# Fiabilité des projections des modèles d'aire de répartition des arbres forestiers d'Europe

2021-2024

**Objectif :** comparer les capacités des modèles corrélatifs et des modèles basés sur les processus à prédire les répartitions passée et future d'espèces d'arbres

**Objectif :** comparer les capacités des modèles corrélatifs et des modèles basés sur les processus à prédire les répartitions passée et future d'espèces d'arbres

**Quelle est l'origine de leur robustesse ?**

- les hypothèses du modèle (relations mécanistes vs. statistiques) ?
- la façon dont les paramètres sont estimés ?

**Objectif :** comparer les capacités des modèles corrélatifs et des modèles basés sur les processus à prédire les répartitions passée et future d'espèces d'arbres

**Quelle est l'origine de leur robustesse ?**

- les hypothèses du modèle (relations mécanistes vs. statistiques) ?
- la façon dont les paramètres sont estimés ?

**Modèles :** PHENOFIT, CASTANEA, FORCEPS ? + plusieurs modèles corrélatifs
Versions plus ou moins complètes, et paramétrées différemment

**Objectif :** comparer les capacités des modèles corrélatifs et des modèles basés sur les processus à prédire les répartitions passée et future d'espèces d'arbres

**Quelle est l'origine de leur robustesse ?**

- ► les hypothèses du modèle (relations mécanistes vs. statistiques) ?
- ► la façon dont les paramètres sont estimés ?

**Modèles :** PHENOFIT, CASTANEA, FORCEPS ? + plusieurs modèles corrélatifs
Versions plus ou moins complètes, et paramétrées différemment

**Périodes :** - passé lointain (Holocène)
- passé proche (1850-2015)
- présent (calibration)
- futur proche (2025-2100)

**Efficiency/accuracy:** ability of the model to provide accurate predictions in conditions that have been used to calibrate the model

**Robustness:** ability of the model to provide accurate projections in external conditions, i.e. other conditions than those used to calibrate the model
$\Longrightarrow$ determines its **transferability in time** and space

(*Janssen and Heuberger, 1995 ; Asse et al., 2020*)

## Correlative vs process-based : non-exhaustive state of the art

### ► Présent

- *Shabani et al. (2016)* :
  Climex vs. 5 CSDMs (calibration : outside Australia, evaluation : Australia)
- *Higgins et al. (2020)* :
  TTR-SDM vs. Maxent (calibration : Australia, evaluation : outside Australia)
- *Gritti et al. (2013)* :
  STASH, LPJ, et PHENOFIT
- *Magarey et al. (2018) :*
  TTR-SDM vs. Maxent (calibration : outside US, evaluation : US)

### ► Futur

- *Cheaib et al. (2012) :*
  PHENOFIT, CASTANEA et 2 CSDMs…
- *Morin and Thuiller (2009) :*
  PHENOFIT vs. Biomod
- *Kearney et al. (2010) :*
  NicheMapper vs. Maxent
- *Keenan et al. (2011) :*
  GOTILWA+ vs. Biomod
- *Schneiderman et al. (2015) :*
  LINKAGES 2.2 vs. TreeAtlas model
- Kramer et al. (2010) :
  LPJ-GUESS vs. Biomod

*And so on…*

Asse et al. (2020). Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate. *Agricultural and Forest Meteorology*

Asse et al. (2020). Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate. *Agricultural and Forest Meteorology*

Modèle phénologique corrélatif *vs.* modèles basés sur les processus paramétrisés différemment :

- entièrement à partir des données **(backward)**
- en fixant certains paramètres grâce à des données expérimentales **(**partly **forward)**

Asse et al. (2020). Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate. *Agricultural and Forest Meteorology*

Modèle phénologique corrélatif *vs.* modèles basés sur les processus paramétrisés différemment :

- ► entièrement à partir des données **(backward)**
- ► en fixant certains paramètres grâce à des données expérimentales **(**partly **forward)**

« *Our results show that robustness of process-based models can come additionally from forward parameter estimation, but also that forward parameter estimation is not necessarily the Holy Grail that we should seek.* »

Asse et al. (2020). Process-based models outcompete correlative models in projecting spring phenology of trees in a future warmer climate. *Agricultural and Forest Meteorology*

