.org/10.5281/zenodo.4013452.

ASSESSING THE RESPONSES

You can generate similar plots from your data directly in R with one line of code!

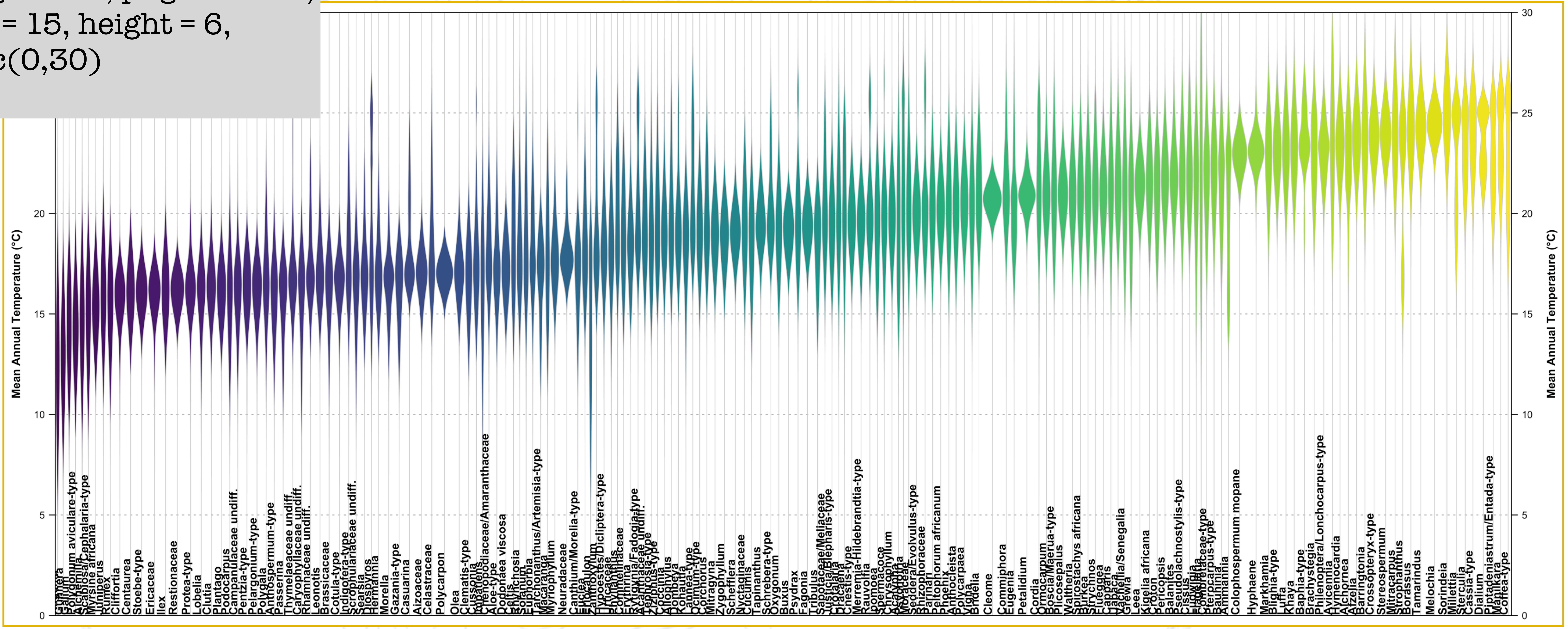
72. ASSESSING THE RESPONSES

```
plot_taxaCharacteristics(
        rcnstrctn,
        taxanames='Ericaceae',
        save = TRUE,
        filename = 'Figure 7.png',
        as.png=TRUE, png.res=600,
        width=4.5, height=8.13,
        add_modern=TRUE
)

plot_taxaCharacteristics(
        rcnstrctn,
        taxanames='Aizoaceae',
        save = TRUE,
        filename = 'Figure 7.png',
        as.png=TRUE, png.res=600,
        width=4.5, height=8.13,
        add_modern=TRUE
)
```



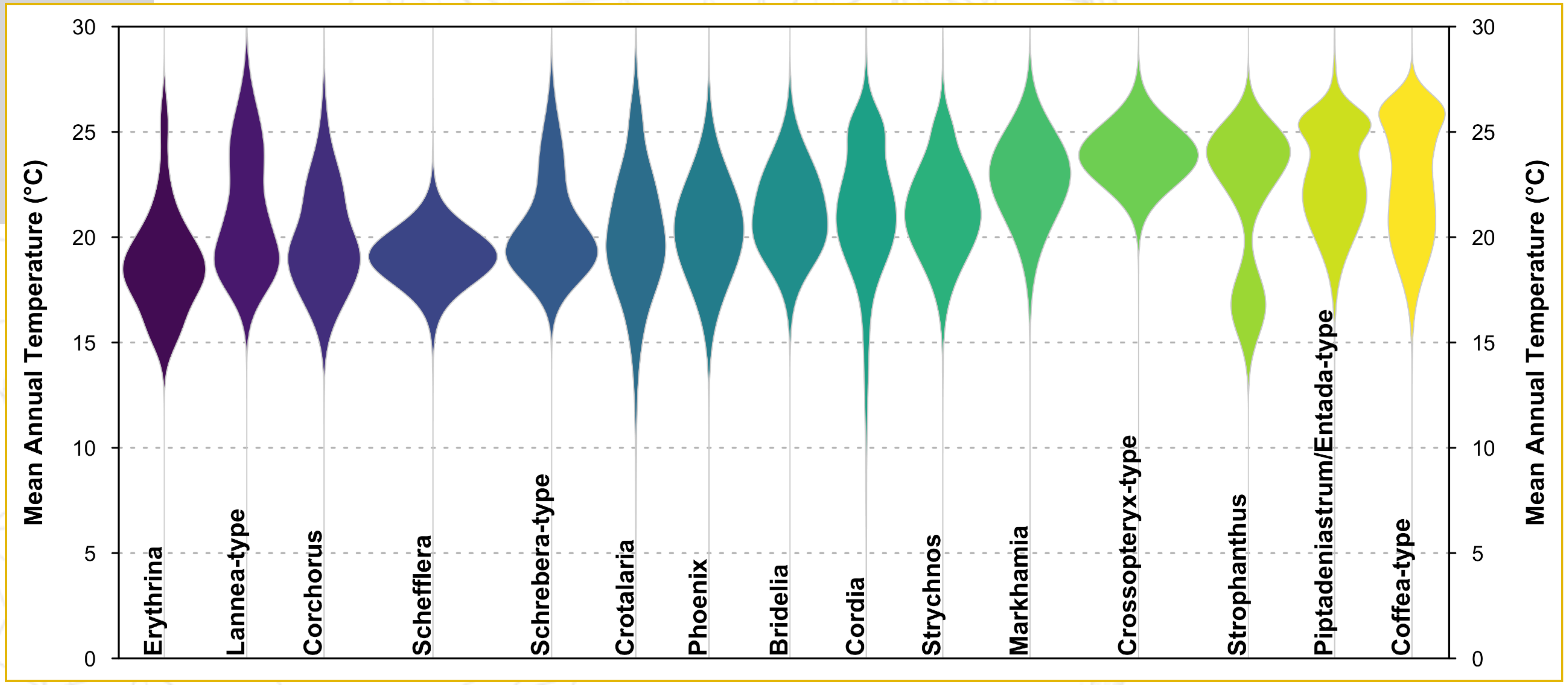
73. ASSESSING THE RESPONSES

```
plot_violinPDFs(rcnstrctn,
          save = TRUE,
          filename = 'Figure 8.png',
          as.png=TRUE, png.res=600,
          width = 15, height = 6,
          ylim=c(0,30)
)
```



74. Assessing the Responses

```
tax <- sample(rcnstrctn$input$taxa.name, 15)

plot_violinPDFs(rcnstrctn,
        taxanames = tax,
        save = TRUE,
        filename = 'Figure 8.png',
        as.png=TRUE, png.res=600,
        width = 15, height = 6,
        ylim=c(0,30)
)
```

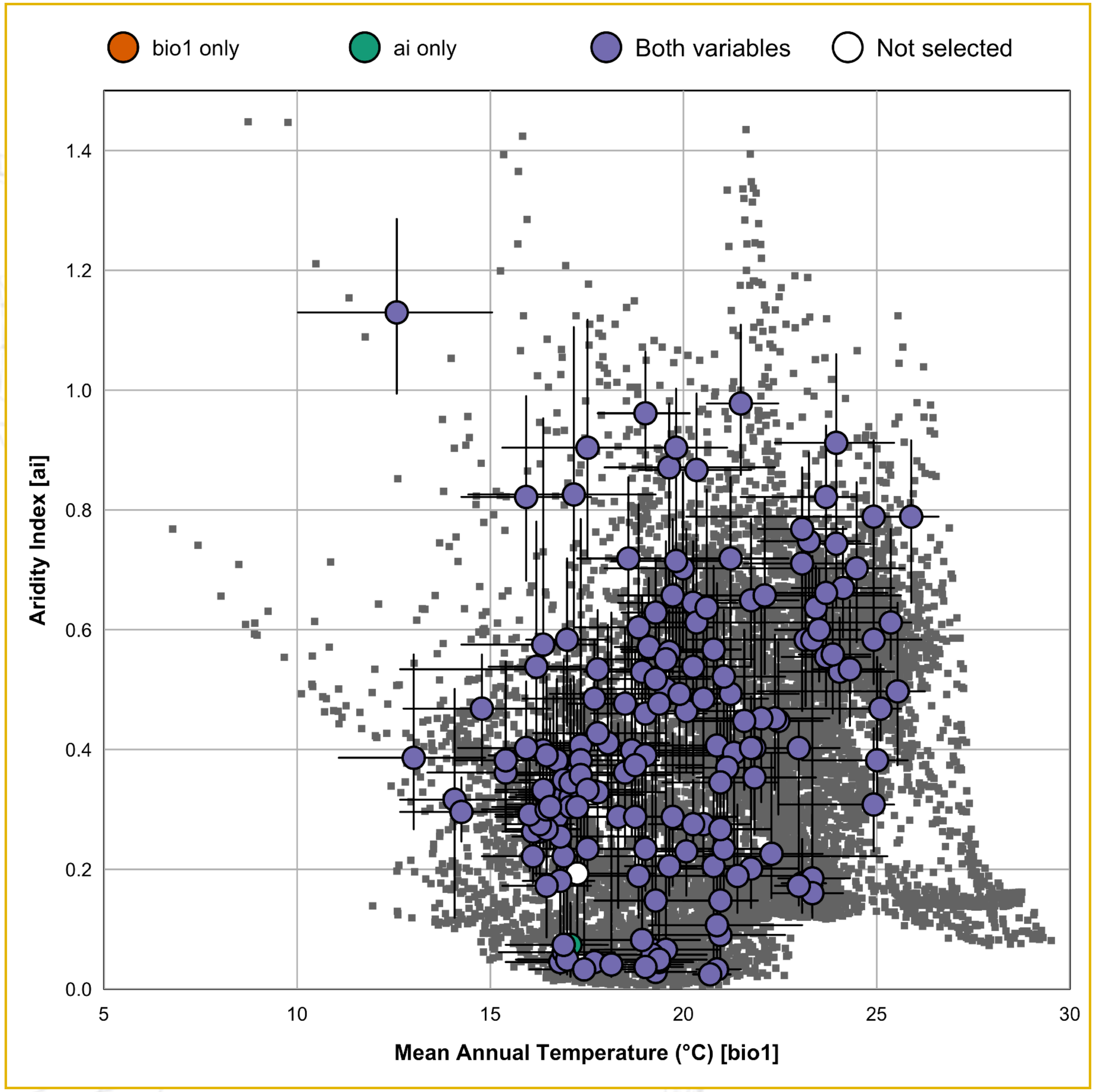


75. ASSESSING THE RESPONSES

A new plot in the last version!

