Presentation projet

Partie 1 calibration

Partie 1 calibration inverse

- modèles : **PHENOFIT, CASTANEA**

Partie 1 calibration inverse

- modèles : **PHENOFIT, CASTANEA**

- données climatiques : **ERA5-Land**

- données d'occurrence : **EU-Forest** (+GBIF, WOODIV)

Partie 1 calibration inverse

- modèles : **PHENOFIT, CASTANEA**

- données climatiques : **ERA5-Land**

- données d'occurrence : **EU-Forest** (+GBIF, WOODIV)

- algorithme évolutionnaire : **CMA-ES**

Partie 1 calibration inverse

- modèles : **PHENOFIT, CASTANEA**

- données climatiques : **ERA5-Land**

- données d'occurrence : **EU-Forest** (+GBIF, WOODIV)

- algorithme évolutionnaire : **CMA-ES**

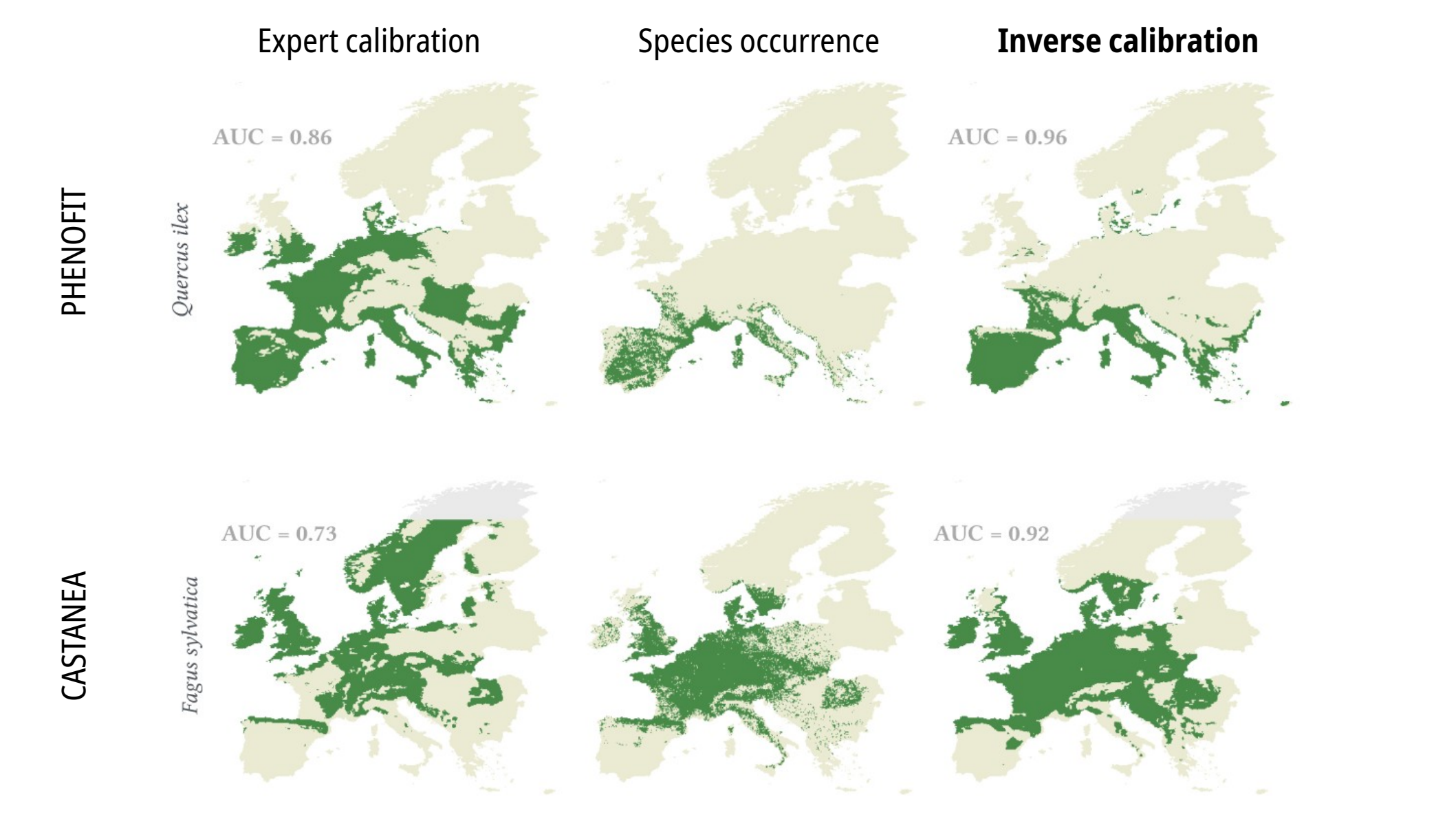
- article méthodo publié dans **Methods in Ecology and Evolution**

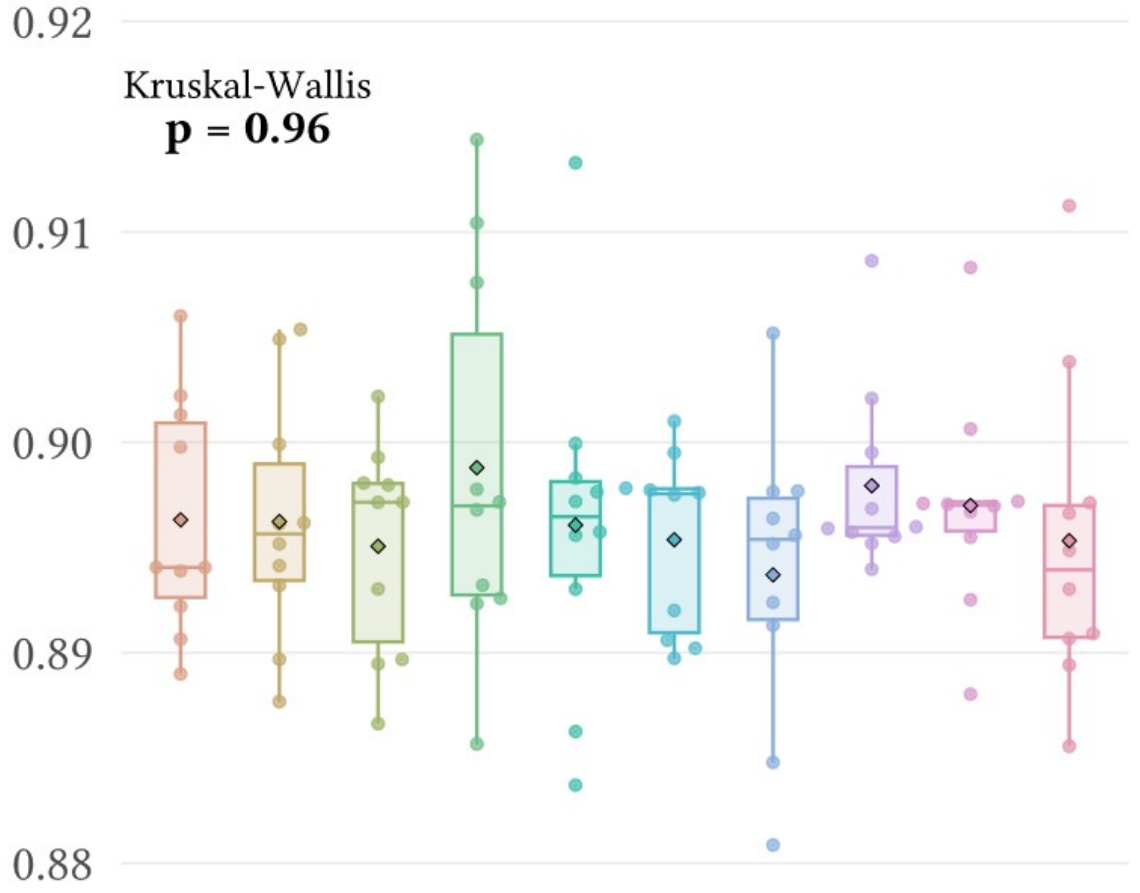
Received: 30 November 2022 | Accepted: 20 April 2023

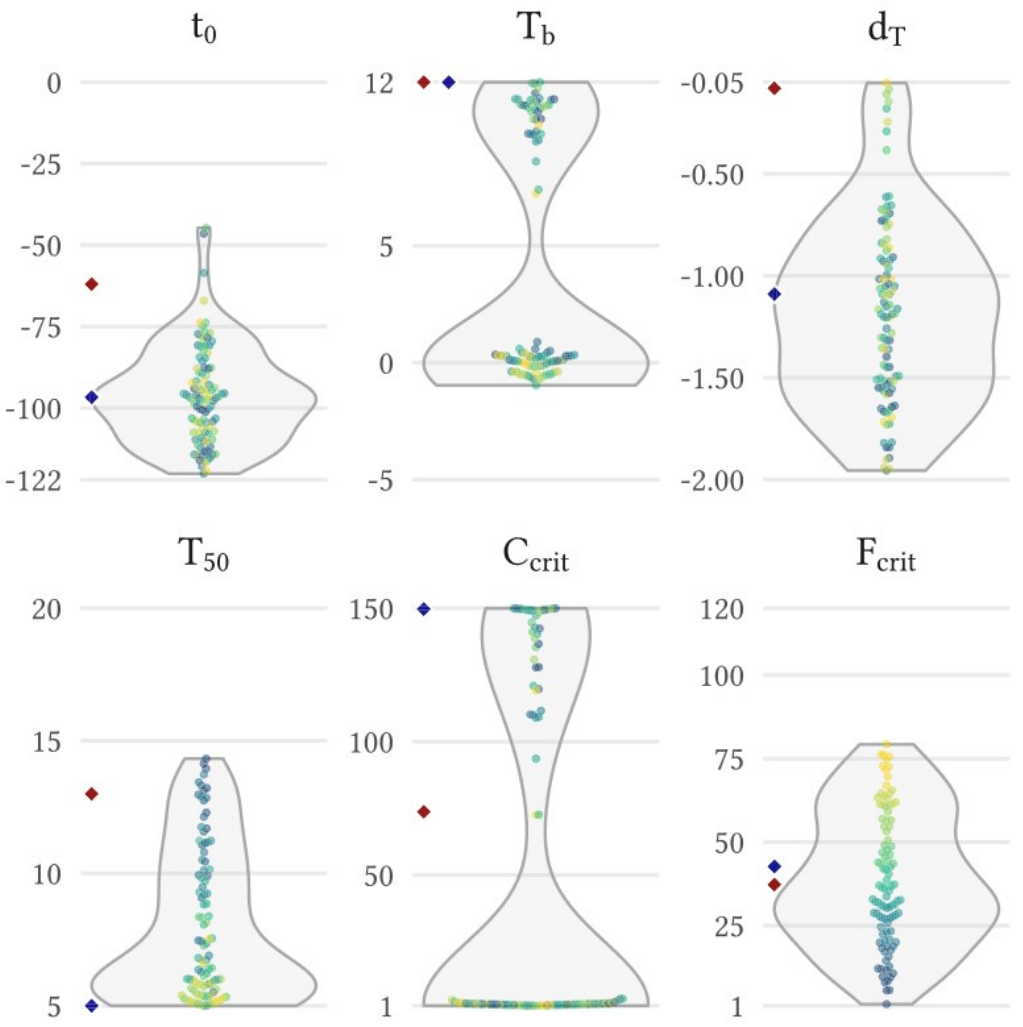
DOI: 10.1111/2041-210X.14119

RESEARCH ARTICLE

Estimating process-based model parameters from species distribution data using the evolutionary algorithm CMA-ES