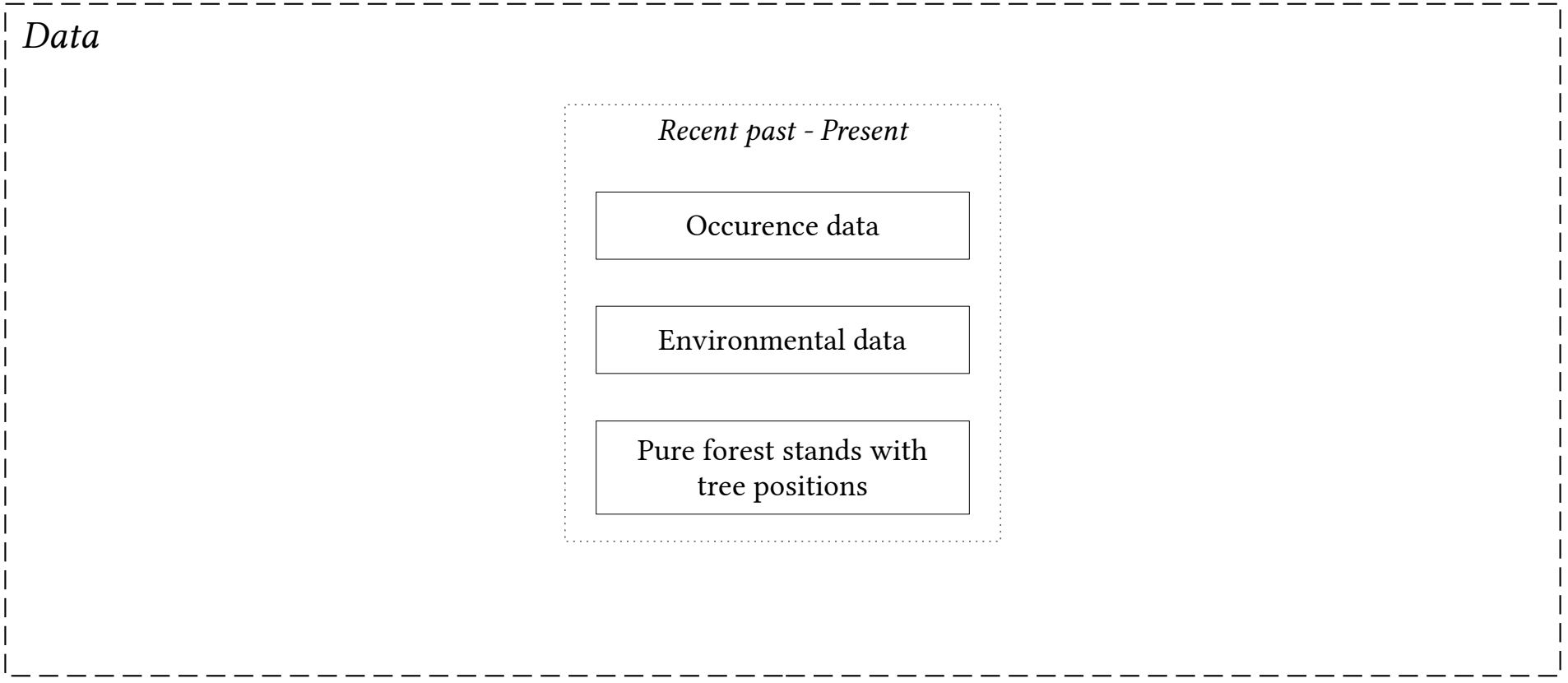
Modèle phénologique corrélatif *vs.* modèles basés sur les processus paramétrisés différemment :

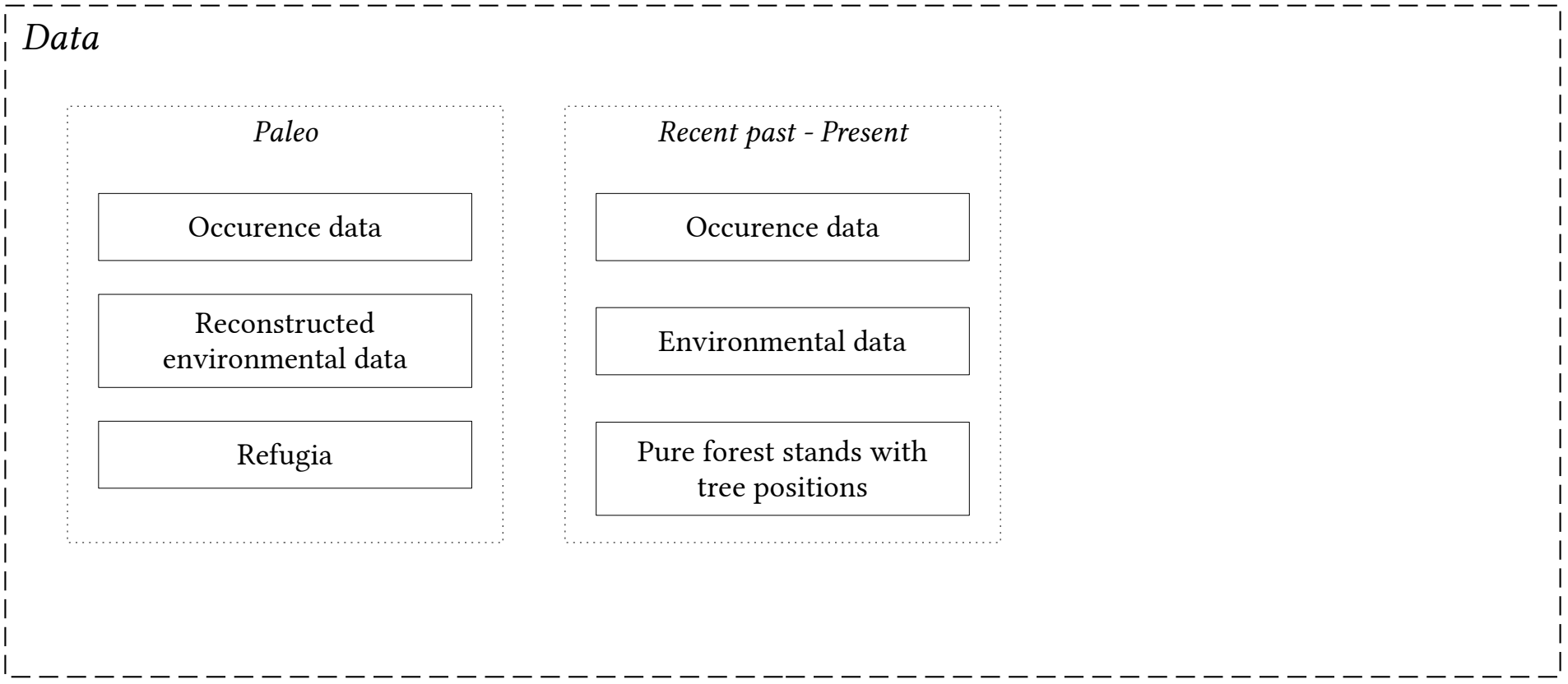
- ► entièrement à partir des données **(backward)**
- ► en fixant certains paramètres grâce à des données expérimentales **(**partly **forward)**