(If working with more than one variable)

```
plot_scatterPDFs (
        rcnstrctn,
        filename = 'Figure 12.png',
        save=TRUE,
        as.png=TRUE,
        png.res=600,
        xlim = c(5, 30),
        ylim = c(0.0, 1.5),
        uncertainties = 0.5
)
```



Excluding taxa from selected variables

```
rcnstrctn <- excludeTaxa(rcnstrctn,
        taxa = c('Aizoaceae', 'Chenopodiaceae/Amaranthaceae'),
        climate=c('bio1')
)


rcnstrctn <- excludeTaxa(rcnstrctn,
        taxa = c('Ericaceae'),
        climate=c('ai')
)
```

SELECTING SPECIFIC TAXA FOR EACH VARIABLE

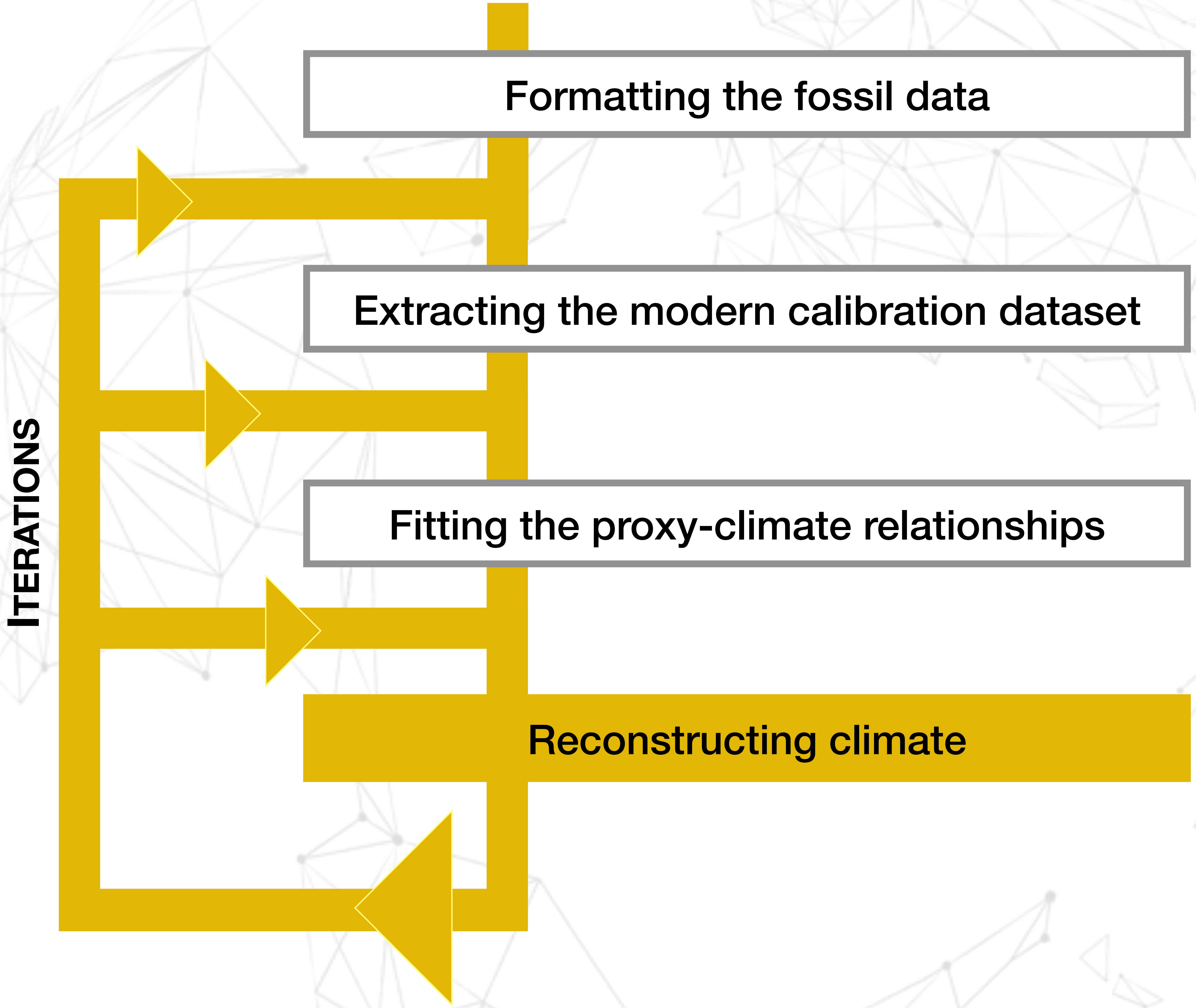
Excluding taxa from selected variables

```
rcnstrctn <- excludeTaxa(rcnstrctn,
          taxa = c('Aizoaceae', 'Chenopodiaceae/Amaranthaceae'),
          climate=c('bio1')
)


rcnstrctn <- excludeTaxa(rcnstrctn,
          taxa = c(‘Ericaceae'),
          climate=c('ai')
)
```

Be careful not to tune the results to your expectations!

(High diversity, unclear relationships, non-climatic drivers, truncated distributions, …)

77. BASIC FRAMEWORK



Previous slides

`crest.get_modern_data(...)`

+ diagnostic tools

`crest.calibrate(...)`

+ diagnostic tools

`crest.reconstruct(...)`

+ diagnostic tools

78. RECONSTRUCTING CLIMATE

Parameters:

- The minimum value to count as a 'true' observation
- The type of weighting to use

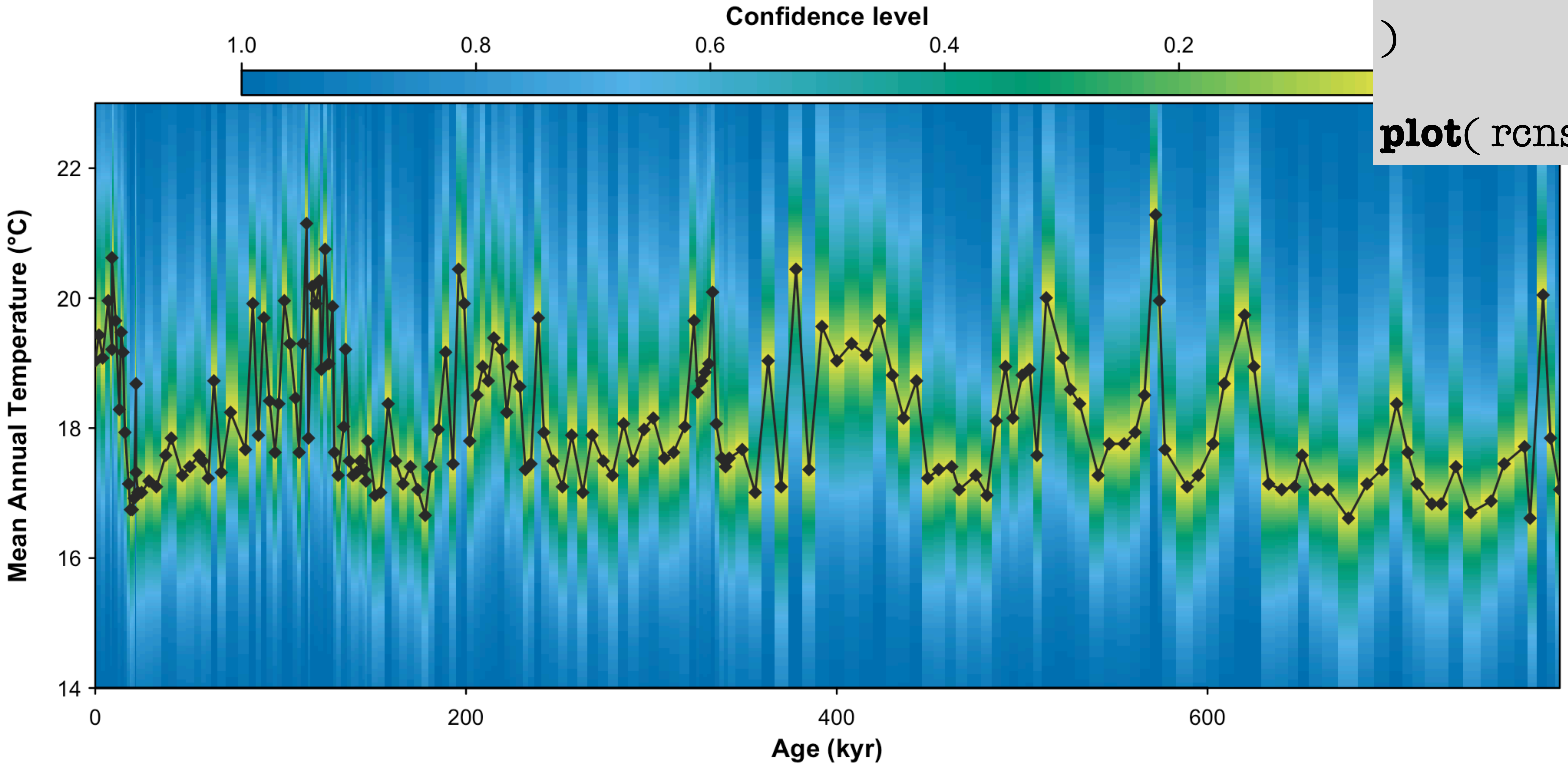
```
rcnstrctn <- crest.reconstruct(
            rcnstrctn,
            presenceThreshold = 0,
            taxWeight = "normalisation",
            verbose = TRUE
)
```

78. RECONSTRUCTING CLIMATE

Parameters:

- The minimum value to count as a 'true' observation
- The type of weighting to use

```
rcnstrctn <- crest.reconstruct(
            rcnstrctn,
            presenceThreshold = 0,
            taxWeight = "normalisation",
            verbose = TRUE
)

plot( rcnstrctn, simplify=FALSE )
```

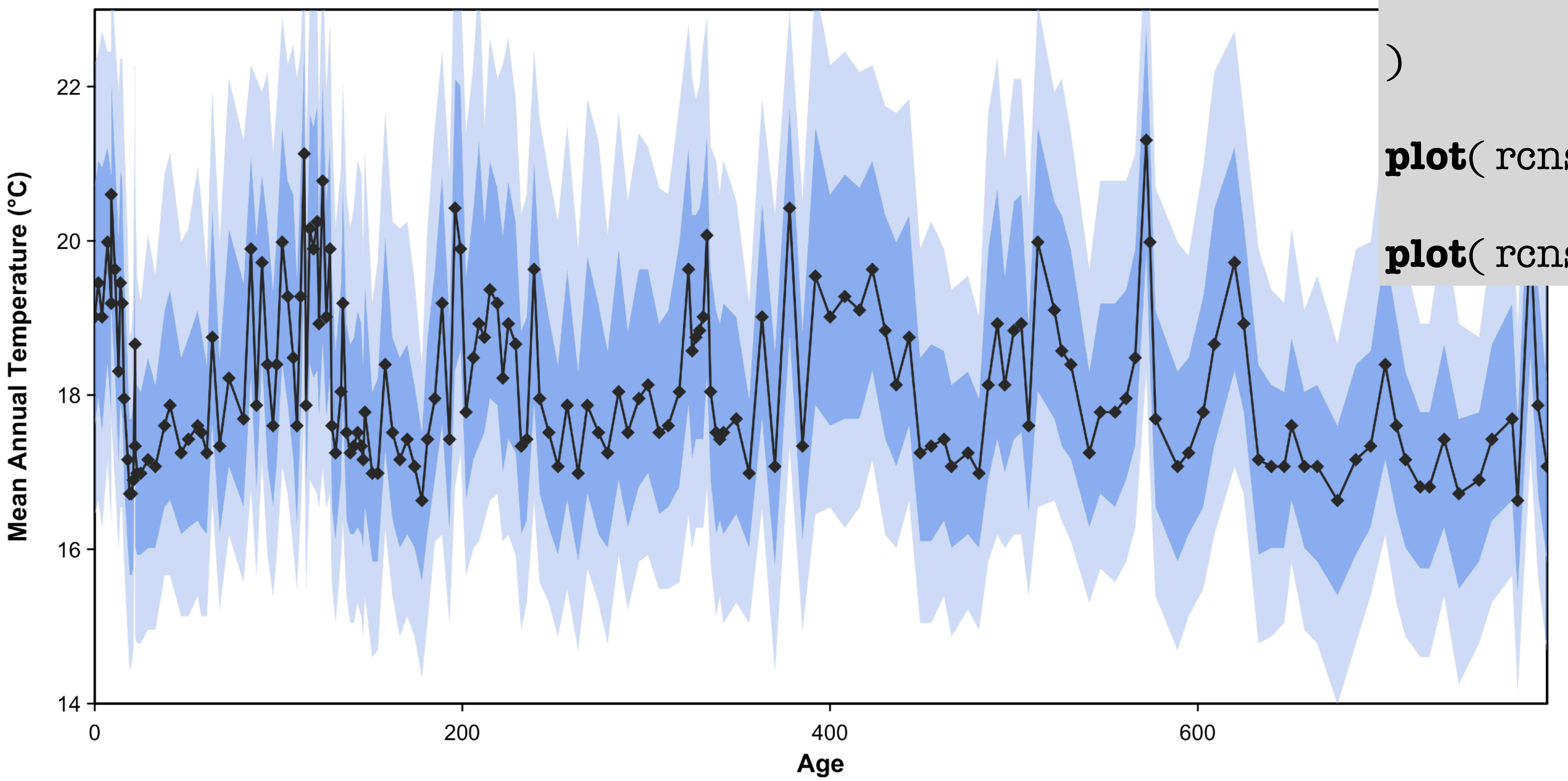


78. RECONSTRUCTING CLIMATE

Parameters:

- The minimum value to count as a ‘true’ observation
- The type of weighting to use

```
rcnstrctn <- crest.reconstruct(
        rcnstrctn,
        presenceThreshold = 0,
        taxWeight = "normalisation",
        verbose = TRUE
)

plot( rcnstrctn, simplify=FALSE )

plot( rcnstrctn, simplify=TRUE )
```