Expert calibration
Species occurrence
Inverse calibration
PHENOFIT
CASTANEA
Quercus ilex
AUC = 0.86
AUC = 0.96
Fagus sylvatica
AUC = 0.73
AUC = 0.92


0.92
Kruskal-Wallis
p = 0.96
0.91
0.90
0.89
0.88


t_0
0
-25
-50
-75
-100
-122
T_b
12
5
0
-5
d_T
-0.05
-0.50
-1.00
-1.50
-2.00
T_{50}
20
15
10
5
C_{crit}
150
100
50
1
F_{crit}
120
100
75
50
25
1

Partie 1 calibration des modèles corrélatifs

- approche ***presence-only*** inspirée de Valavi et al. (2022)

Ecological Monographs, 92(1), 2022, e01486

Predictive performance of presence-only species distribution models: a benchmark study with reproducible code

Roozbeh Valavi,[1,3] Gurutzeta Guillera-Arroita,[2] José J. Lahoz-Monfort,[2] and Jane Elith[2]

Partie 1 calibration des modèles corrélatifs

- approche ***presence-only*** inspirée de Valavi et al. (2022)

Ecological Monographs, 92(1), 2022, e01486

Predictive performance of presence-only species distribution models: a benchmark study with reproducible code

Roozbeh Valavi,[1,3] Gurutzeta Guillera-Arroita,[2] José J. Lahoz-Monfort,[2] and Jane Elith[2]

- données d'occurrence : **EU-Forest** (+GBIF, WOODIV)

- **background points** : « *sampled irrespective of the location of species records* »
50.000 points

Partie 1 calibration des modèles corrélatifs

- approche ***presence-only*** inspirée de Valavi et al. (2022)

Ecological Monographs, 92(1), 2022, e01486

Predictive performance of presence-only species distribution models: a benchmark study with reproducible code

Roozbeh Valavi,[1,3] Gurutzeta Guillera-Arroita,[2] José J. Lahoz-Monfort,[2] and Jane Elith[2]

- données d'occurrence : **EU-Forest** (+GBIF, WOODIV)

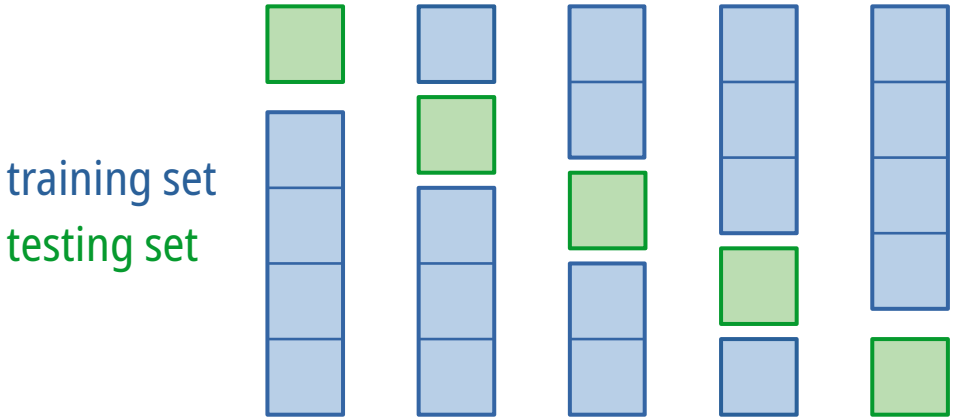
- **background points** : « *sampled irrespective of the location of species records* »
50.000 points

- modèles : GLM with lasso penalty, GAM, BRT, RandomForest
(+ biomod et Maxent)

- ***5-fold environmental block cross-validation*** pour estimer les erreurs

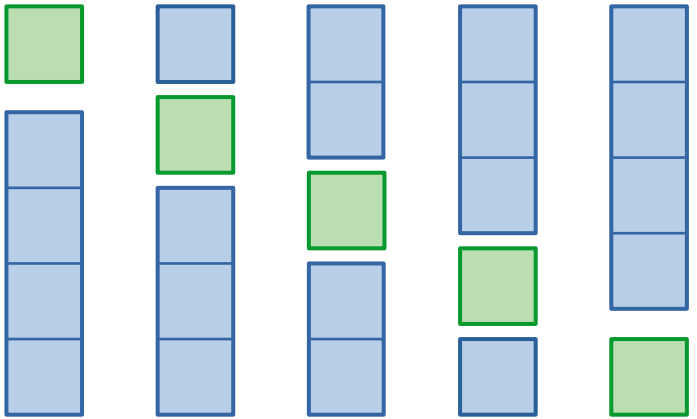
- 5-fold environmental block cross-validation pour estimer les erreurs

- 5-fold environmental block cross-validation pour estimer les erreurs



- 5-fold environmental block cross-validation pour estimer les erreurs

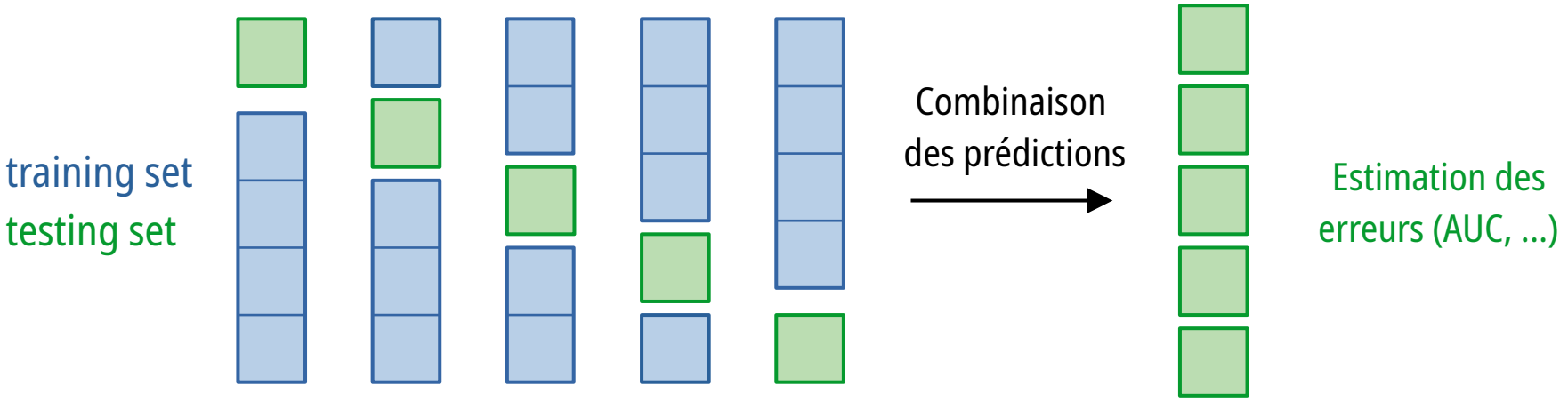
training set
testing set



Combinaison
des prédictions

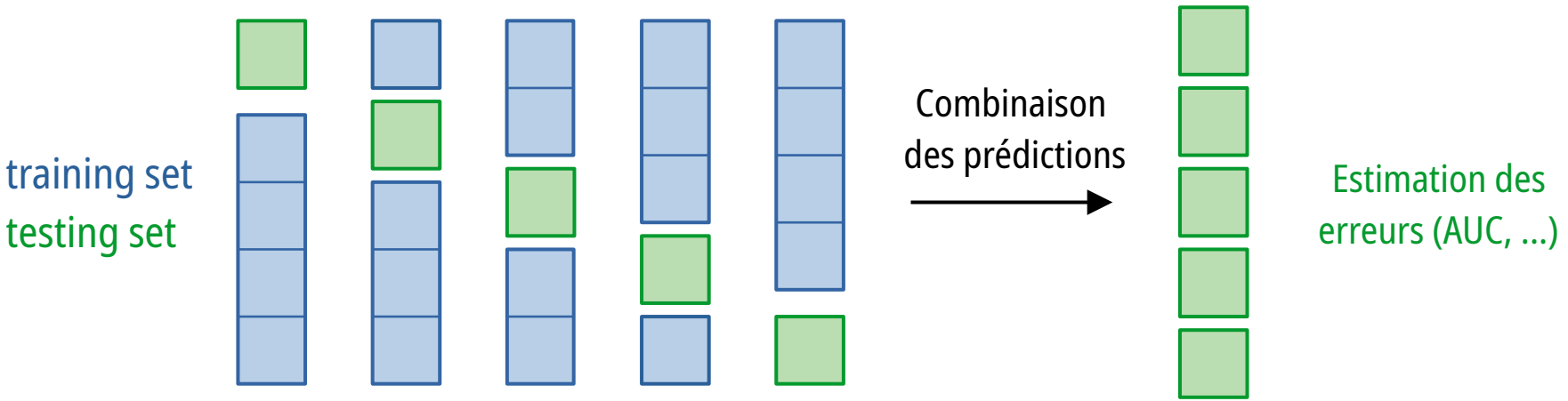
Estimation des
erreurs (AUC, ...)