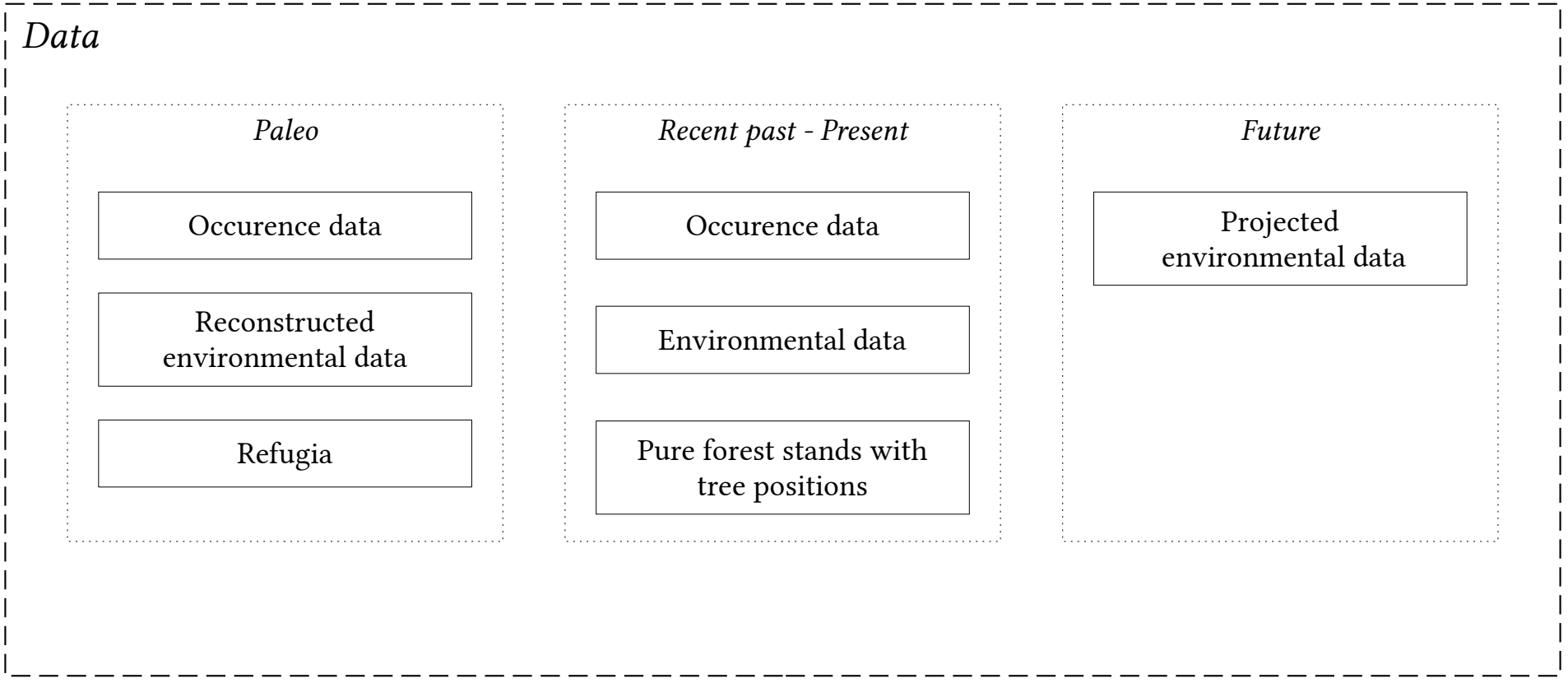
« *Our results show that robustness of process-based models can come additionally from forward parameter estimation, but also that forward parameter estimation is not necessarily the Holy Grail that we should seek.* »