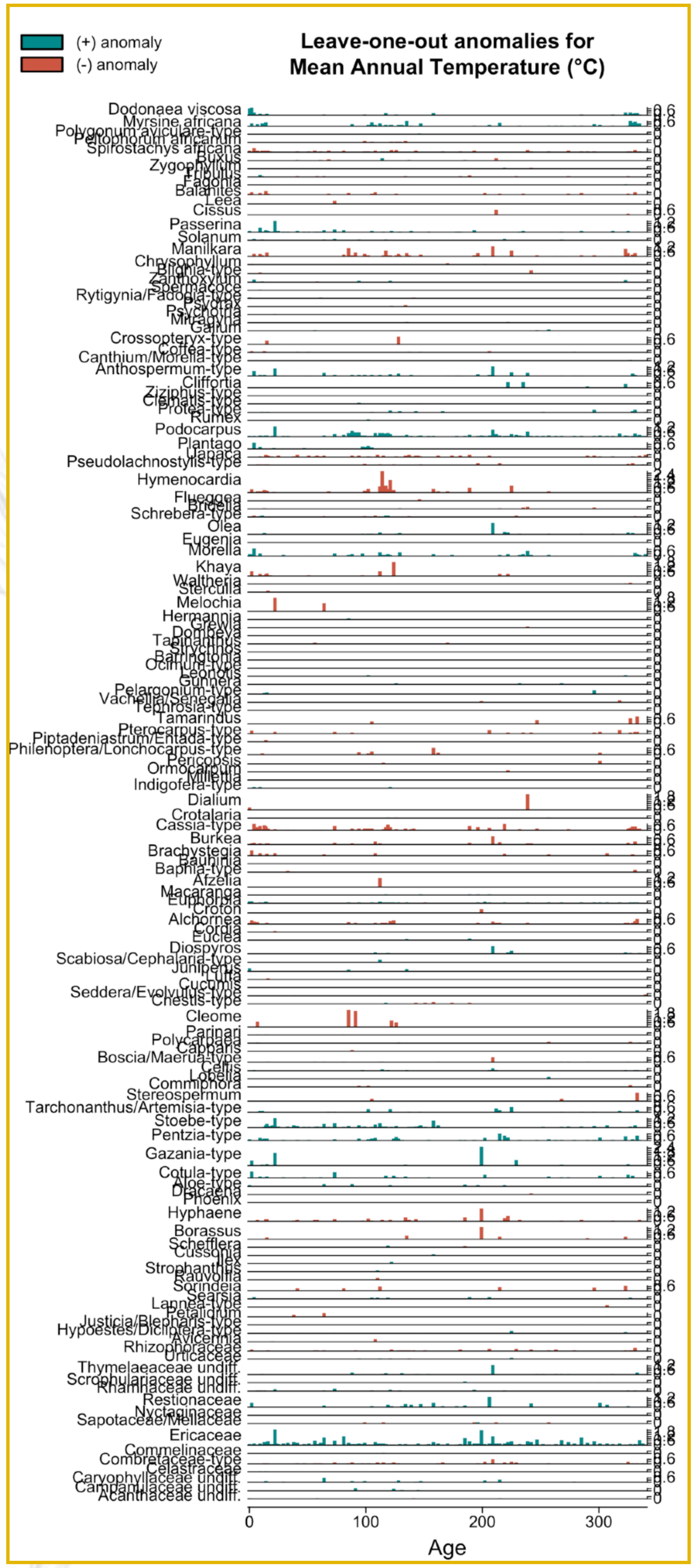
79. ASSESSING RECONSTRUCTIONS

A powerful way to evaluate the results:
the leave-one-out reconstructions

$\text{LOO}(\text{taxon}_x) = \text{Recon}(\text{without taxon}_x) - \text{Recon}(\text{all taxa})$

```
rcnstrctn <- loo(rcnstrctn,
  verbose = TRUE
)

plot_loo (rcnstrctn,
      taxanames=rcnstrctn$inputs$taxa.name,
      climate = 'bio1',
      sort = NA, filter = 0,
      xlim=c(0, 340), save = TRUE,
      filename = 'Figure 11.png',
      as.png=TRUE, png.res=600,
      width=3.5, height=8,
      bar_width=3, col_neg='coral3',
      col_pos='darkcyan'
)
```

- Can be hard to read
- Too many taxa
- Many uninformative taxa
- Taxa in random order

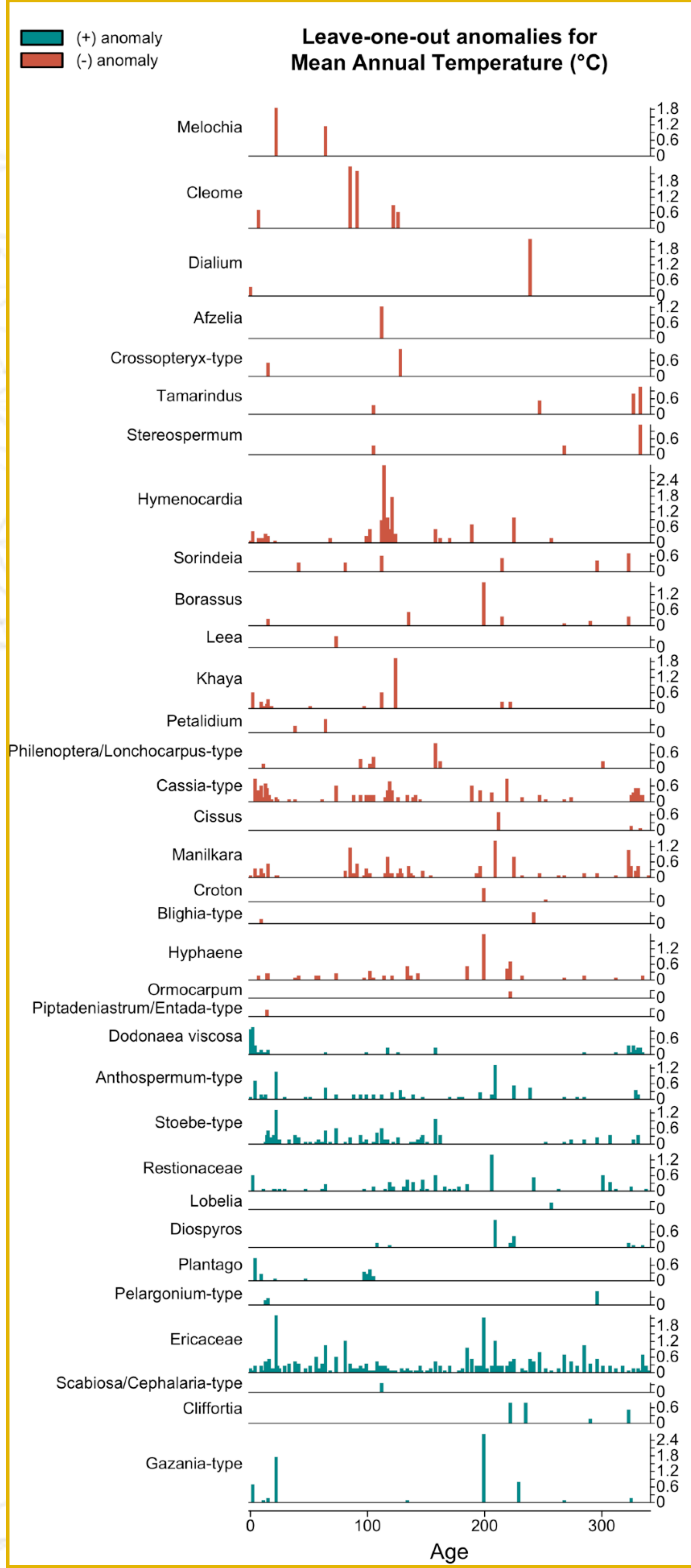


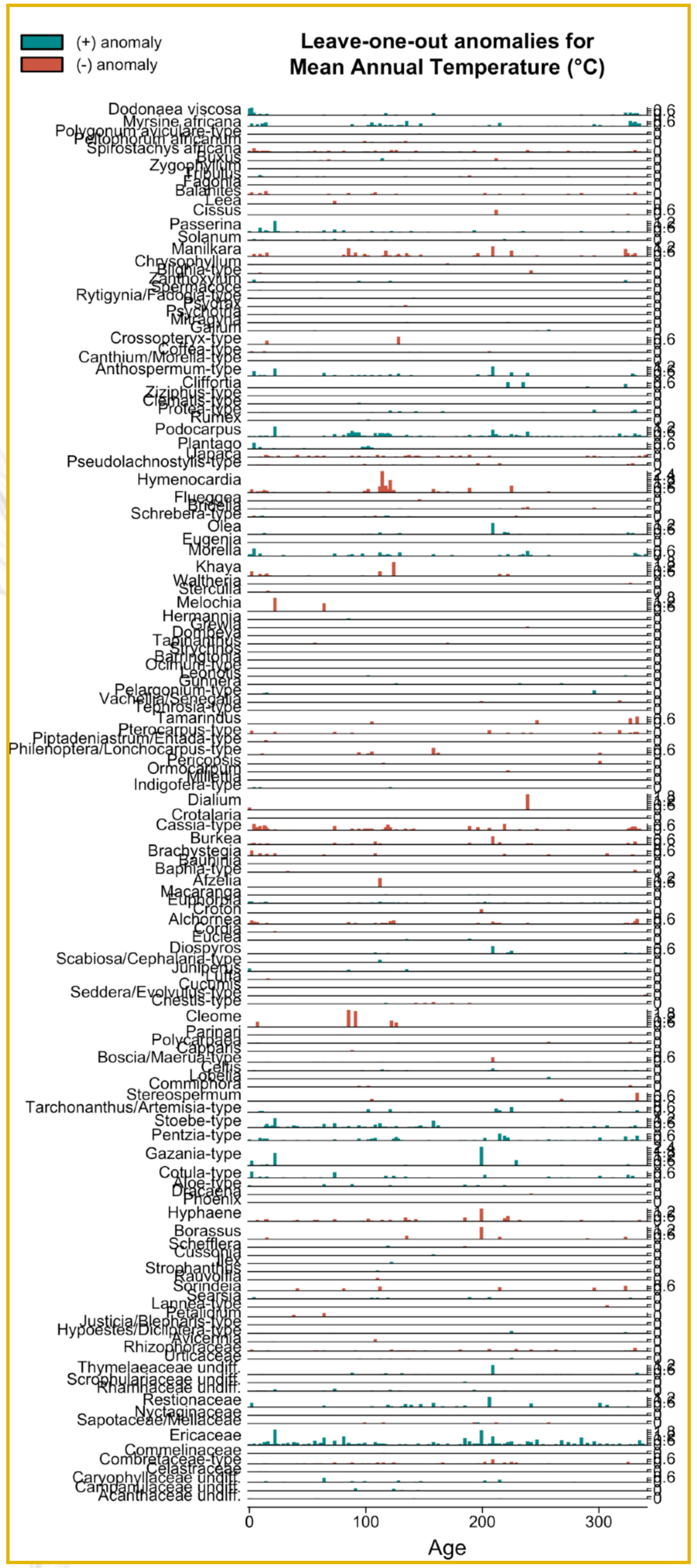
79. ASSESSING RECONSTRUCTIONS

A powerful way to evaluate the results:
the leave-one-out reconstructions

$$\mathrm{LOO}(\mathrm{taxon}_x) = \mathrm{Recon}(\text{without taxon}_x) - \mathrm{Recon}(\text{all taxa})$$

```
rcnstrctn <- loo(rcnstrctn,
  verbose = TRUE
)

plot_loo (rcnstrctn,
      taxanames=rcnstrctn$inputs$taxa.name,
      climate = 'bio1',
      sort = 'decr', filter = 0.25,
      xlim=c(0, 340), save = TRUE,
      filename = 'Figure 11.png',
      as.png=TRUE, png.res=600,
      width=3.5, height=8,
      bar_width=3, col_neg='coral3',
      col_pos='darkcyan'
)
```