- ***5-fold environmental block cross-validation*** pour estimer les erreurs



- ***final model*****:** *« all the available training data can be used to fit a new model »*
Favorise la qualité des prédictions (plutôt que l'exactitude des erreurs)

- *5-fold environmental block cross-validation* pour estimer les erreurs



- *final model*: *« all the available training data can be used to fit a new model »*
Favorise la qualité des prédictions (plutôt que l'exactitude des erreurs)

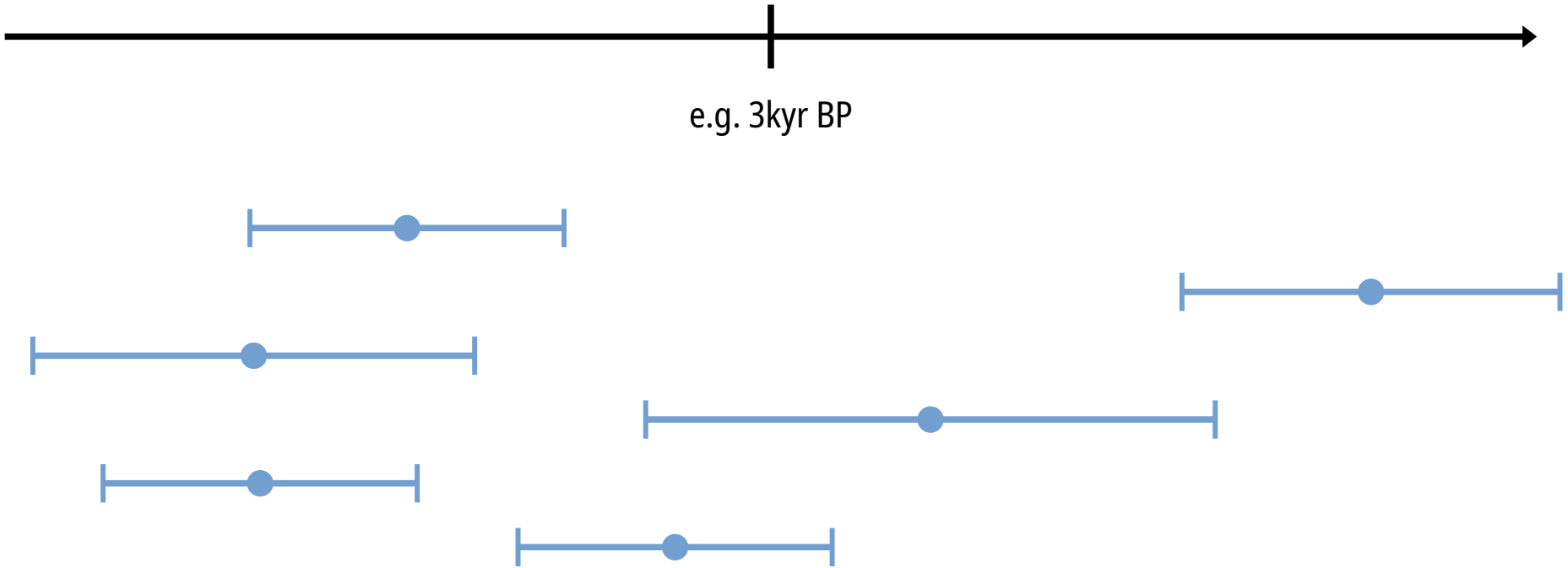
Over-estimated
AUC

Exemple avec *Fagus sylvatica*

	Lasso GLM	GAM	Random Forest	BRT
« True » AUC	0.68	0.79	0.81	0.79
Over-estimated AUC	0.92	0.93	0.99	0.96