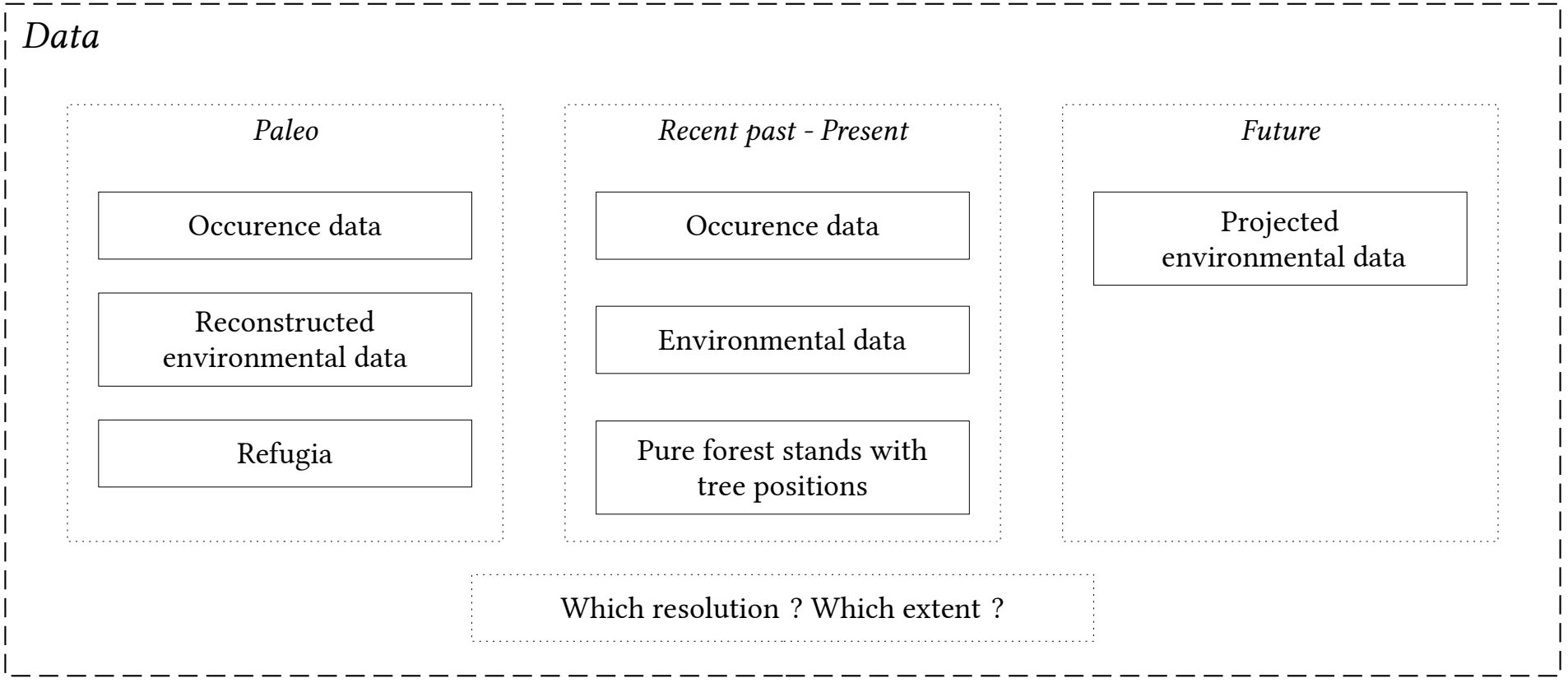
**⟹ Et pour les modèles d'aire de répartition ?**