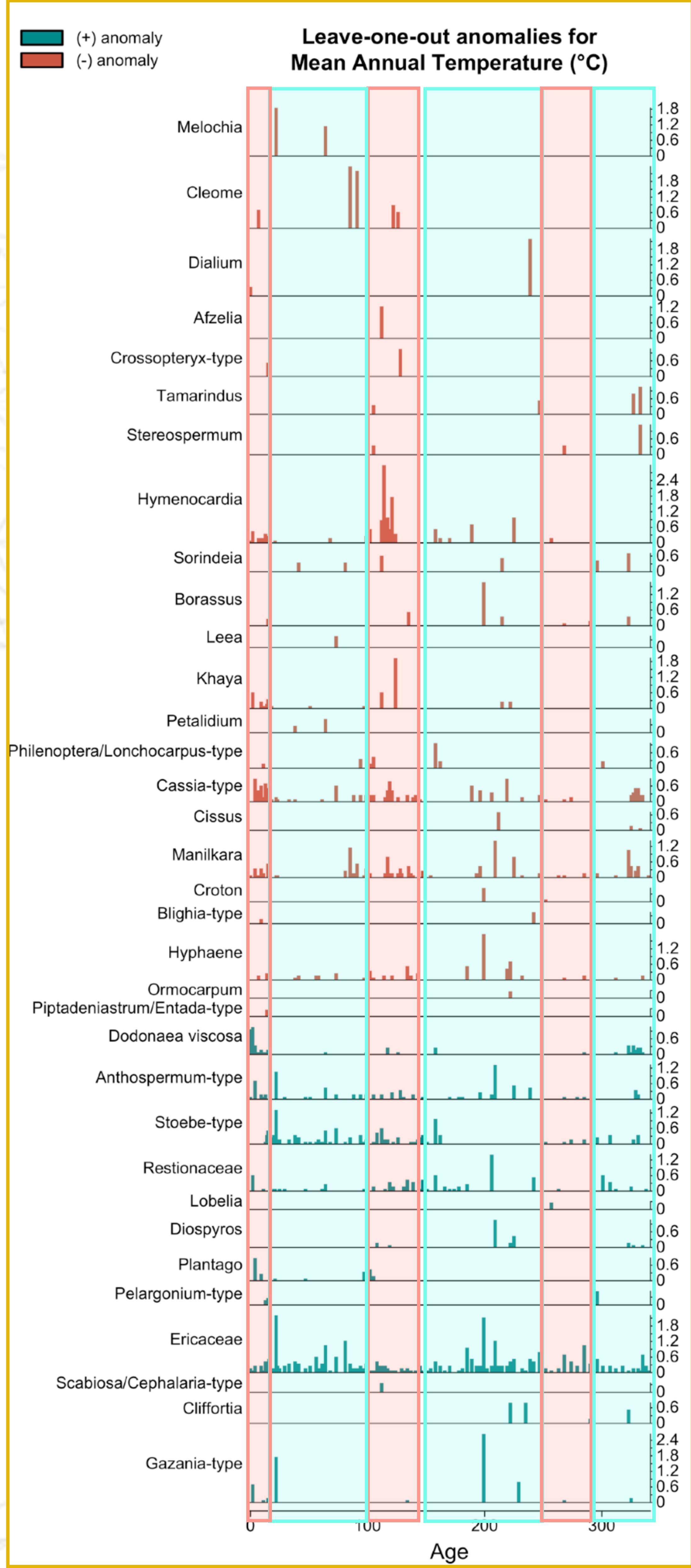


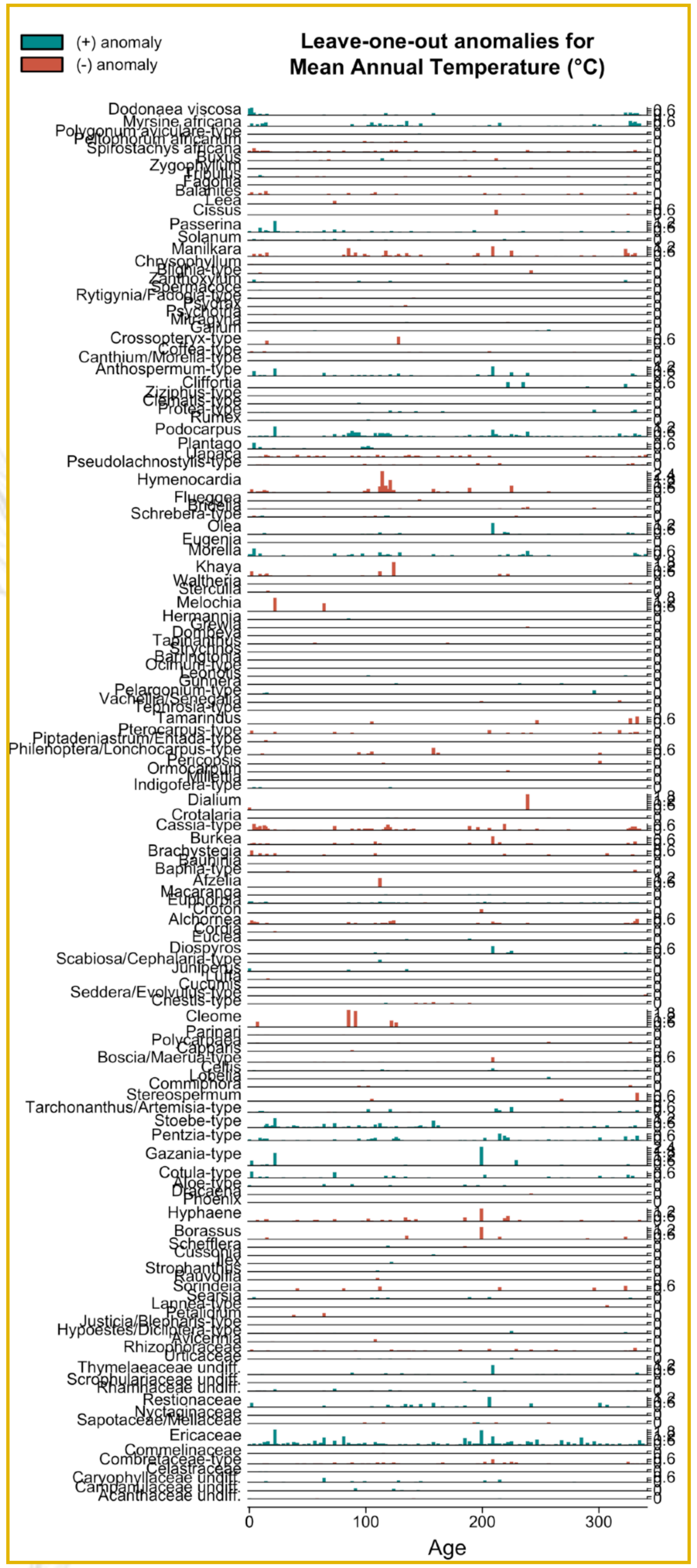
79. ASSESSING RECONSTRUCTIONS

A powerful way to evaluate the results:
the leave-one-out reconstructions

$$\mathrm{LOO}(\mathrm{taxon}_x) = \mathrm{Recon}(\mathrm{without\ taxon}_x) - \mathrm{Recon}(\mathrm{all\ taxa})$$

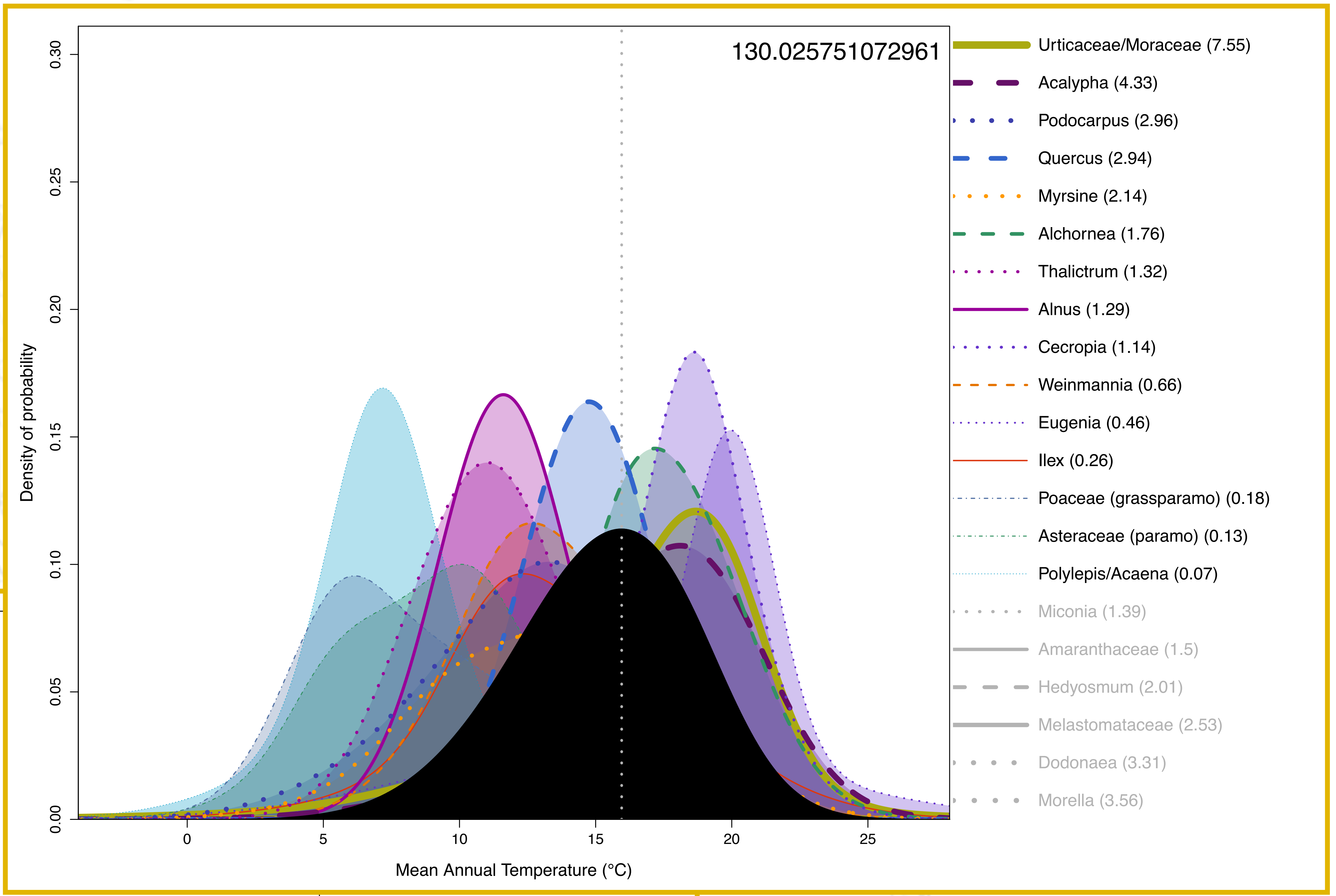
```
rcnstrctn <- loo(rcnstrctn,
  verbose = TRUE
)

plot_loo (rcnstrctn,
        taxanames=rcnstrctn$inputs$taxa.name,
        climate = 'bio1',
        sort = 'decr', filter = 0.25,
        xlim=c(0, 340), save = TRUE,
        filename = 'Figure 11.png',
        as.png=TRUE, png.res=600,
        width=3.5, height=8,
        bar_width=3, col_neg='coral3',
        col_pos='darkcyan'
)
```

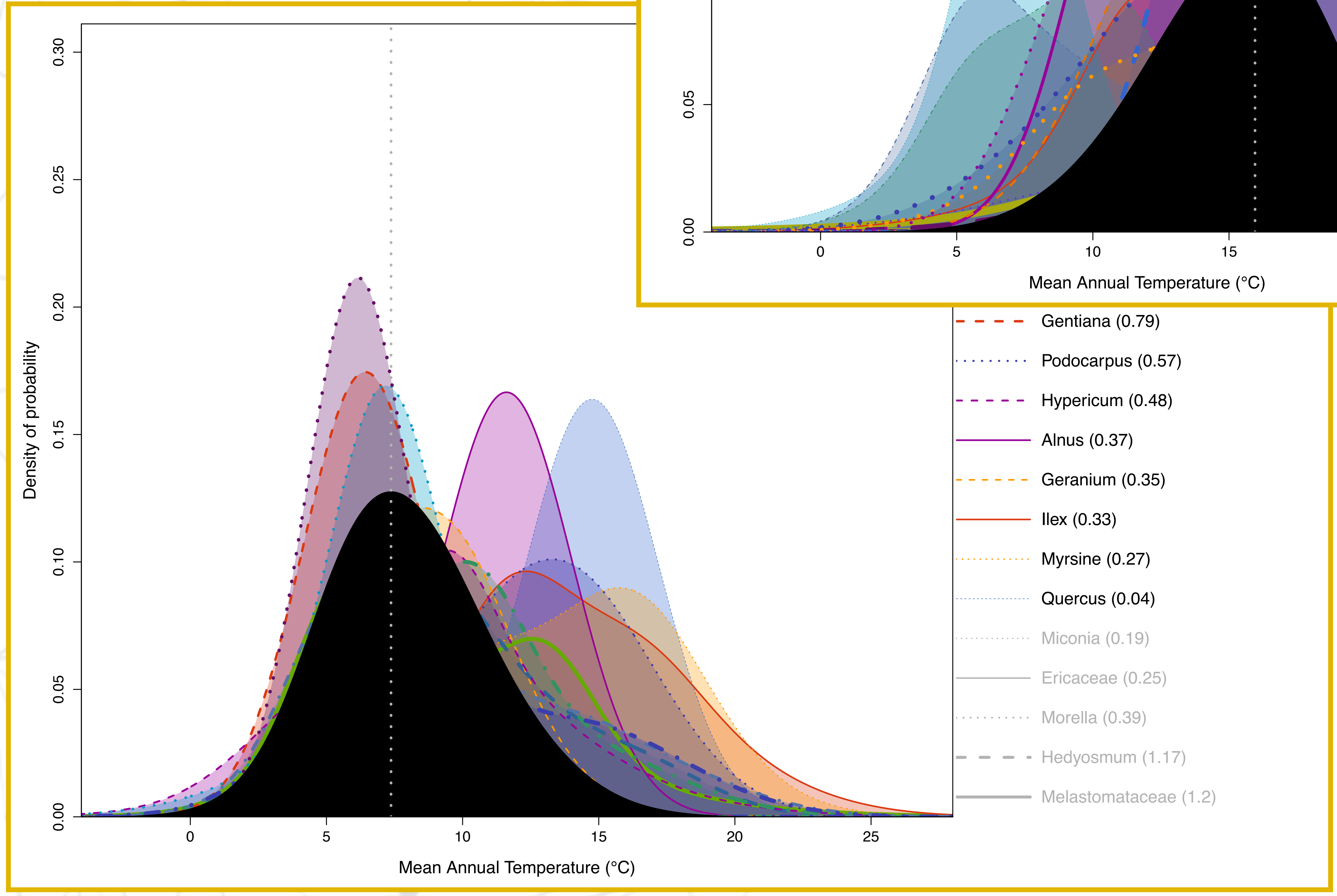




80. ASSESSING THE RESPONSES

```
plot_combinedPDFs(rcnstrctn,
            samples=c(1,2,3),
            climate='bio1',
            only.present=TRUE,
            only.selected=FALSE,
            save = TRUE,
            filename = 'Figure 8.png',
            as.png=TRUE, png.res=600
)
```





SUMMARY

Optimising all of these factors will often be a trade-off.

You will need to find the "sweet spot" between all the possible parameters where everything is as coherent and cohesive as possible.

81. SUMMARY

Optimising all of these factors will often be a trade-off.

You will need to find the "sweet spot" between all the possible parameters where everything is as coherent and cohesive as possible.

Back to rule #2: it is an ITERATIVE process.

Do not aim to get it right in the first place: be curious about the data, try different modelling options, look at the data from different angles, and try to observe the data without any preconceived ideas.