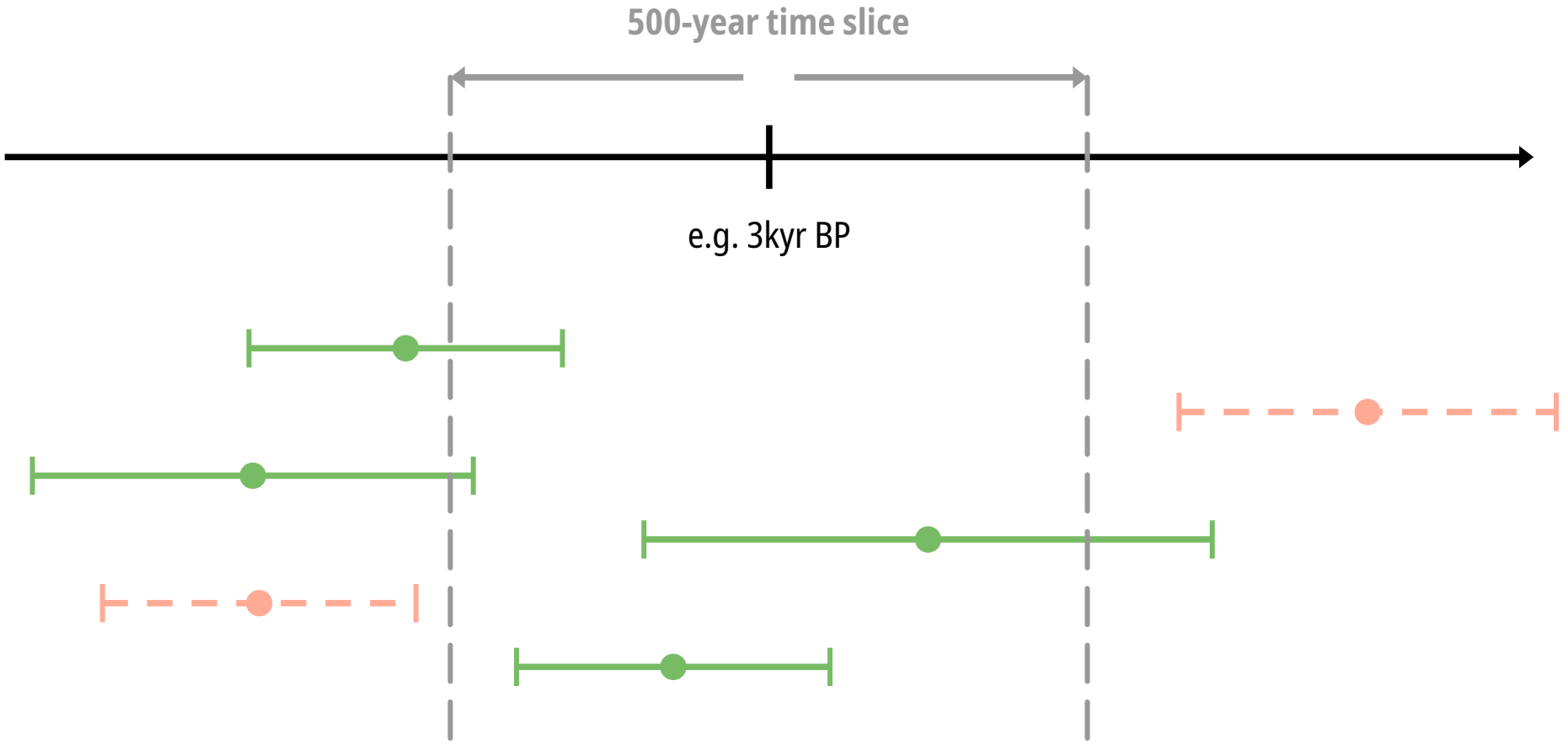
Partie 2 paleo

Temporal coverage per pollen site

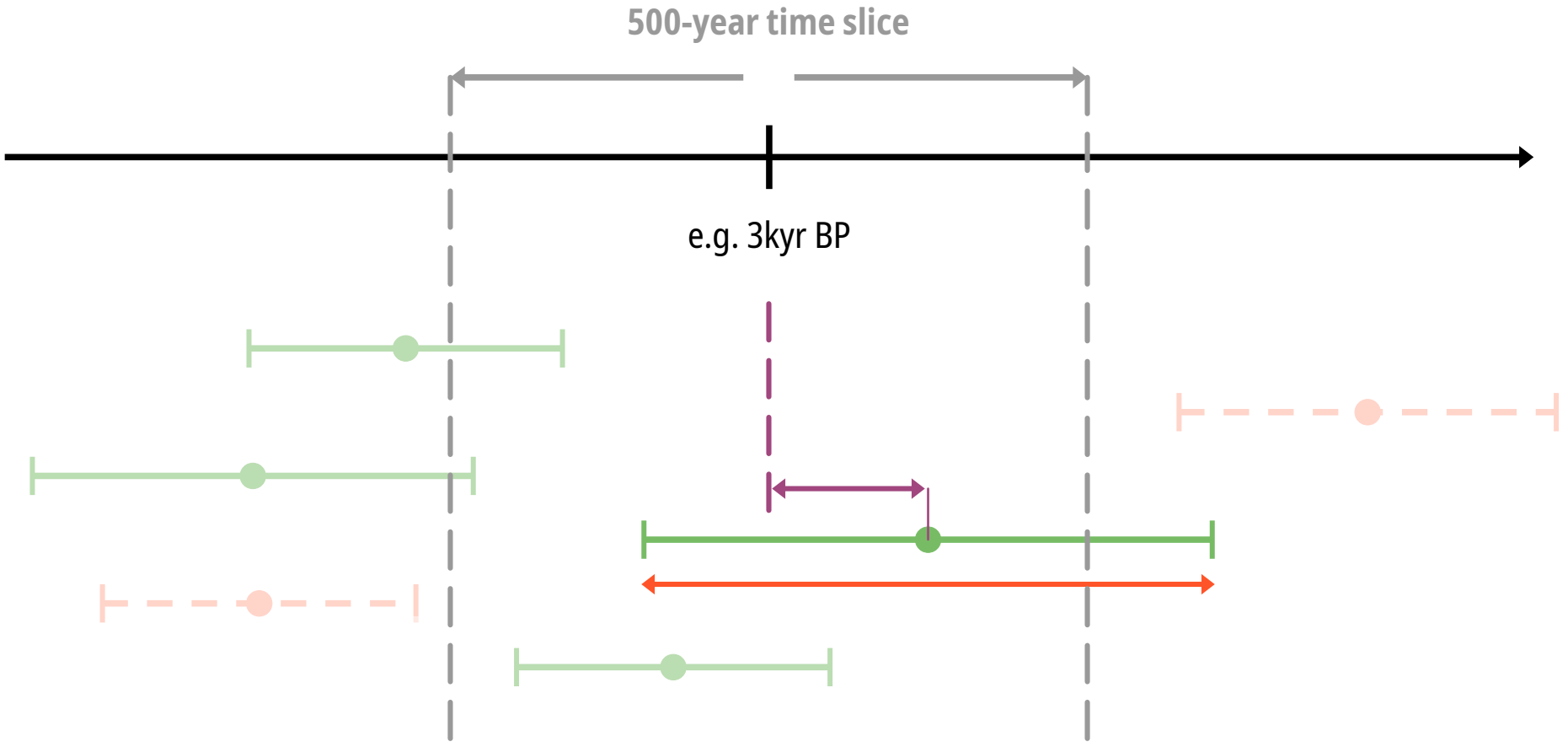


Multiple samples within the same site

Temporal coverage per pollen site

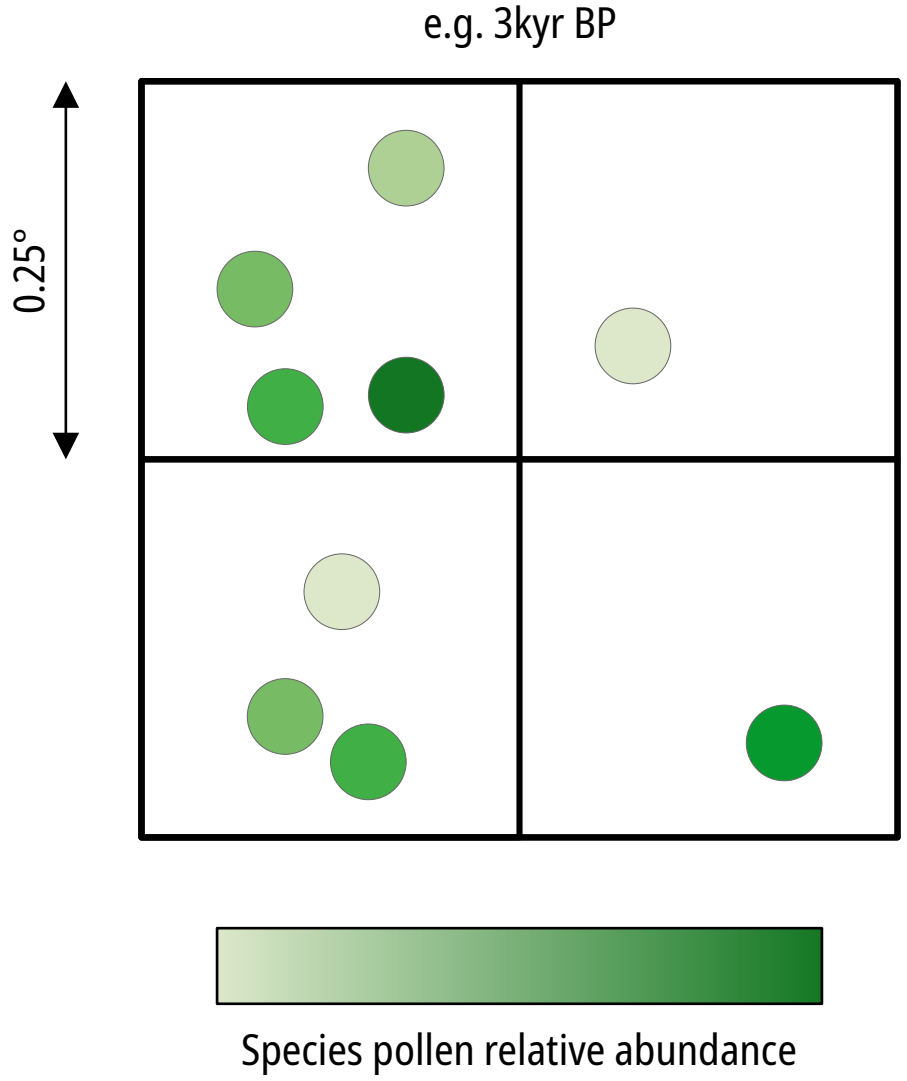


Temporal coverage per pollen site

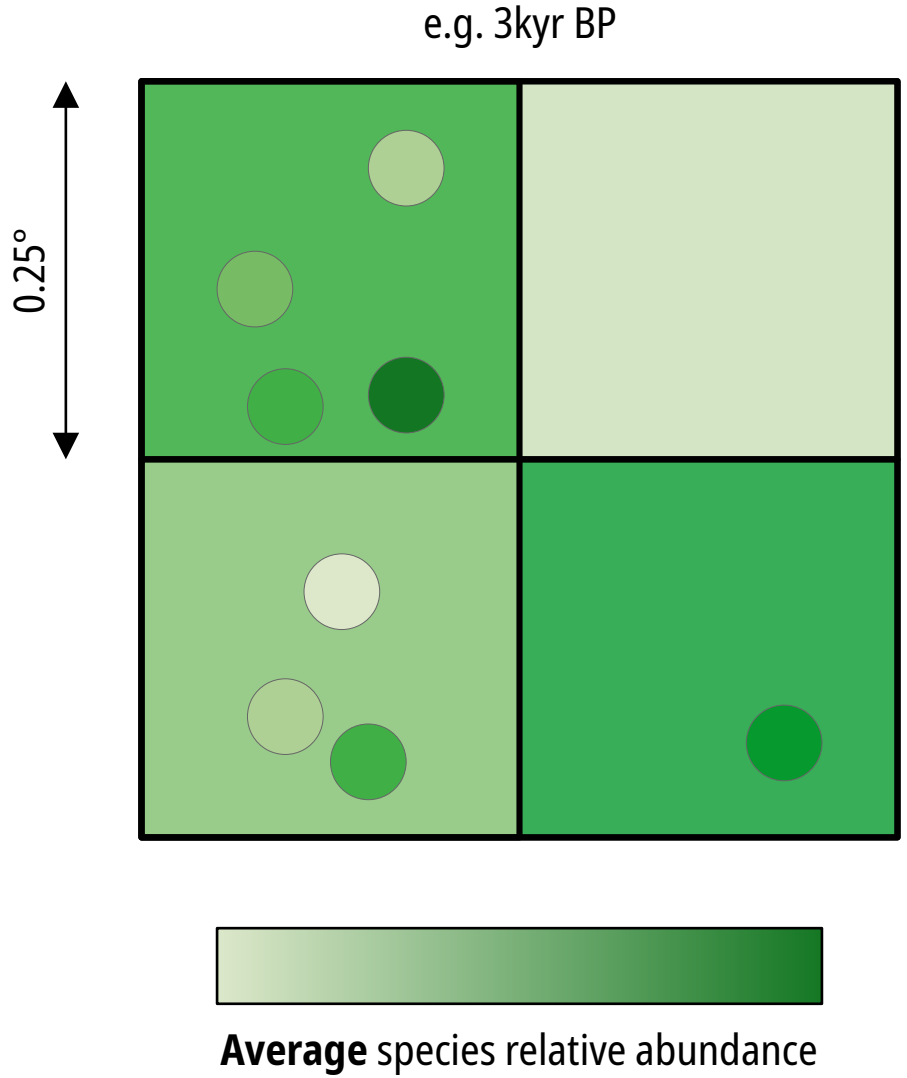


Weighted mean of pollen counts, taking into account both **uncertainty** and **time distance**

Spatial upscaling

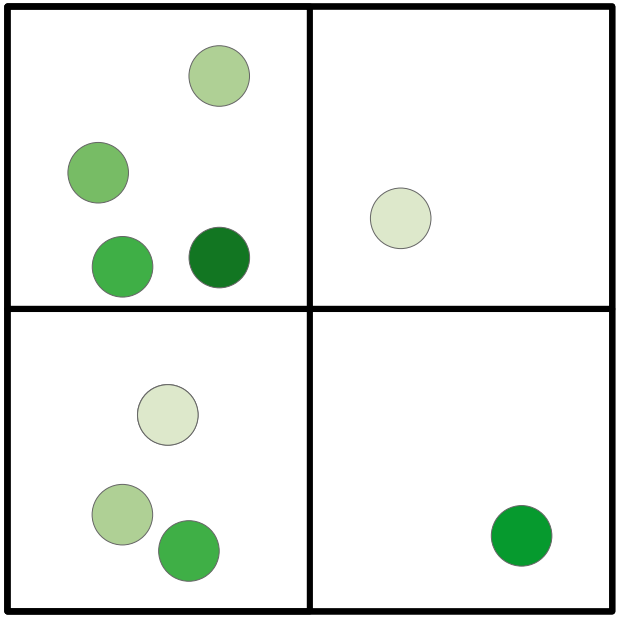


Spatial upscaling: relative abundance

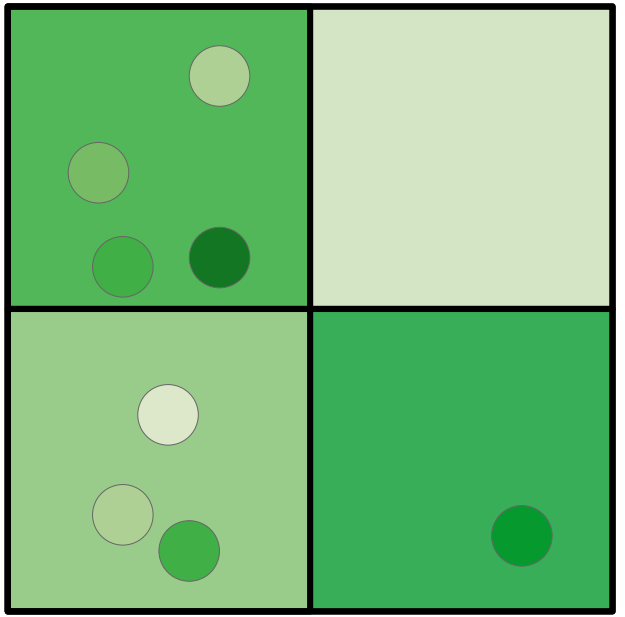


Spatial upscaling: presence/absence

« one is enough »



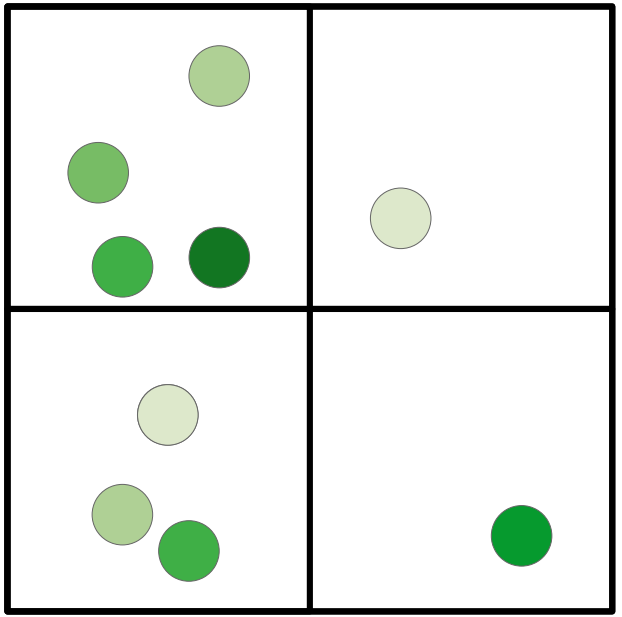
average presence



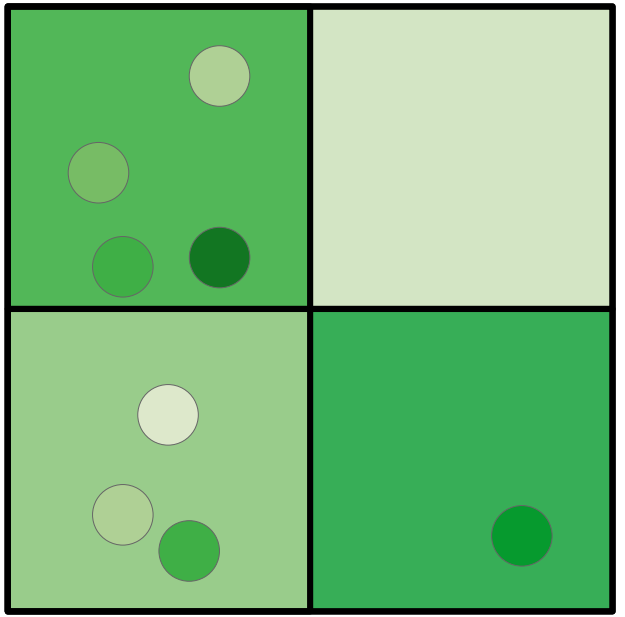
Presence threshold

Spatial upscaling: presence/absence

« one is enough »