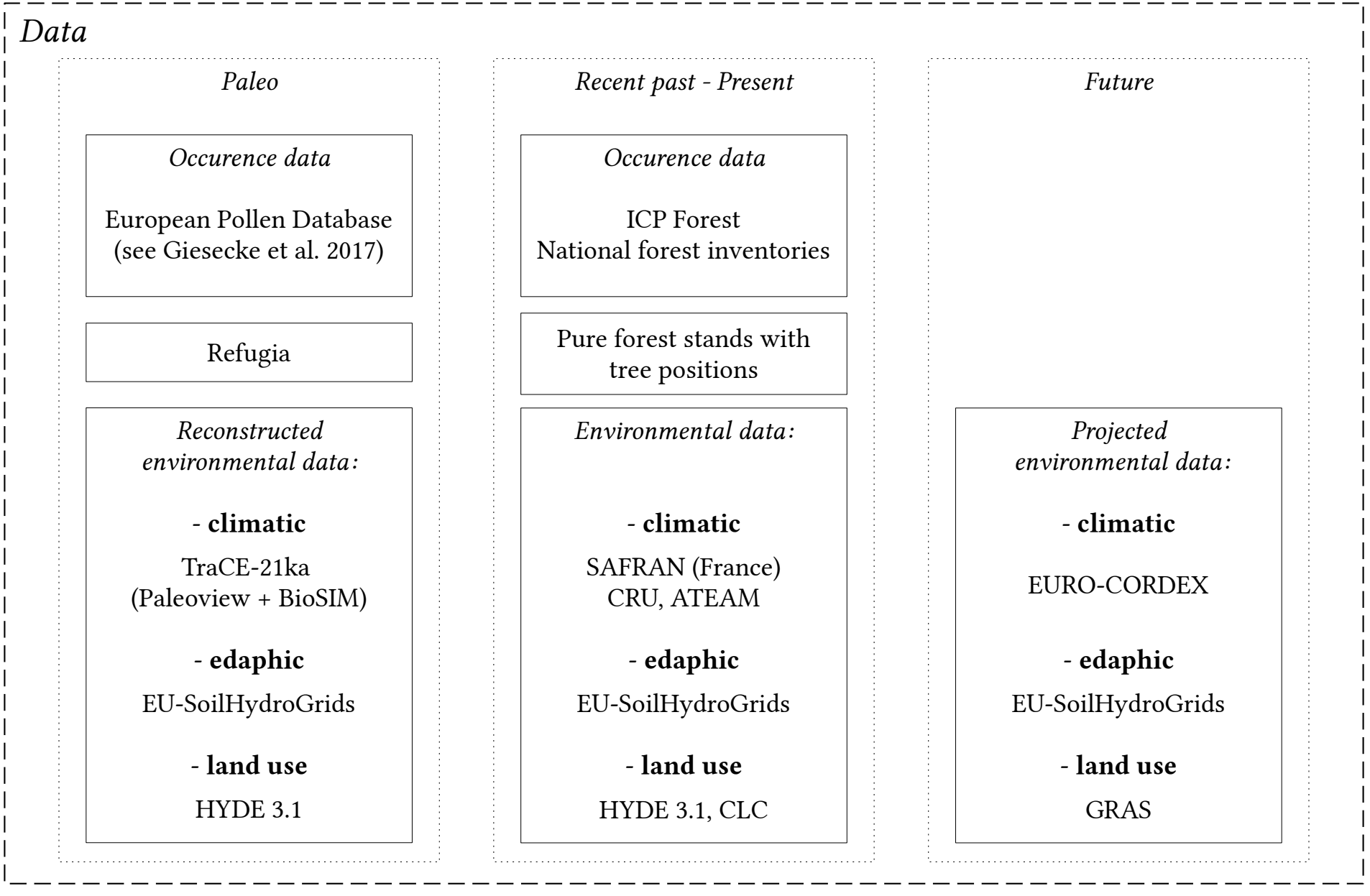

Data


Data
Recent past - Present
Occurence data
Environmental data
Pure forest stands with tree positions