After the analysis

83. BASIC FRAMEWORK

ITERATIONS

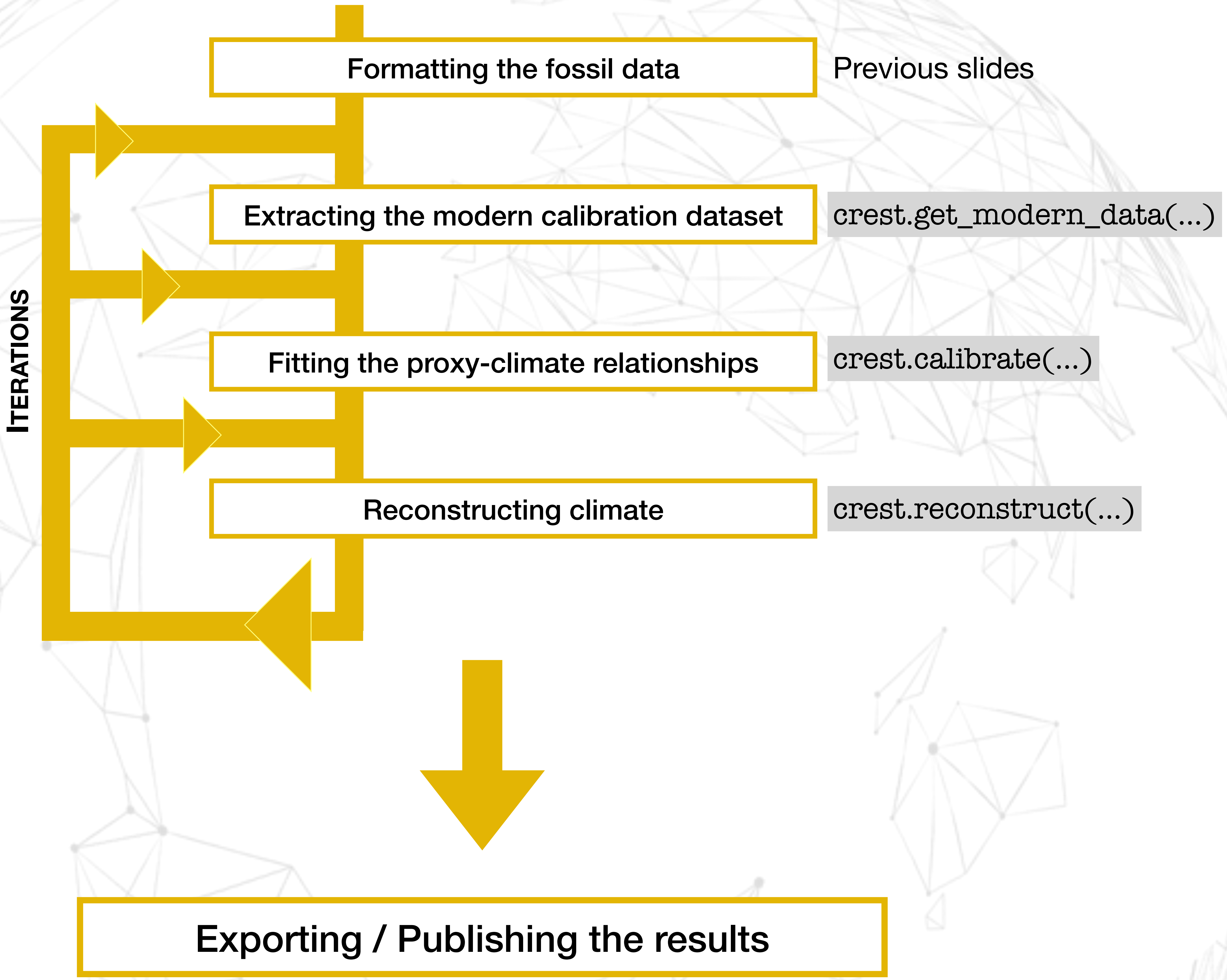
Formatting the fossil data — Previous slides

Extracting the modern calibration dataset — `crest.get_modern_data(...)`

Fitting the proxy-climate relationships — `crest.calibrate(...)`

Reconstructing climate — `crest.reconstruct(...)`

83. BASIC FRAMEWORK



EXPORTING THE RESULTS

What should I export to ensure transparency & reproducibility

84 EXPORTING THE RESULTS

What should I export to ensure transparency & reproducibility

- The data as used for the reconstruction (percentages, presence/absence, …)
- The modelling assumptions for the pdfs and reconstructions.
- The species-proxy table
- The calibration dataset (use the diagnostic tool!)
- The version of the package!

84. EXPORTING THE RESULTS

What should I export to ensure transparency & reproducibility

```
export(rcnstrctn,
        dataname='final_recon',
        loo=TRUE,
        fullUncertainties=FALSE,
        pdfs=FALSE
)
```

- The data as used for the reconstruction (percentages, presence/absence, …)
- The modelling assumptions for the pdfs and reconstructions.
- The species-proxy table
- The calibration dataset (use the diagnostic tool!)
- The version of the package!

84. EXPORTING THE RESULTS

What should I export to ensure transparency & reproducibility

```
export(rcnstrctn,
        dataname='final_recon',
        loo=TRUE,
        fullUncertainties=FALSE,
        pdfs=FALSE
)
```

- The data as used for the reconstruction (percentages, presence/absence, …)
- The modelling assumptions for the pdfs and reconstructions.
- The species-proxy table
- The calibration dataset (use the diagnostic tool!)
- The version of the package!

85 EXPORTING THE RESULTS

Every reconstruction is based on several datasets, methods, packages that deserve recognition

cite_crest(rcnstrctn)

This analysis is built on the following **8 references**. Please cite them appropriately.

Methodology:
--> Chevalier, M., Cheddadi, R. and Chase, B.M., 2014, CREST (Climate REconstruction SofTware): a probability density function (PDF)-based quantitative climate reconstruction method. Climate of the Past, 10(6), pp. 2081-2098. doi: 10.5194/cp-10-2081-2014
--> Chevalier, M., 2022, crestr: an R package to perform probabilistic climate reconstructions from palaeoecological datasets, Climate of the Past, 18(4), pp. 821-844, doi:10.5194/cp-18-821-2022
--> Chevalier, M., 2019, Enabling possibilities to quantify past climate from fossil assemblages at a global scale. Global and Planetary Change, 175, pp. 27-35. doi: 10.1016/j.gloplacha.2019.01.016

Distribution data:
--> GBIF.org (Date accessed: 24 September 2020) Magnoliopsida occurrence data. https://doi.org/10.15468/dl.ra49dt.
--> GBIF.org (Date accessed: 24 September 2020) Liliopsida occurrence data. https://doi.org/10.15468/dl.axv3yd.
--> GBIF.org (Date accessed: 24 September 2020) Pinopsida occurrence data. https://doi.org/10.15468/dl.x2r7pa.

Climate data:
--> Fick, S.E. and Hijmans, R.J., 2017, WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37, pp. 4302-4315.
--> Zomer, R.J., Trabucco, A., Bossio, D.A. and Verchot, L. V., 2008, Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agriculture, Ecosystems & Environment, 126, pp. 67-80.

Note: This output is also integrated in export()

Available ressources

87. WHERE TO START?

https://www.manuelchevalier.com/crestr/

crestr 1.3.0.9000 | How it works | Get started | Prepare your data | Calibration data | Reference | Community | Vignettes | News

crestr An R package to perform probabilistic palaeoclimate reconstructions from palaeoecological datasets

The goal of `crestr` is to produce probabilistic reconstructions of past climate change from fossil assemblage data (Chevalier, 2022). The approach is based on the estimation of responses of studies bio-proxy studied to climate parameters using probability density functions (*PDFs*; see Chevalier *et al.* (2014) and Chevalier (2019)). The theory underpinning this package is explained in section *A bit of theory* and is illustrated with an application based on pseudo-data in section *Get Started*. The different vignettes present different aspects of the structure of the package and the data it contains, along with applications based on real data.

> ***NOTE:*** If you notice any bug, or if you would like to see some specific functions implemented, you can contact me at chevalier.manuel@gmail.com.

Installation

The package is currently unavailable from CRAN. However, you can install the latest development version (1.3.0) from GitHub with:

```
if(!require(devtools)) install.packages("devtools")
devtools::install_github("mchevalier2/crestr")
```