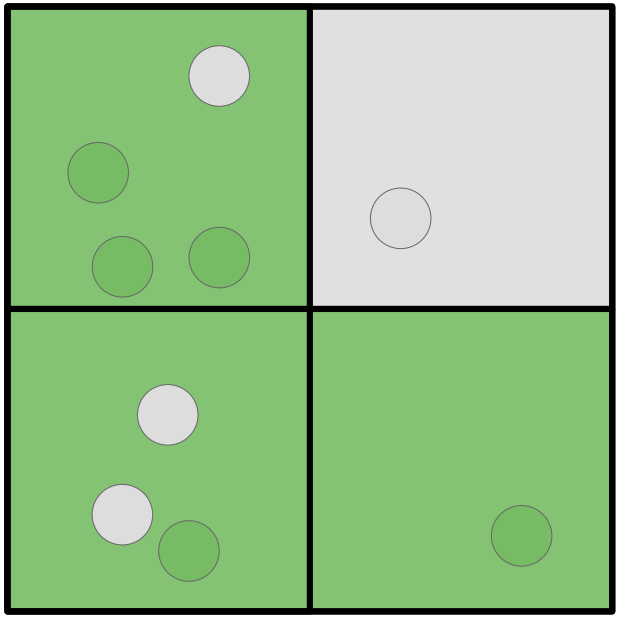
average presence



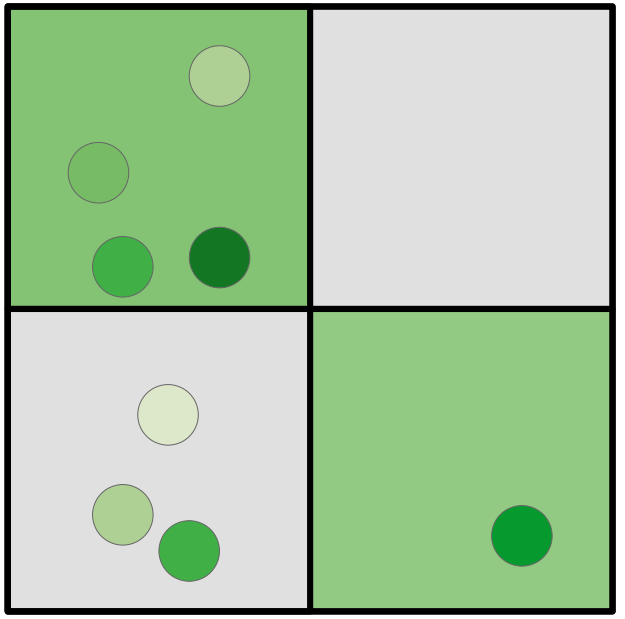
Presence threshold

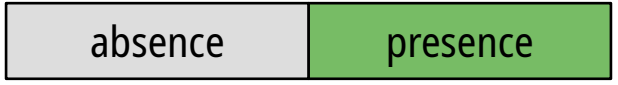
Spatial upscaling: presence/absence

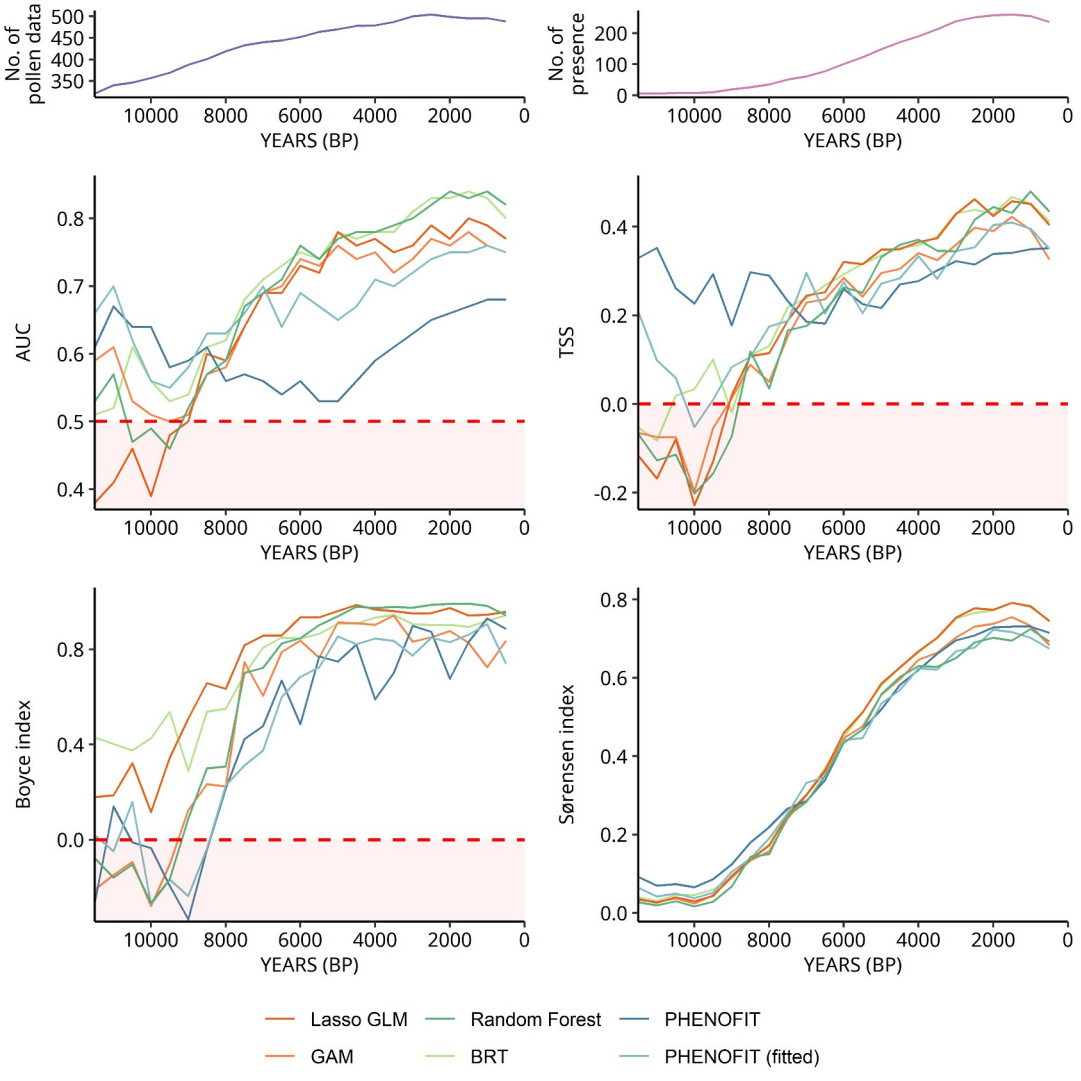
« one is enough »



average presence



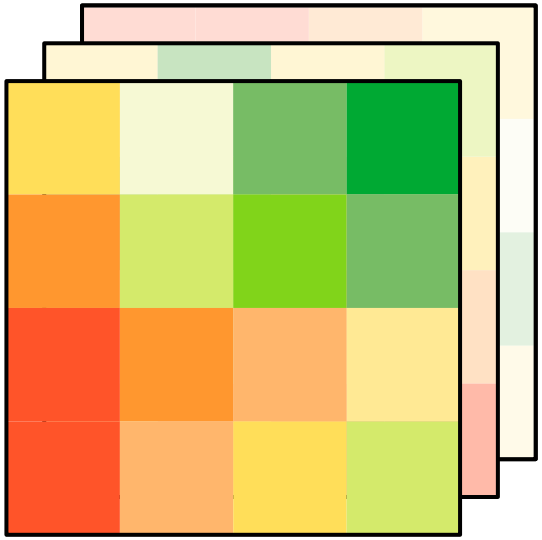



No. of pollen data
500
450
400
350
10000 8000 6000 4000 2000 0
YEARS (BP)
No. of presence
200
100
0
10000 8000 6000 4000 2000 0
YEARS (BP)
AUC
0.8
0.7
0.6
0.5
0.4
10000 8000 6000 4000 2000 0
YEARS (BP)
TSS
0.4
0.2
0.0
-0.2
10000 8000 6000 4000 2000 0
YEARS (BP)
Boyce index
0.8
0.4
0.0
10000 8000 6000 4000 2000 0
YEARS (BP)
Sørensen index
0.8
0.6
0.4
0.2
0.0
10000 8000 6000 4000 2000 0
YEARS (BP)
Lasso GLM
GAM
Random Forest
BRT
PHENOFIT
PHENOFIT (fitted)

CLIMATE APPROACH: Climatic distance

PAST CLIMATE

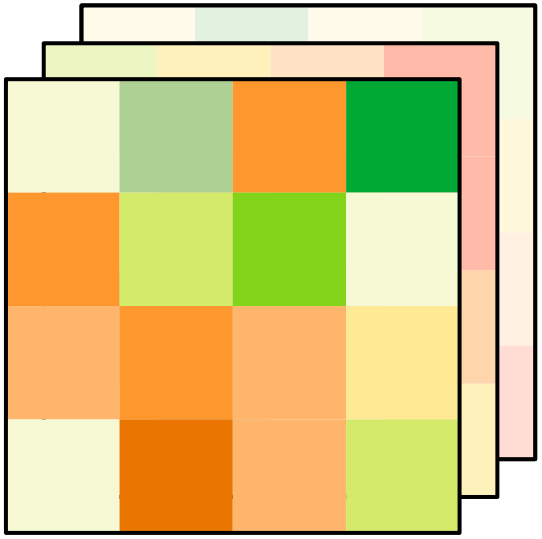
e.g. 9000 years BP



N predictors

BASELINE CLIMATE

1970 - 2000

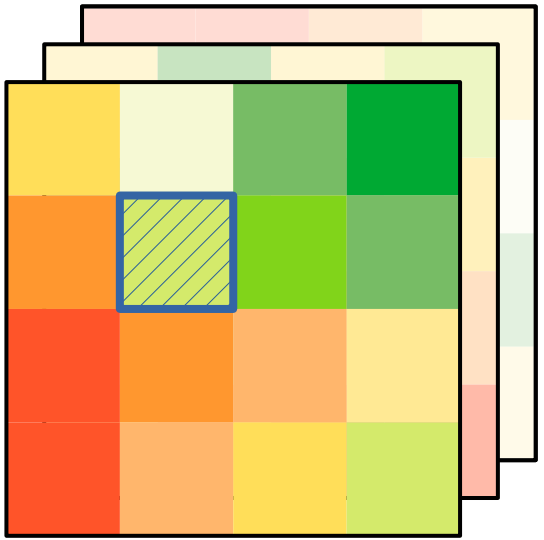


N predictors

CLIMATE APPROACH: Climatic distance

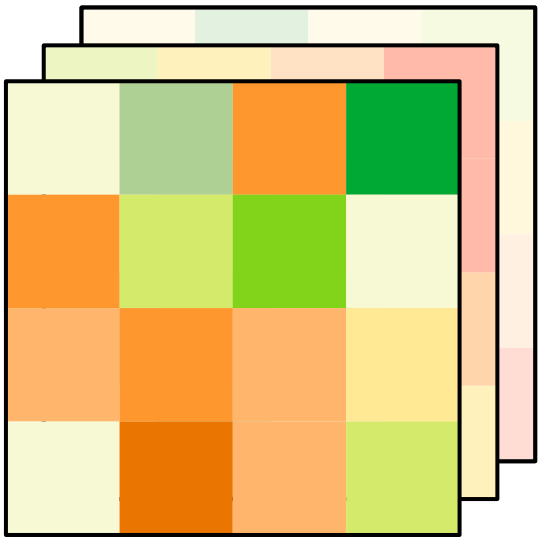
PAST CLIMATE

e.g. 9000 years BP



BASELINE CLIMATE

1970 - 2000

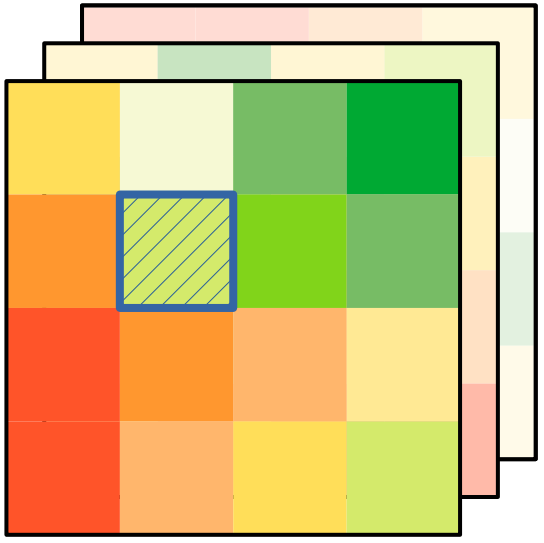


From each cell:

CLIMATE APPROACH: Climatic distance

PAST CLIMATE

e.g. 9000 years BP



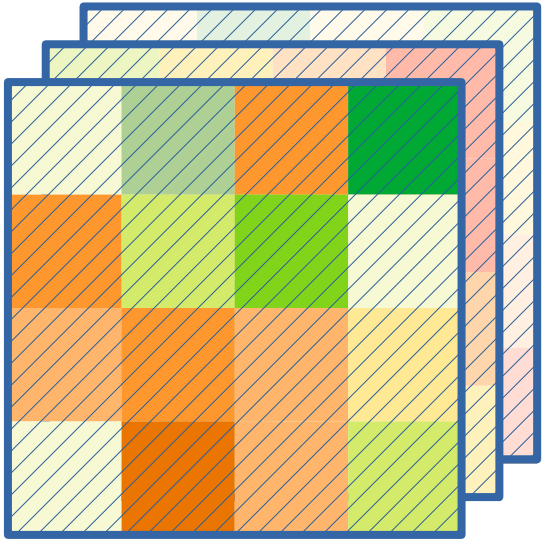
Minimum

Euclidean distance

(N-dimensional)

BASELINE CLIMATE

1970 - 2000



From each cell:

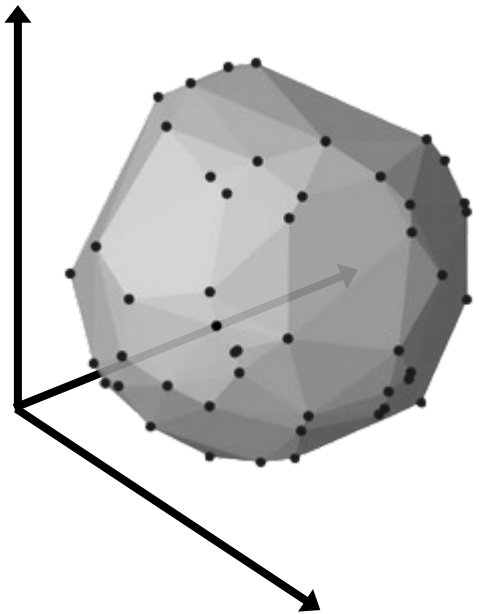
{ ; ; ... }

To the set of all cells
in the calibration period

CLIMATE APPROACH: Hypervolume similarity

PAST CLIMATE

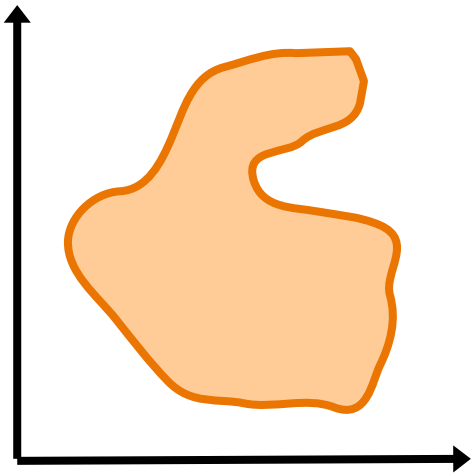
e.g. 9000 years BP



N-dimensional hypervolume

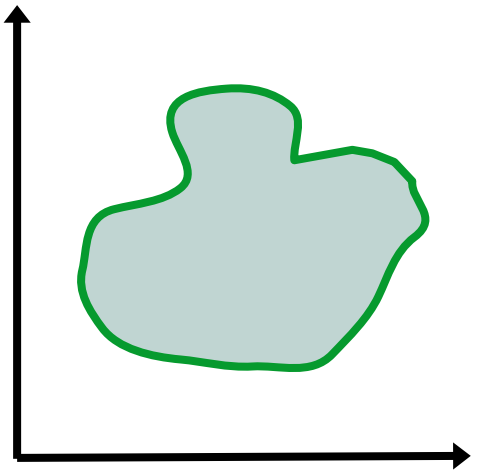
CLIMATE APPROACH: Hypervolume similarity

PAST CLIMATE
e.g. 9000 years BP



N-dimensional hypervolume

BASELINE CLIMATE
1970 - 2000

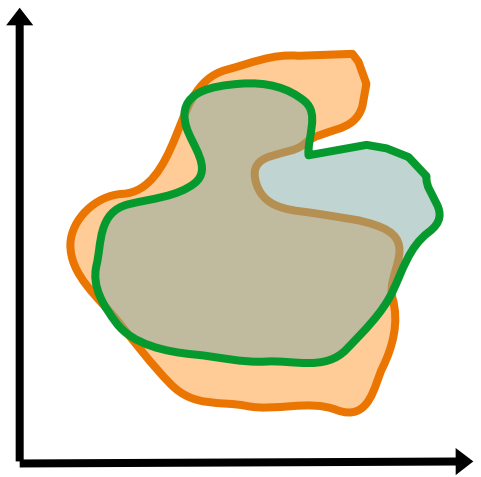


N-dimensional hypervolume

CLIMATE APPROACH: Hypervolume similarity

PAST CLIMATE
e.g. 9000 years BP

BASELINE CLIMATE
1970 - 2000

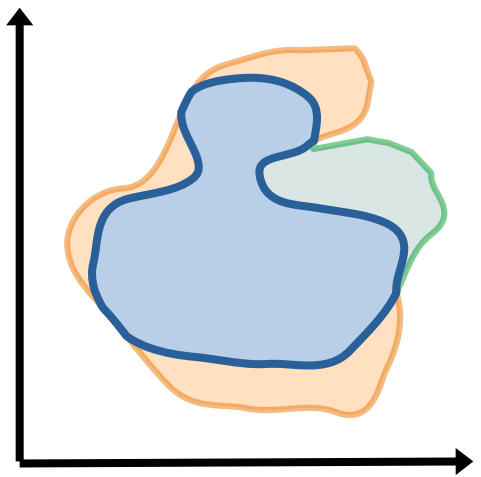


CLIMATE APPROACH: Hypervolume similarity

PAST CLIMATE
e.g. 9000 years BP

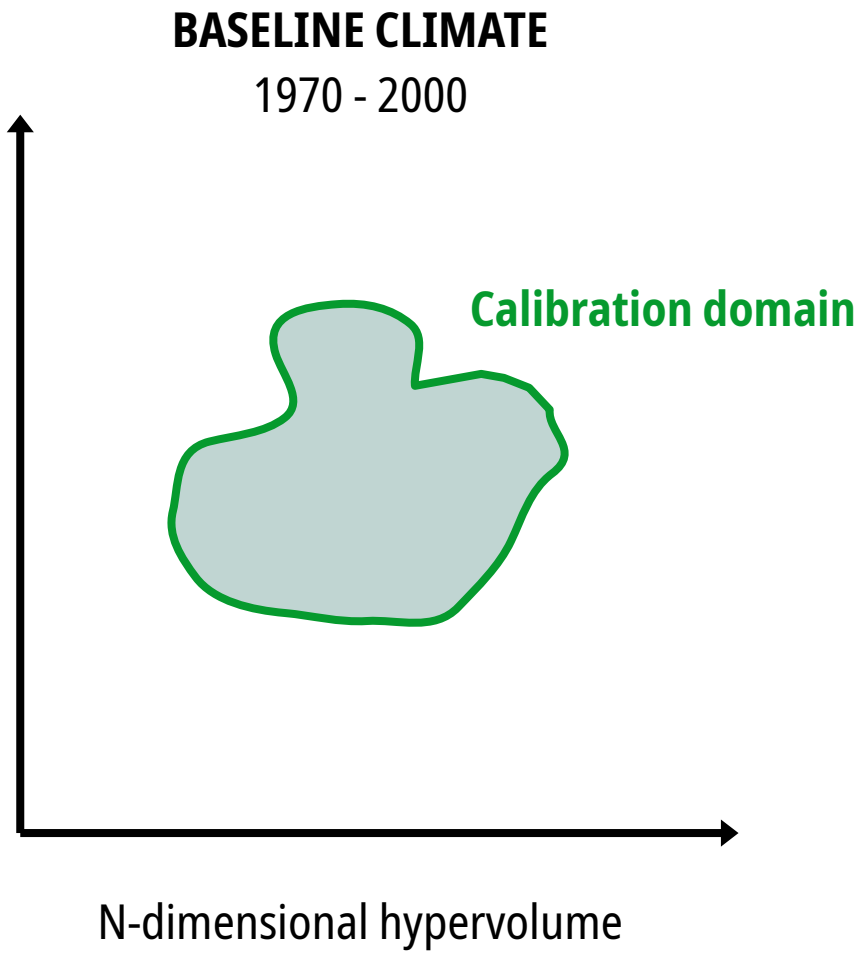
BASELINE CLIMATE
1970 - 2000

Hypervolume overlap

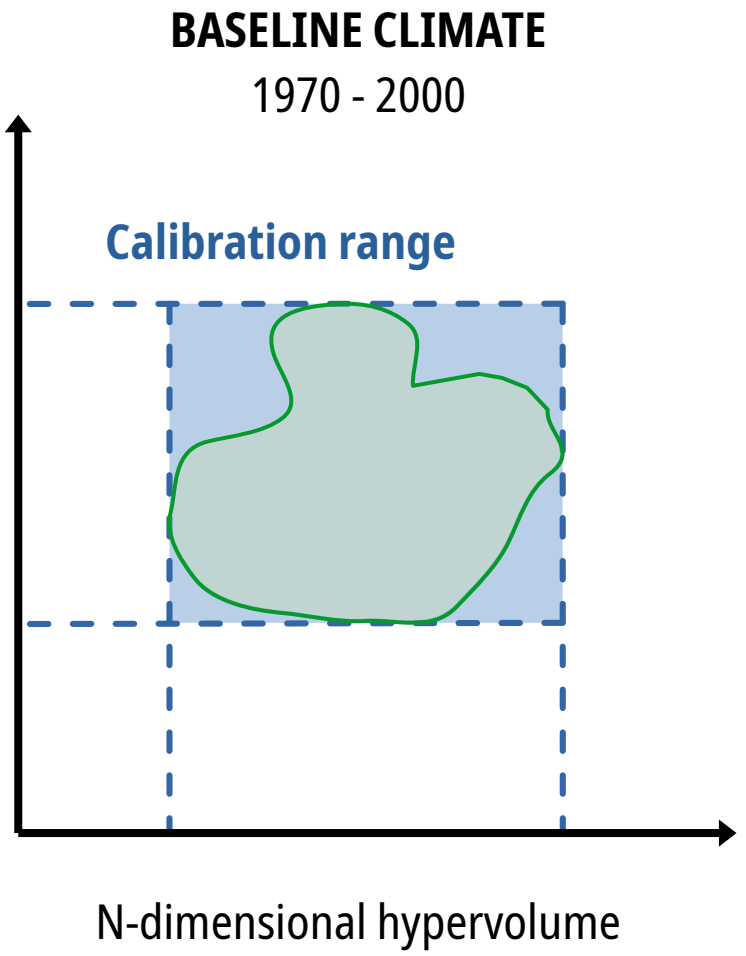


(Sørensen similarity)