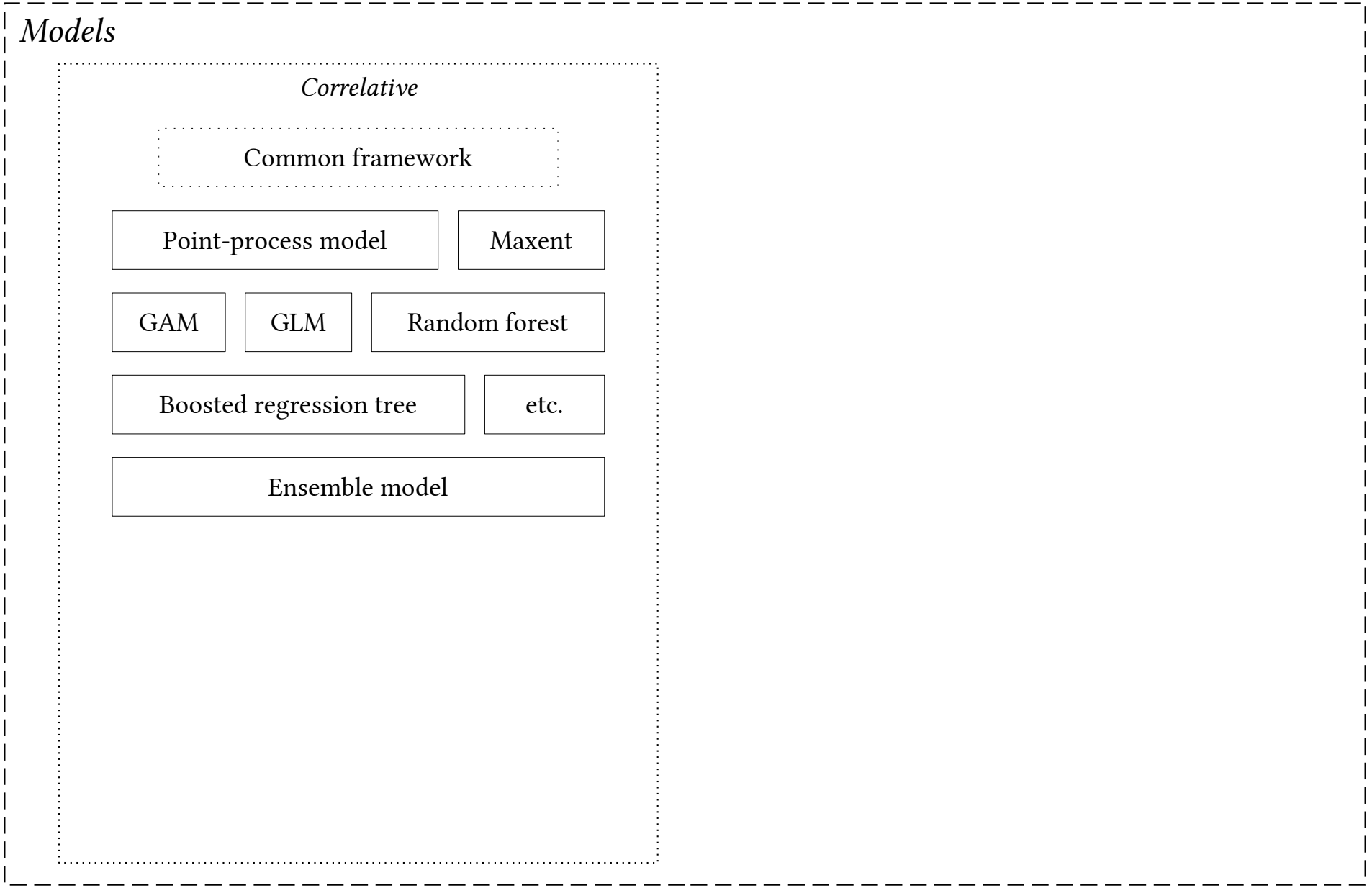

Data
Paleo
Occurence data
Reconstructed environmental data
Refugia
Recent past - Present
Occurence data
Environmental data
Pure forest stands with tree positions

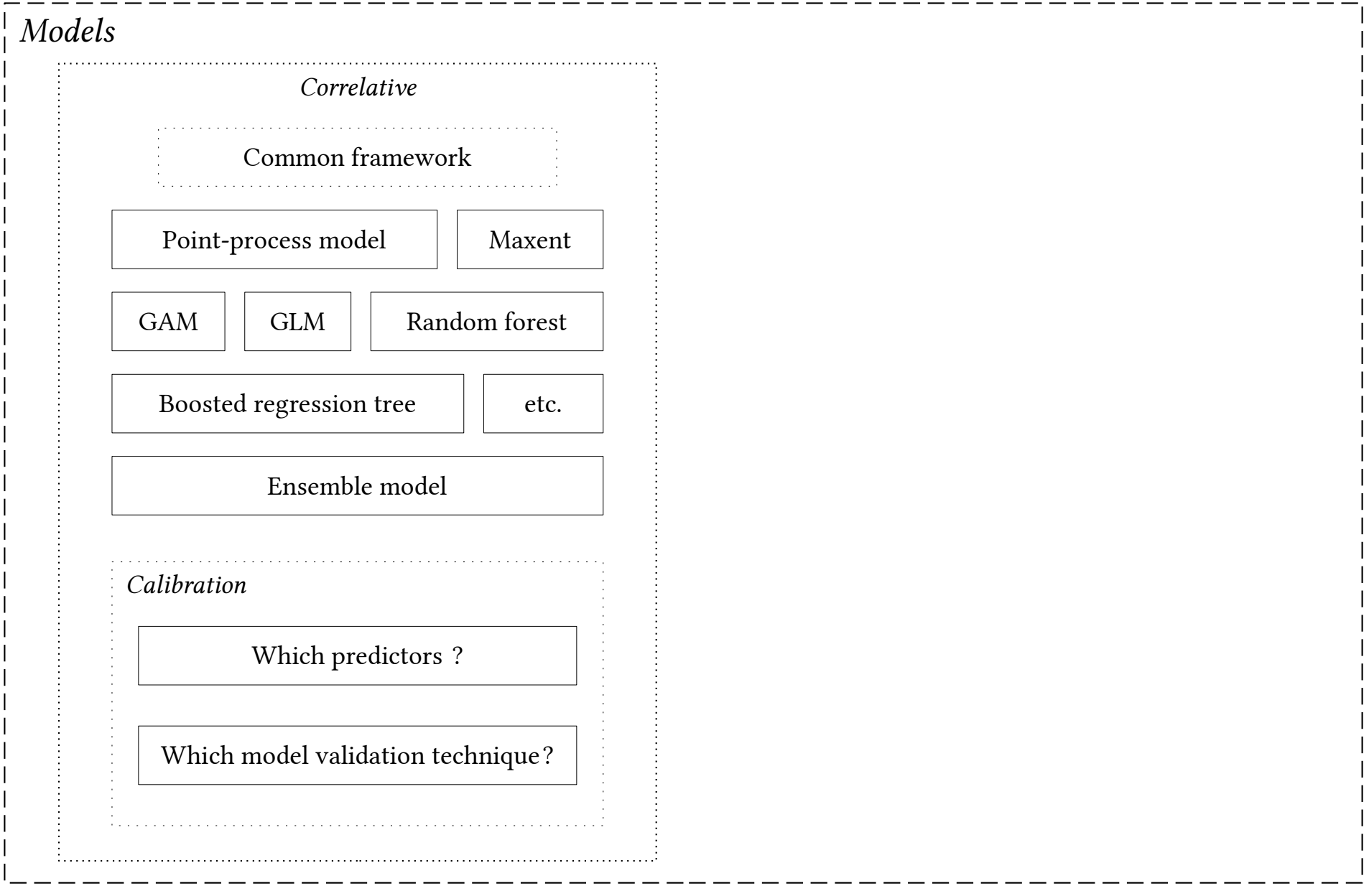

Data
Paleo
Occurence data
Reconstructed environmental data
Refugia
Recent past - Present
Occurence data
Environmental data
Pure forest stands with tree positions
Future
Projected environmental data