Links

View on CRAN

Browse source code

Report a bug

License

Full license

MIT + file LICENSE

Citation

Citing crestr

Developers

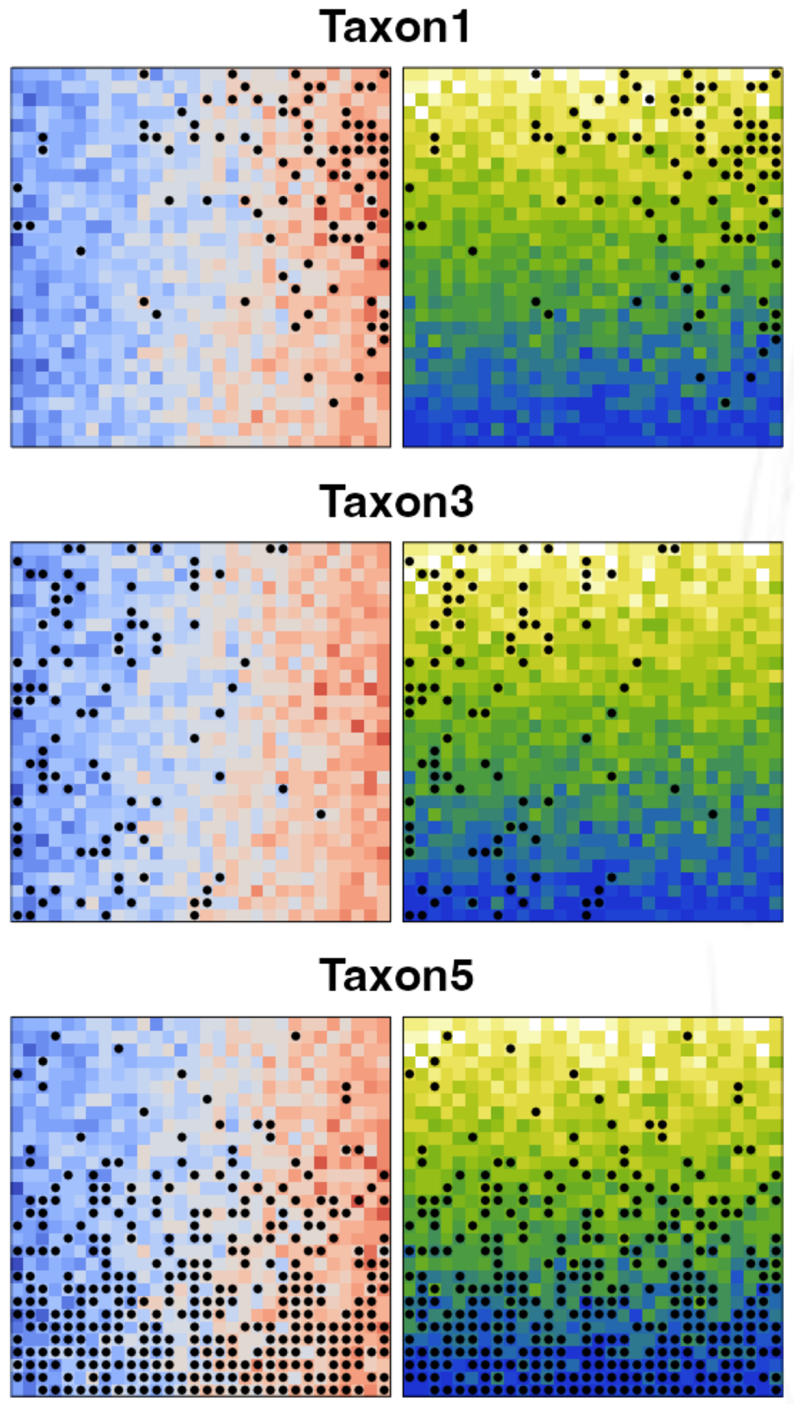
Manuel Chevalier
Author, maintainer

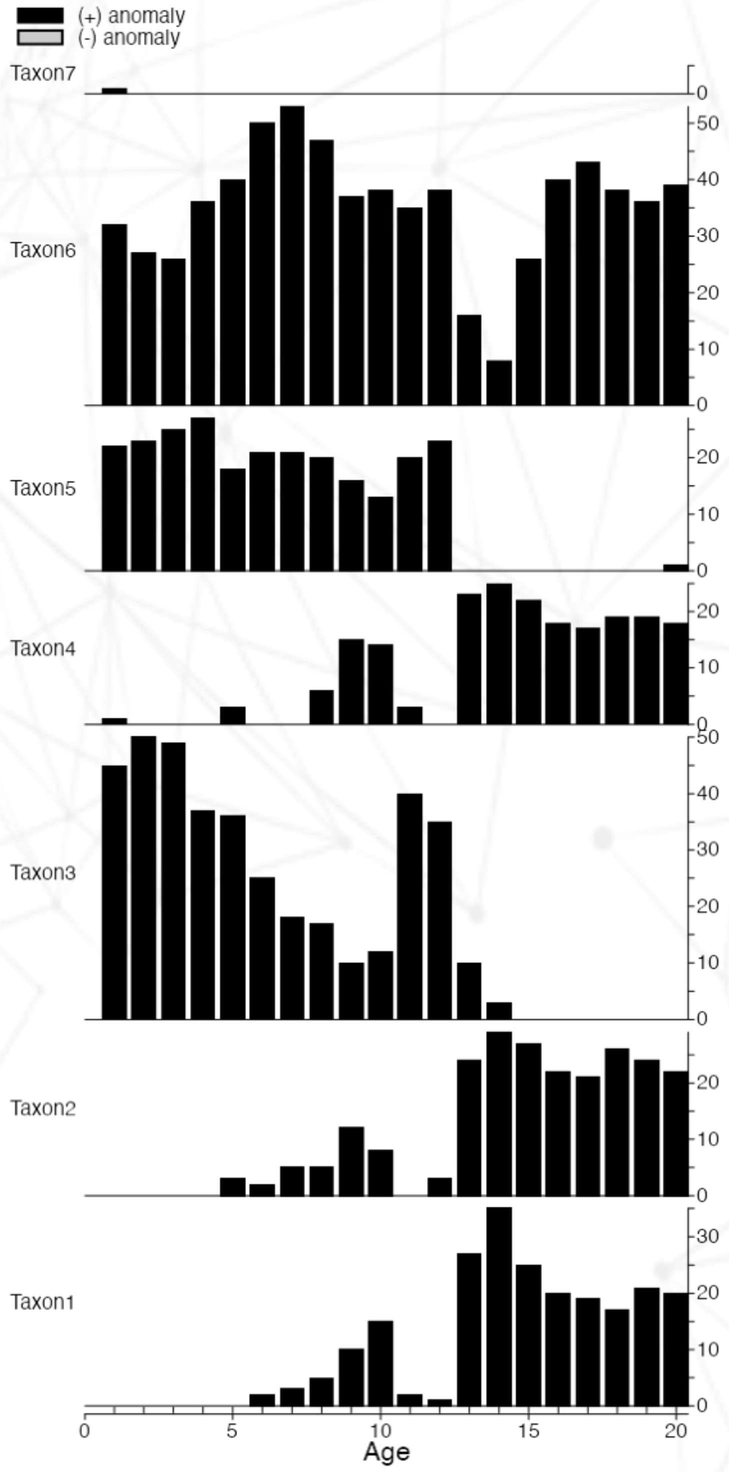
TABLE OF CONTENT

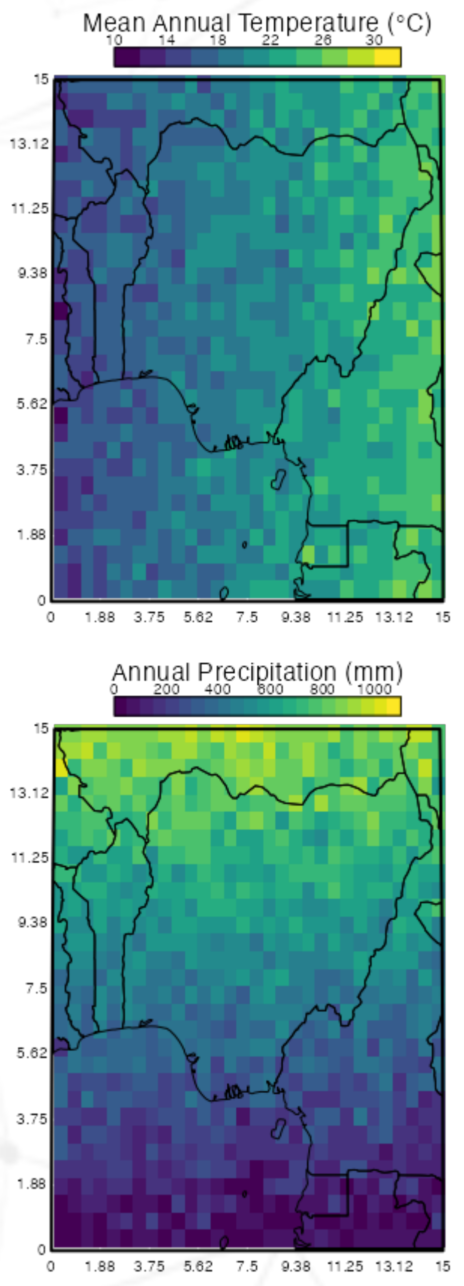
89. YOUR TIME TO SHINE

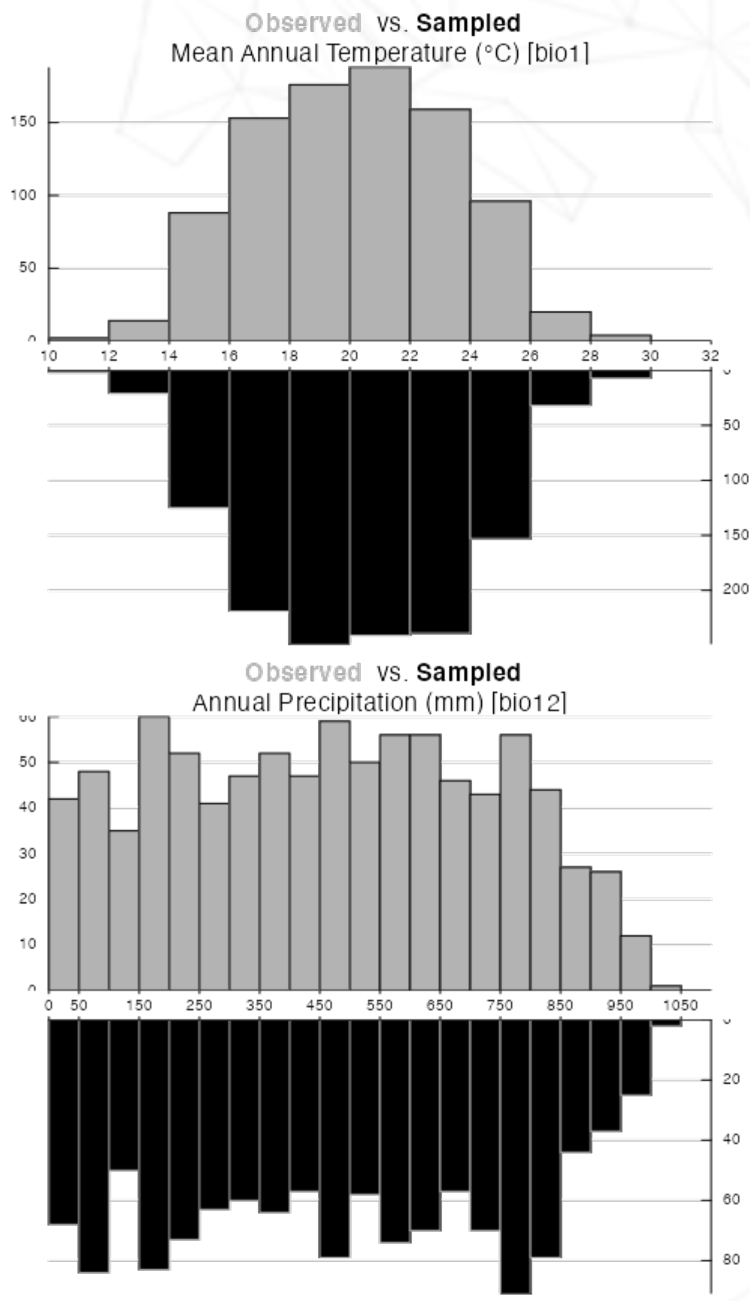
Follow the 'Get started' example from the webpage to get familiar with the crestr environment:

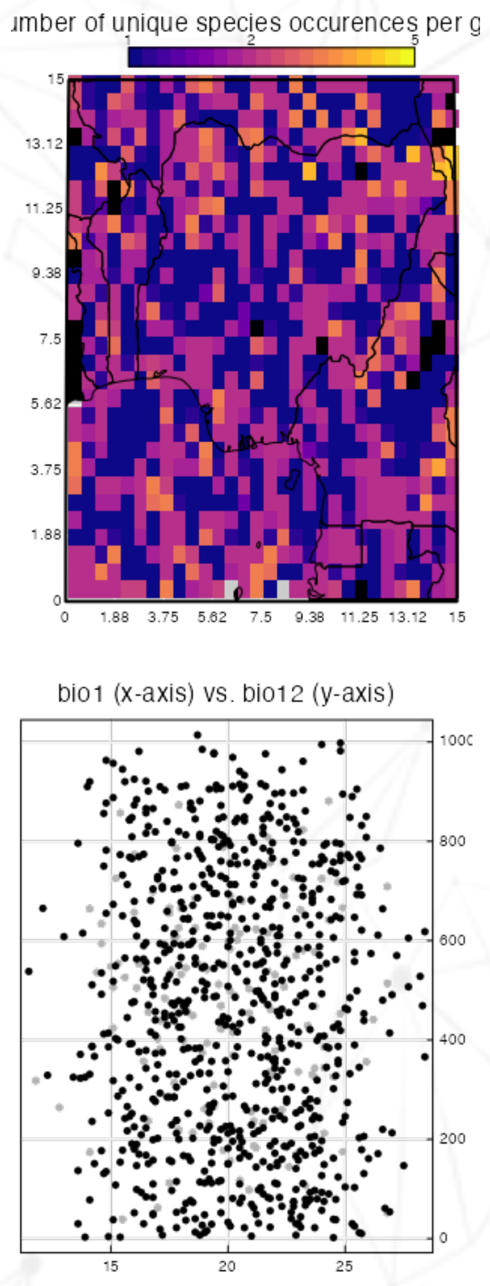
https://www.manuelchevalier.com/crestr/articles/get-started.html

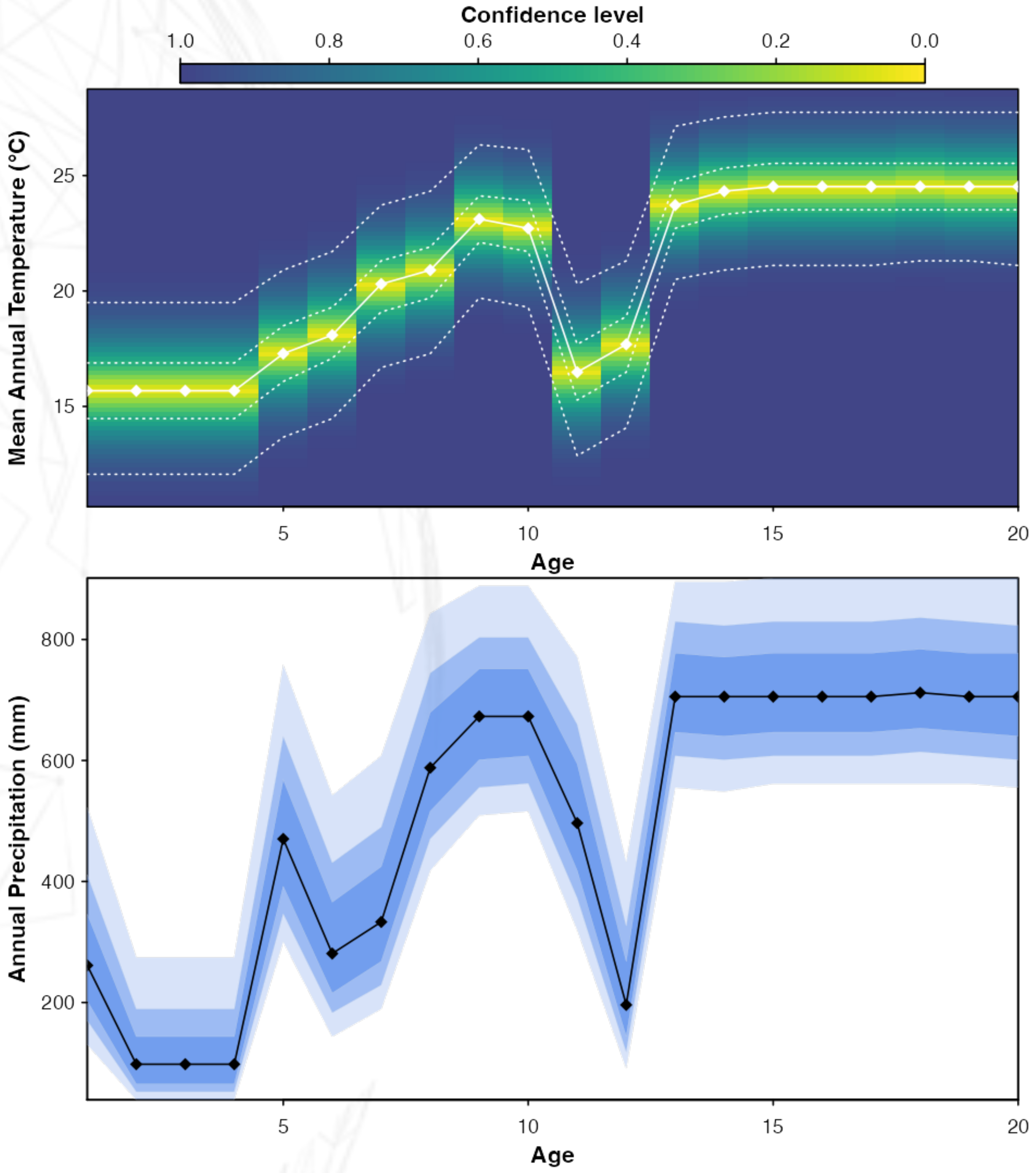










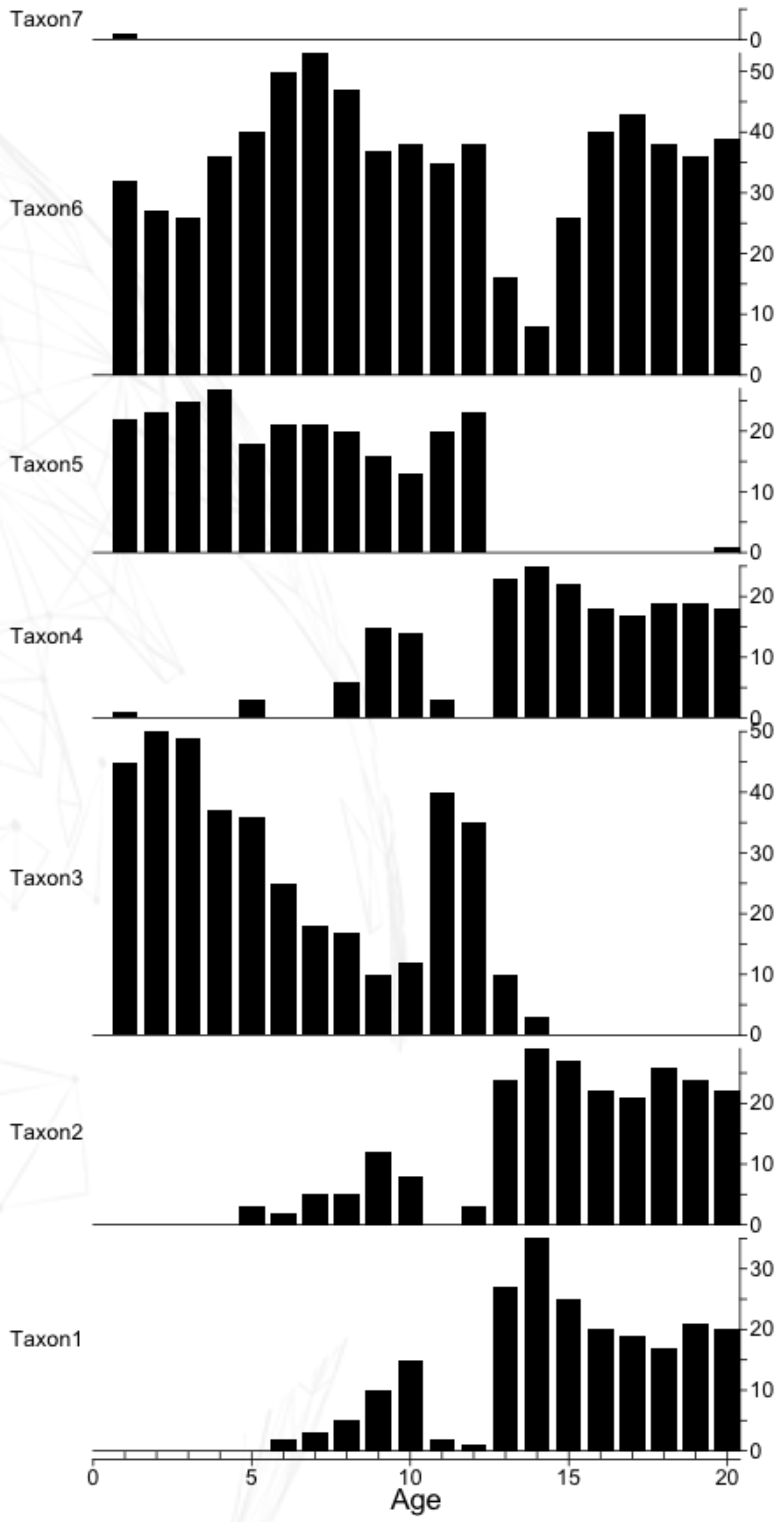


All the necessary data are already included in the package!