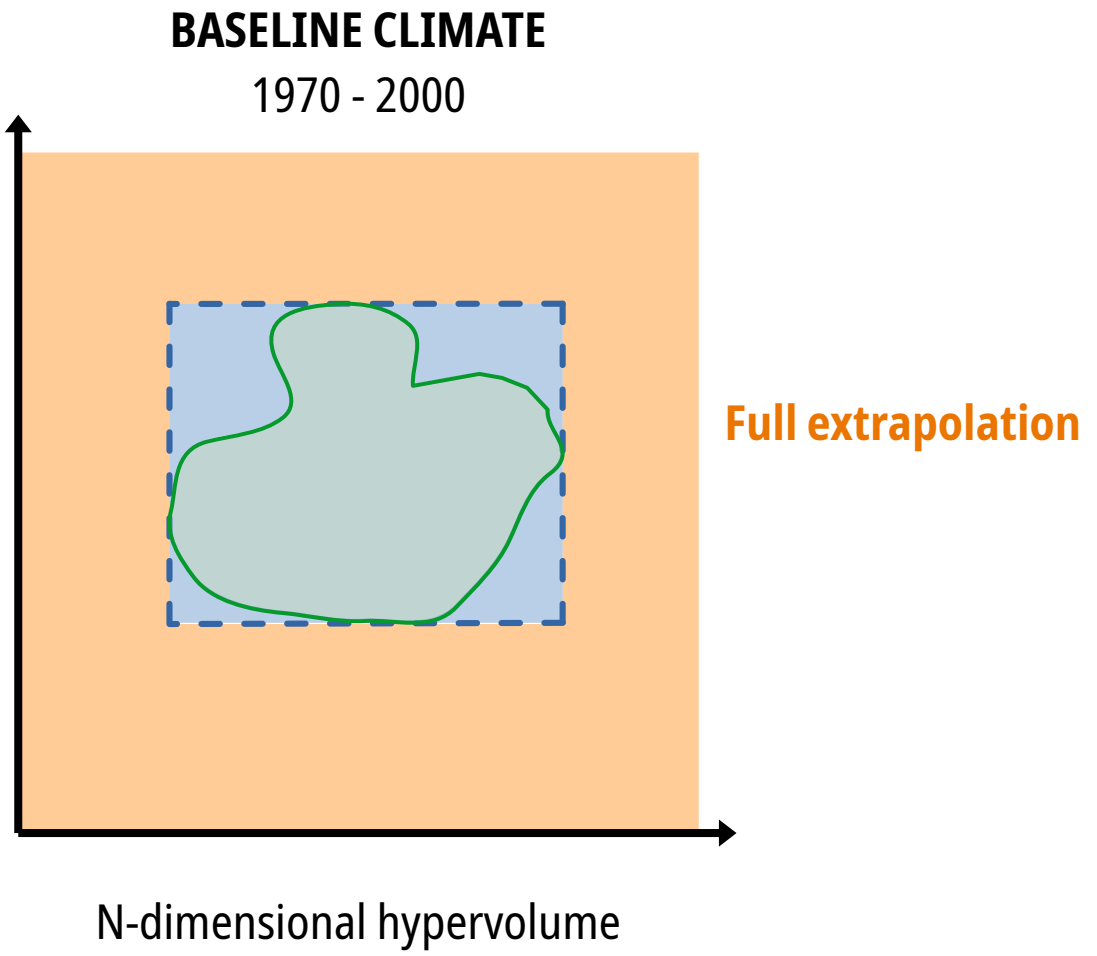
CLIMATE APPROACH: Categorical differenciation



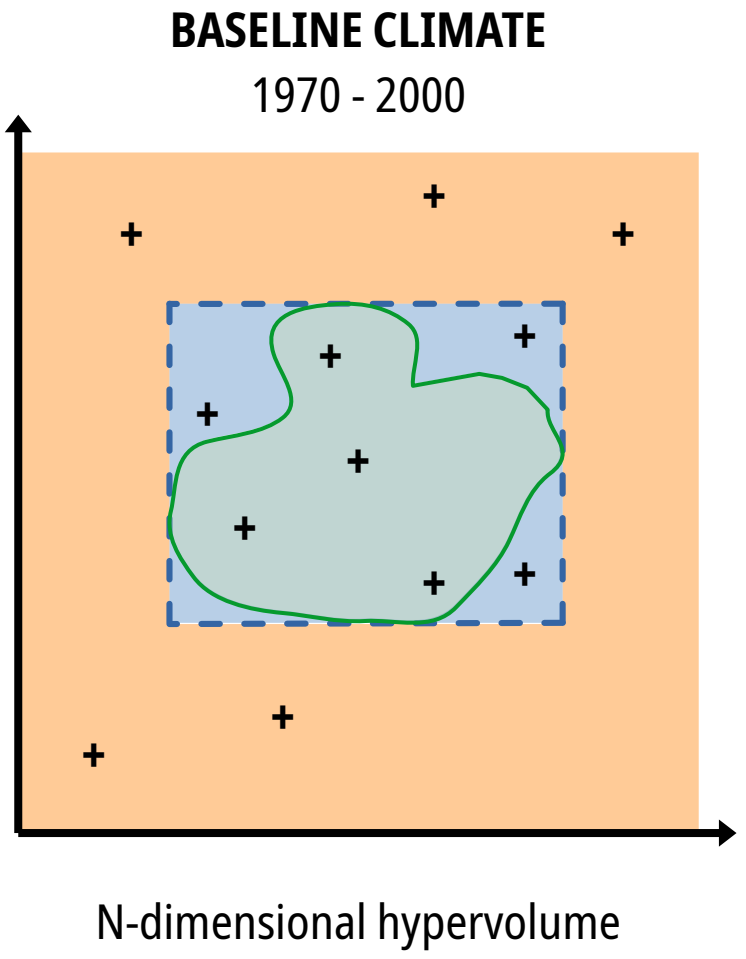
CLIMATE APPROACH: Categorical differenciation



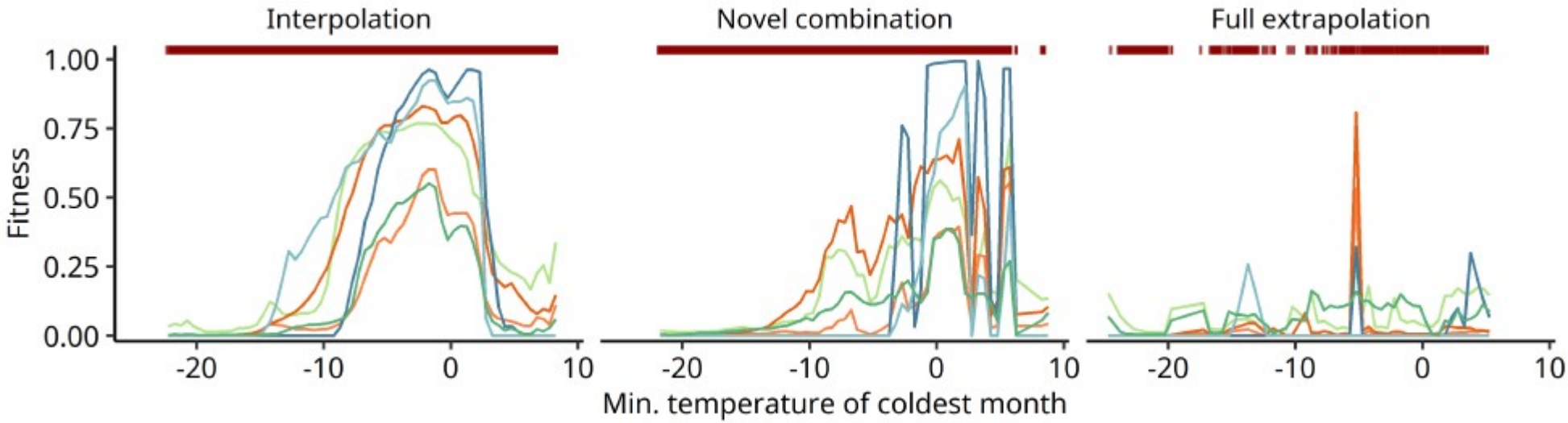
CLIMATE APPROACH: Categorical differenciation