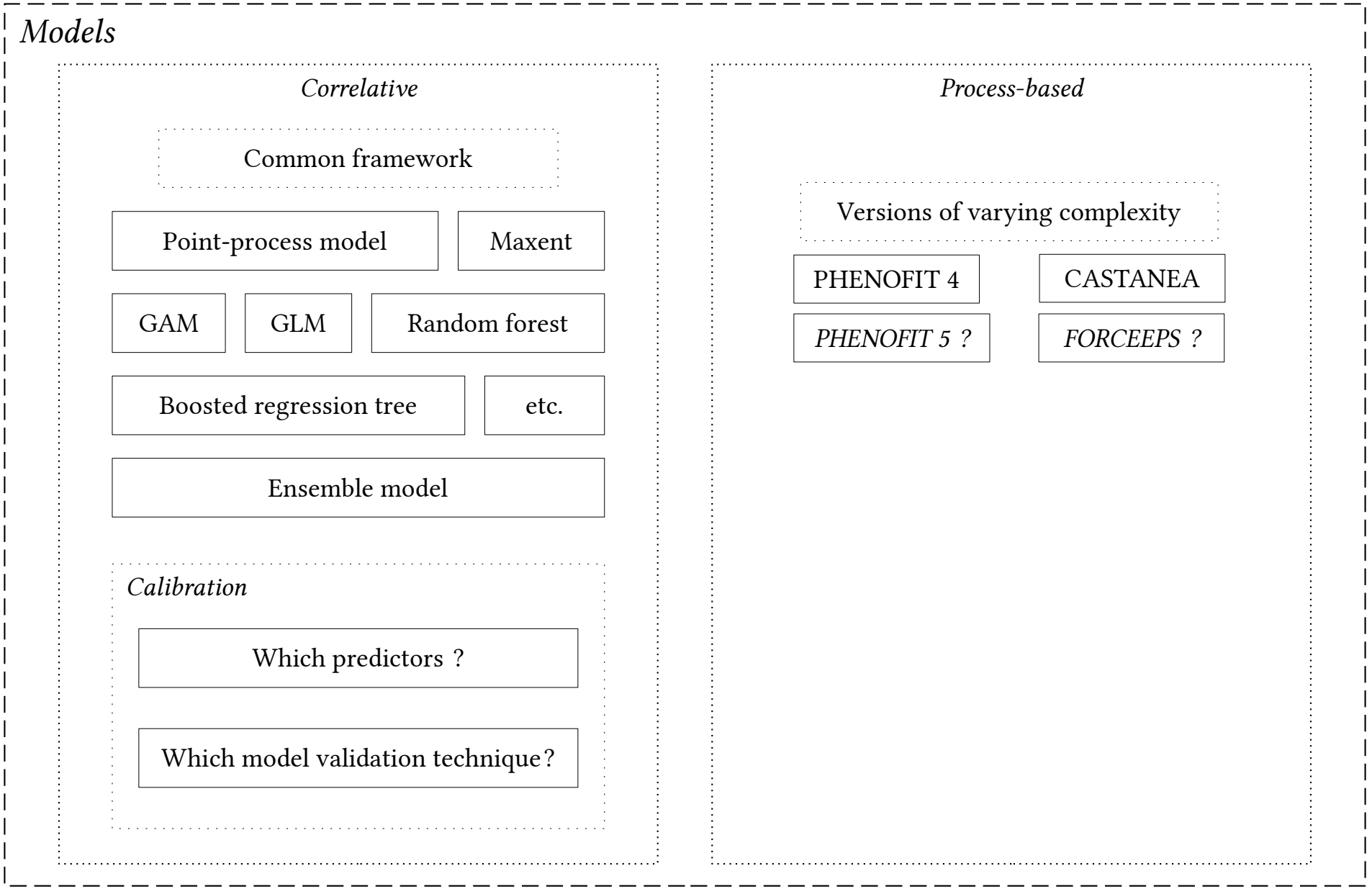

Data
Paleo
Occurence data
Reconstructed environmental data
Refugia
Recent past - Present
Occurence data
Environmental data
Pure forest stands with tree positions
Future
Projected environmental data
Which resolution ? Which extent ?