 EXAMPLE DATASET

Let's pretend we have some fossil data we want to analyse

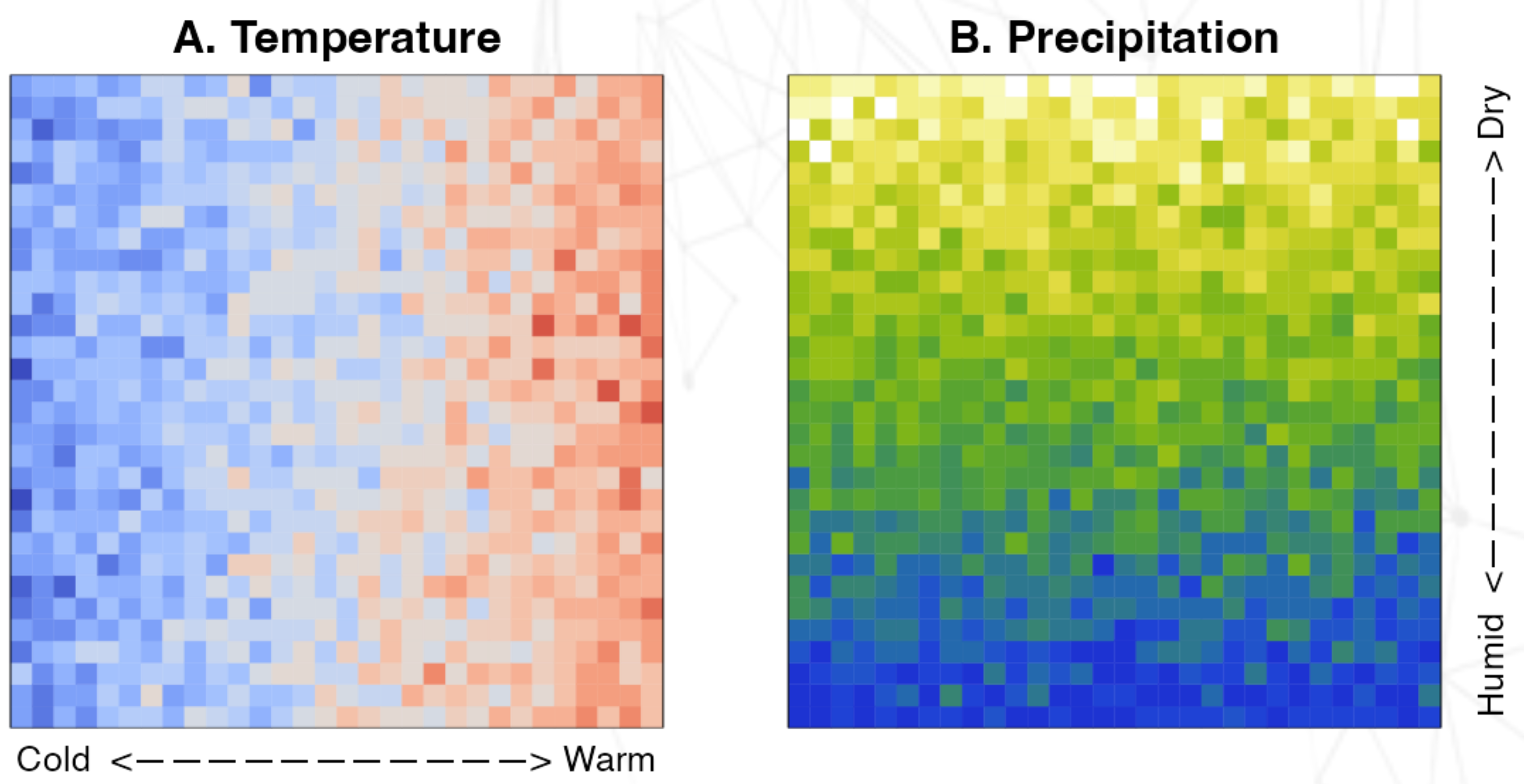
- 7 taxa observed in 20 samples

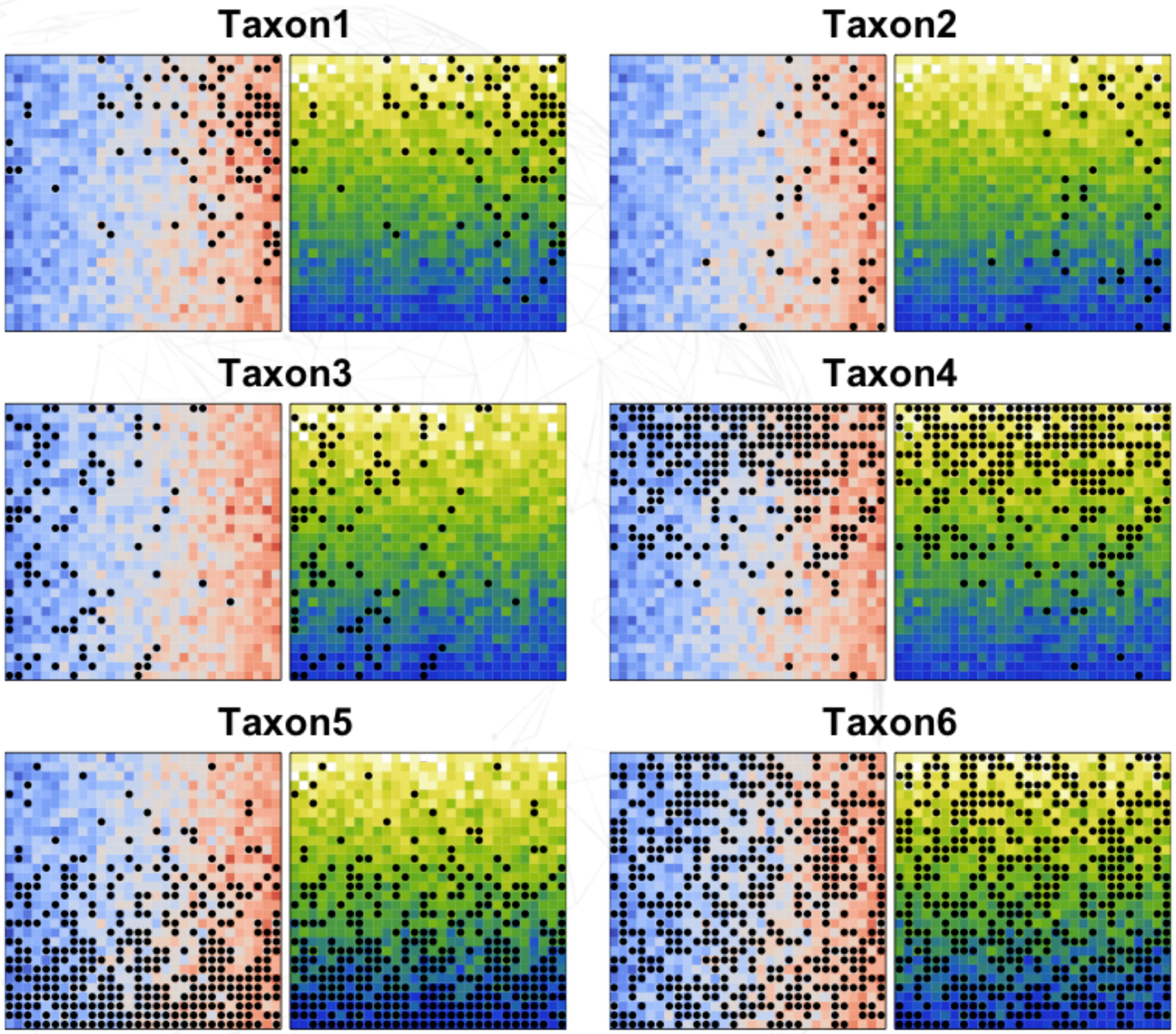


91. EXAMPLE DATASET

And we know the climate associated with their modern distributions

- One issue: No data for Taxon 7

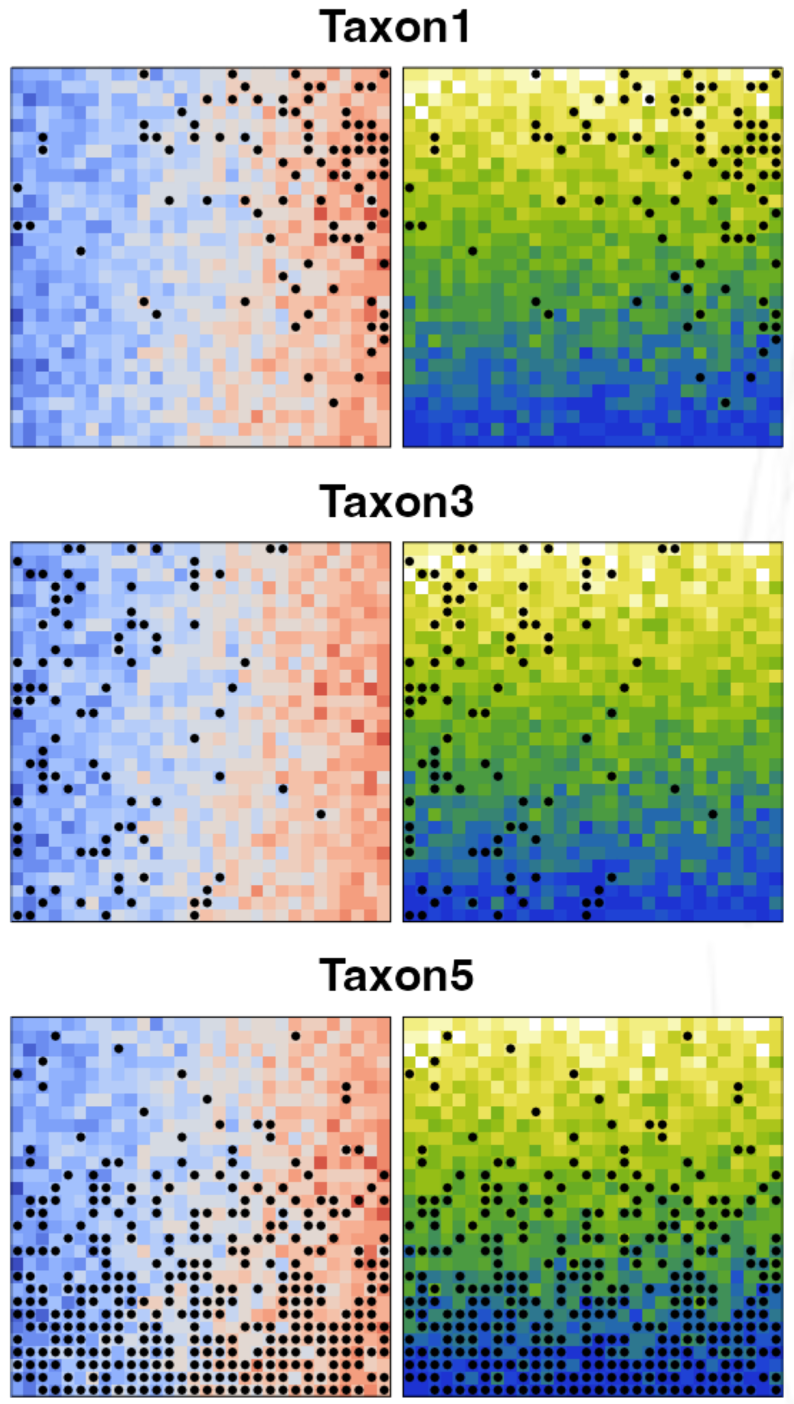


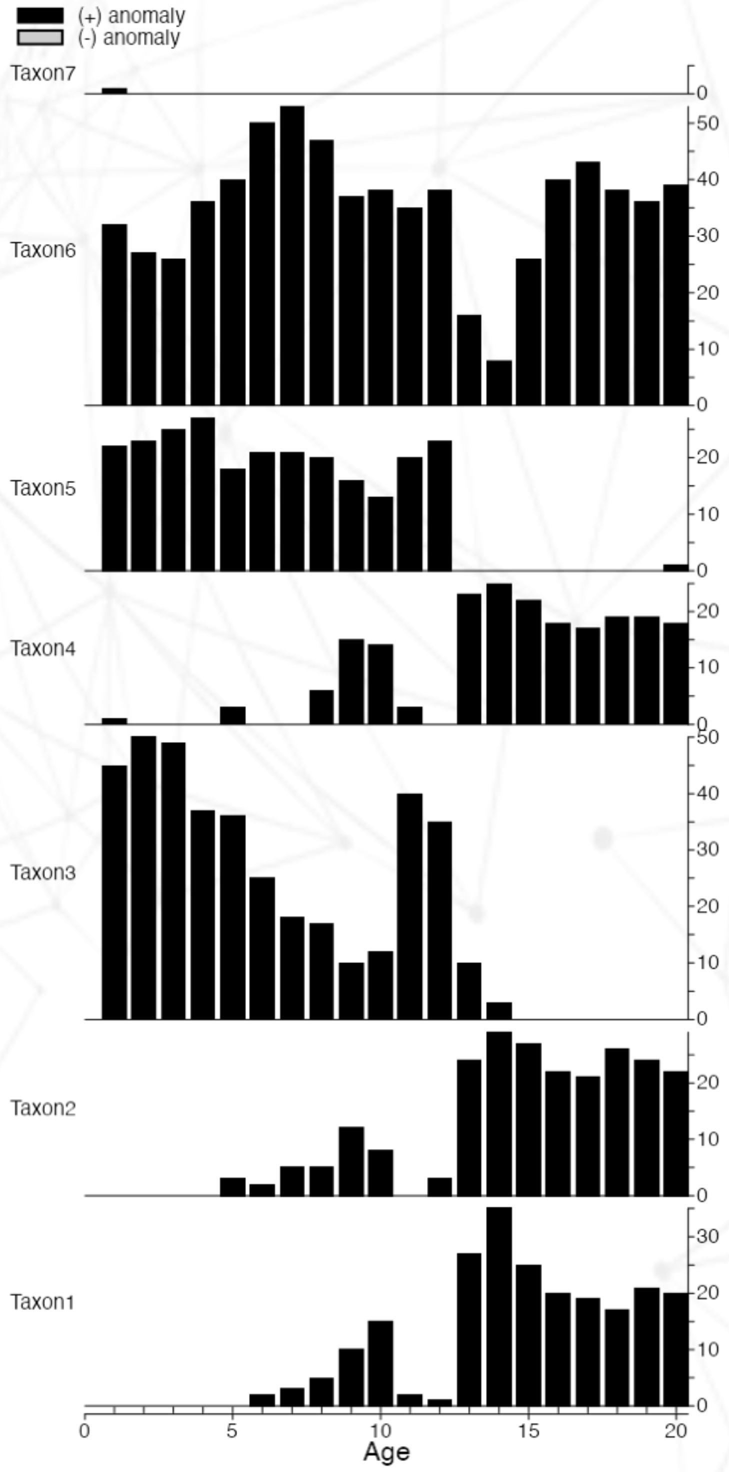