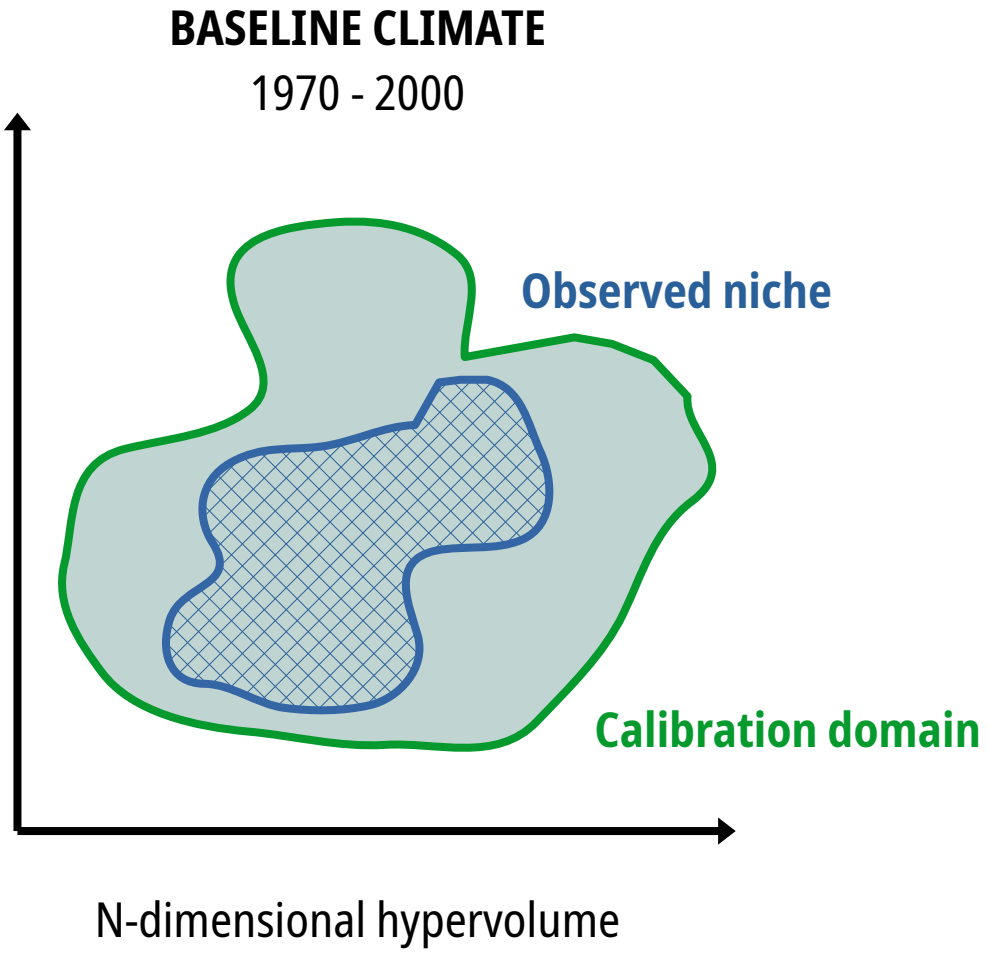
CLIMATE APPROACH: Categorical differenciation



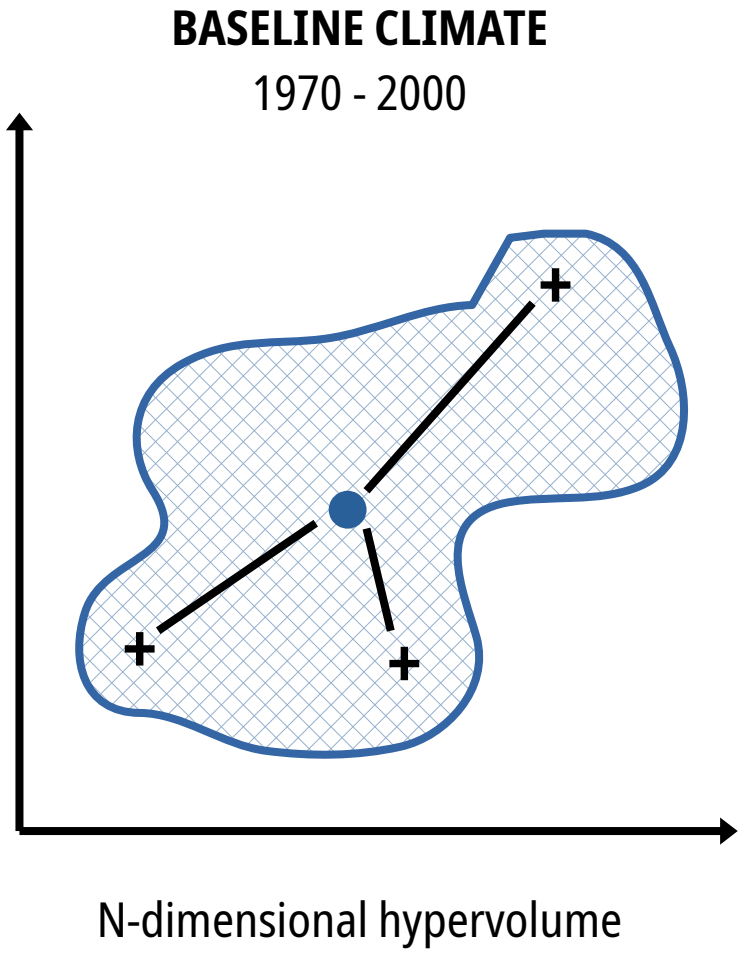
CLIMATE APPROACH: Categorical differenciation and model response



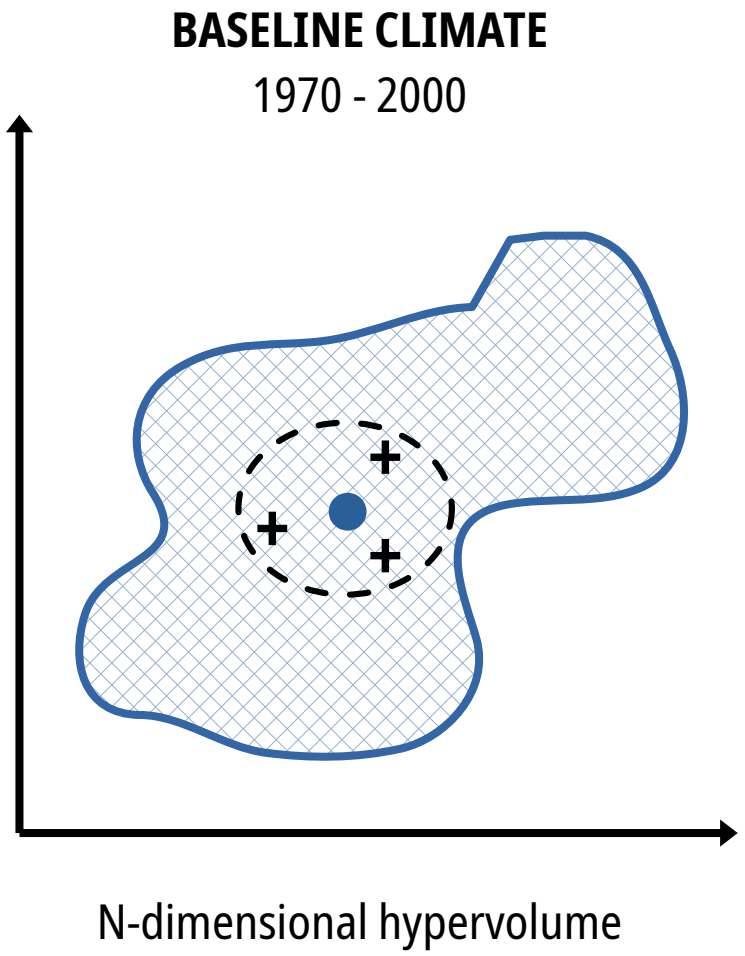
NICHE APPROACH: average distance to centroid



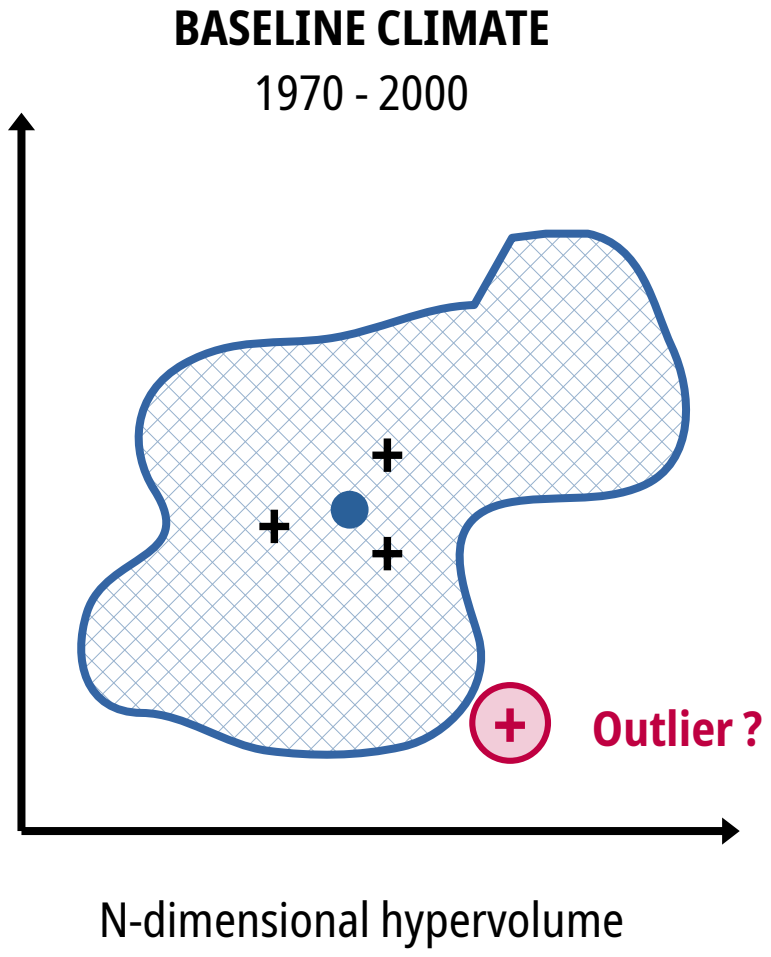
NICHE APPROACH: average distance to centroid



NICHE APPROACH: average distance to centroid



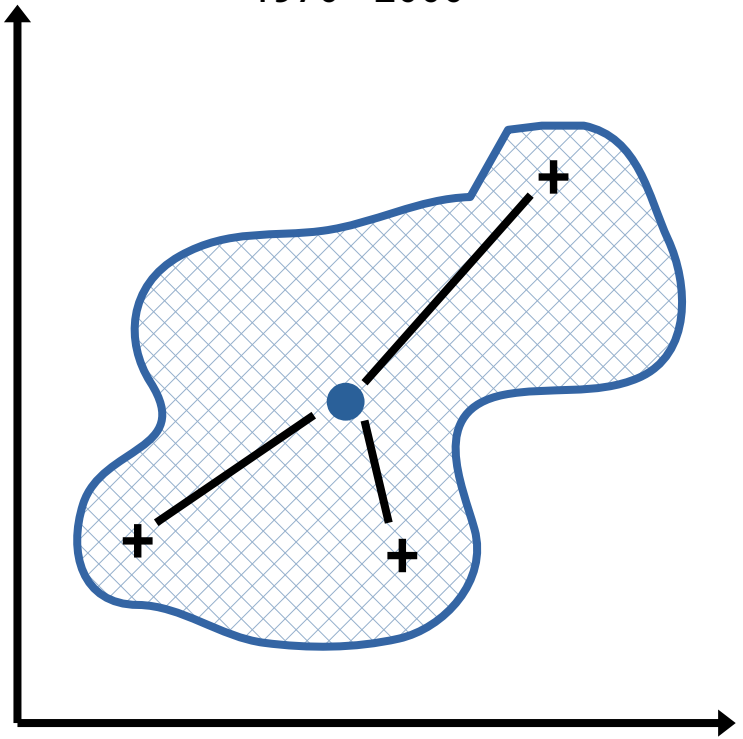
NICHE APPROACH: average distance to centroid



NICHE APPROACH: average distance to centroid

BASELINE CLIMATE

1970 - 2000



Average distance between pollen records and centroid

N-dimensional hypervolume