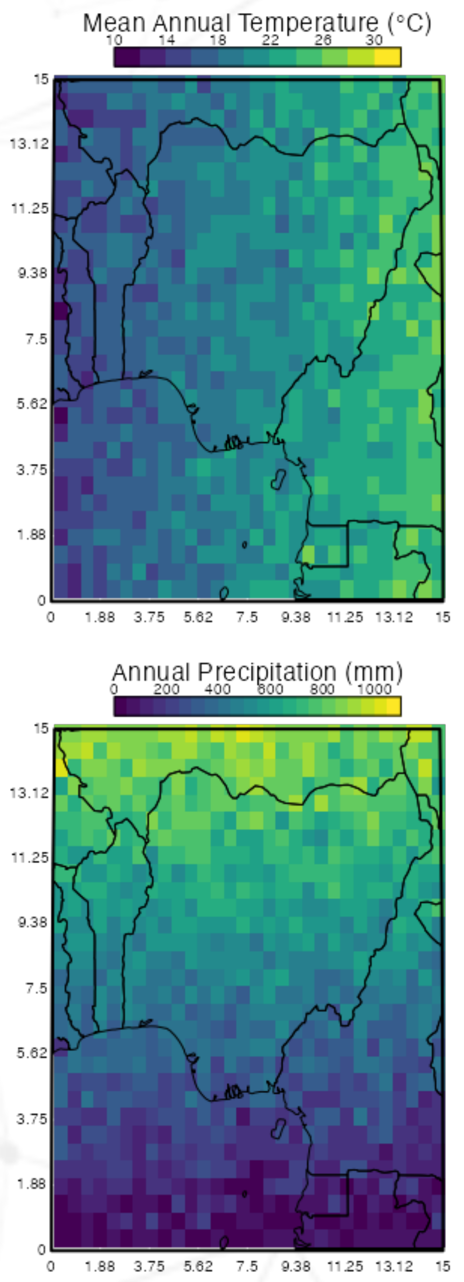
92. YOUR TIME TO SHINE

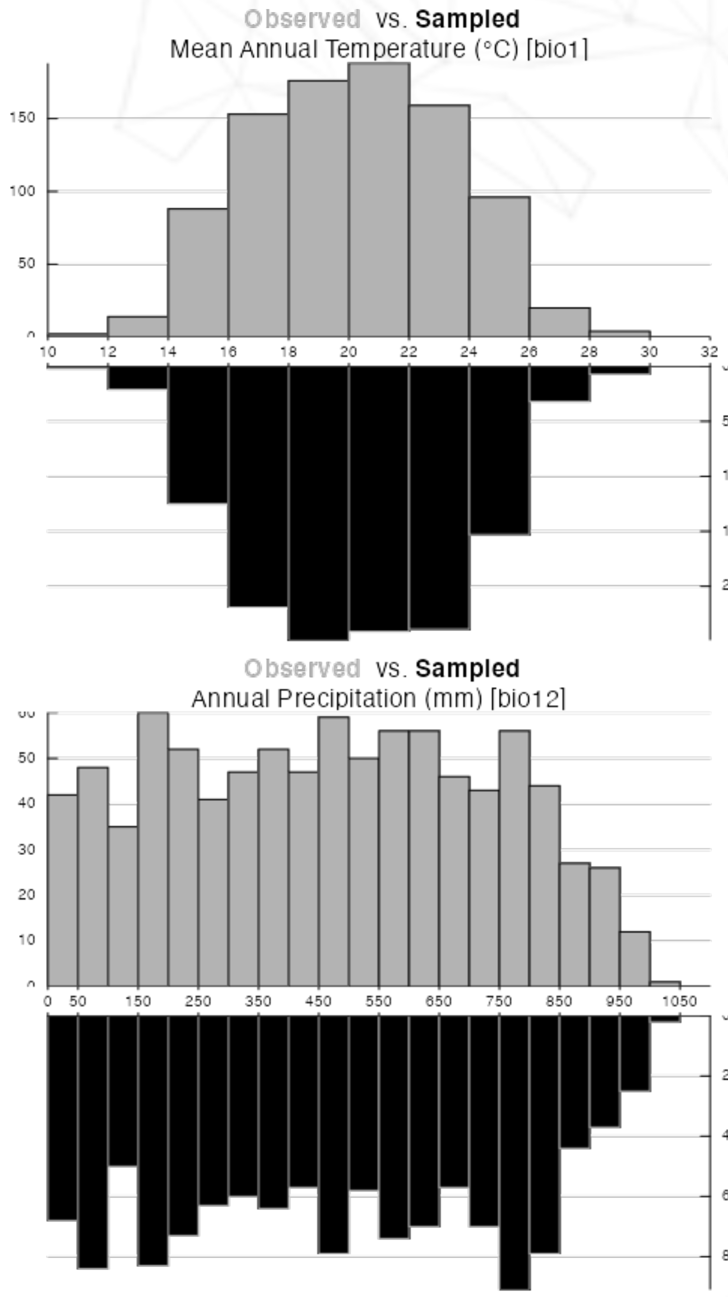
Follow the 'Get started' example from the webpage to get familiar with the crestr environment:

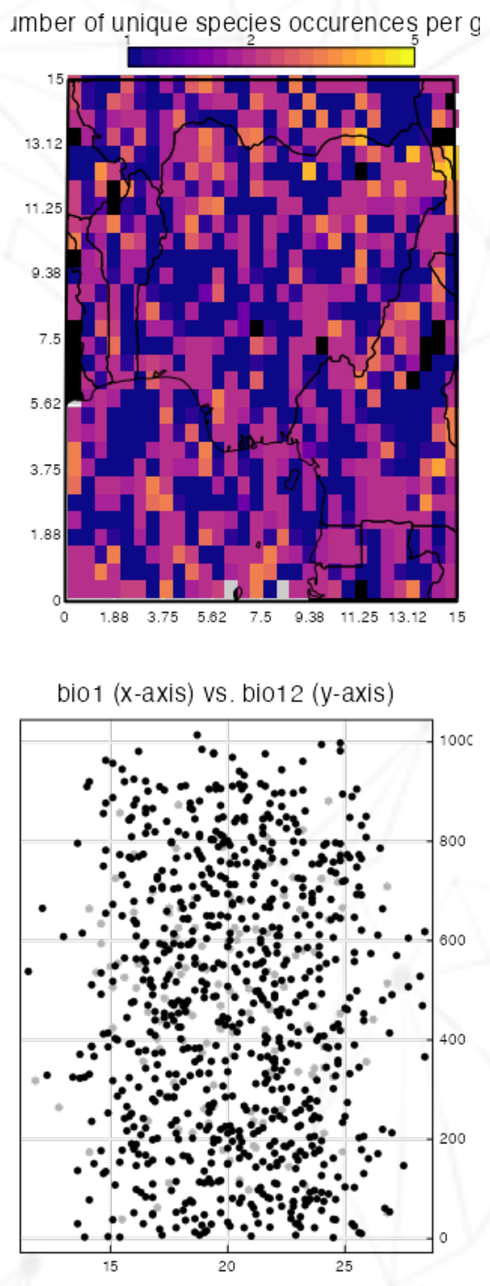
https://www.manuelchevalier.com/crestr/articles/get-started.html

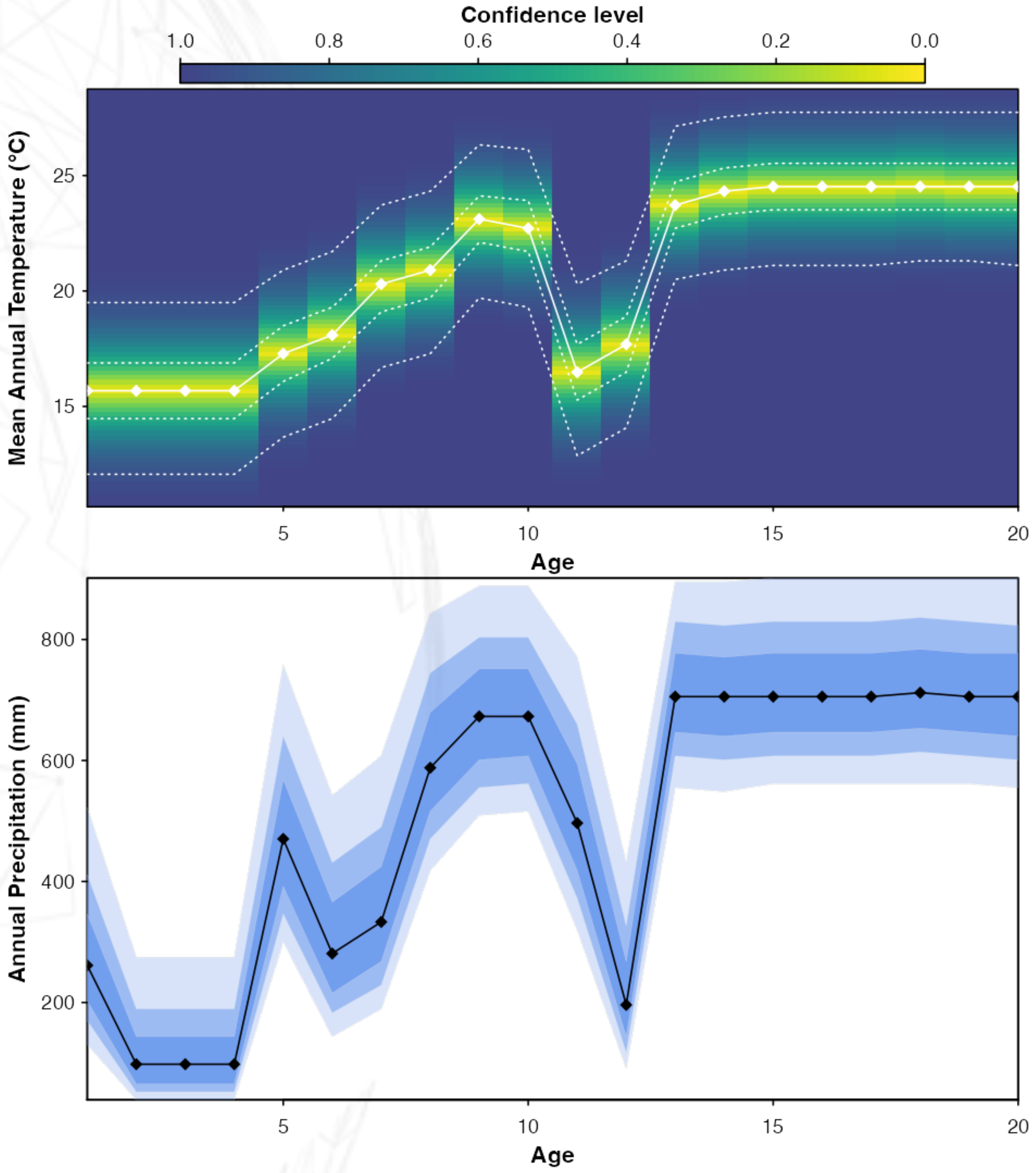












All the necessary data are already included in the package!