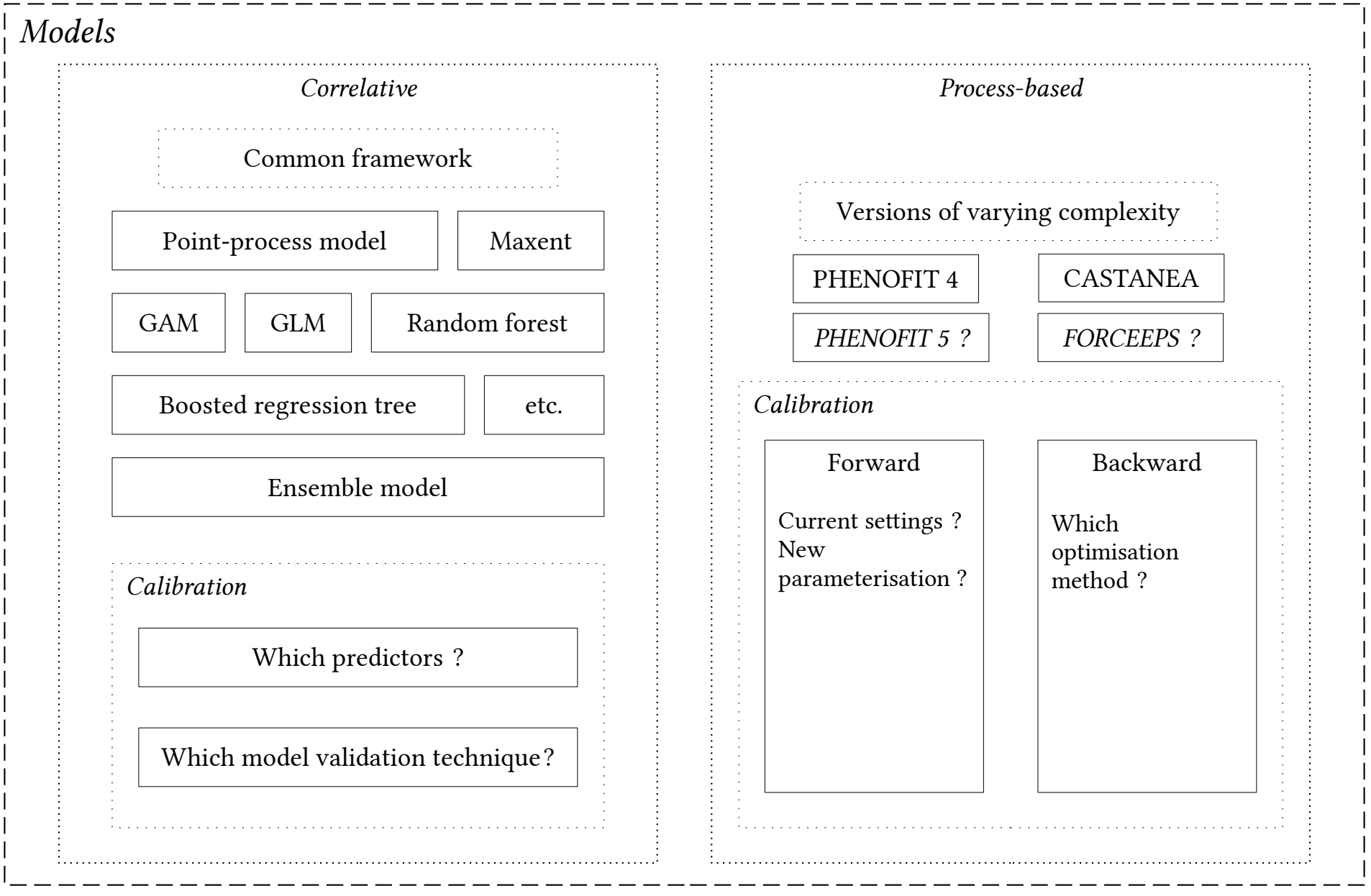

Data
Paleo
Occurence data
European Pollen Database
(see Giesecke et al. 2017)
Refugia
Reconstructed environmental data:
- climatic
TraCE-21ka
(Paleoview + BioSIM)
- edaphic
EU-SoilHydroGrids
- land use
HYDE 3.1
Recent past - Present
Occurence data
ICP Forest
National forest inventories
Pure forest stands with tree positions
Environmental data:
- climatic
SAFRAN (France)
CRU, ATEAM
- edaphic
EU-SoilHydroGrids
- land use
HYDE 3.1, CLC
Future
Projected environmental data:
- climatic
EURO-CORDEX
- edaphic
EU-SoilHydroGrids
- land use
GRAS

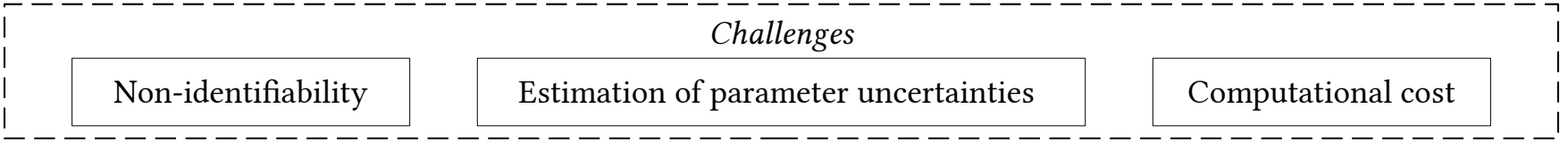
*Models*

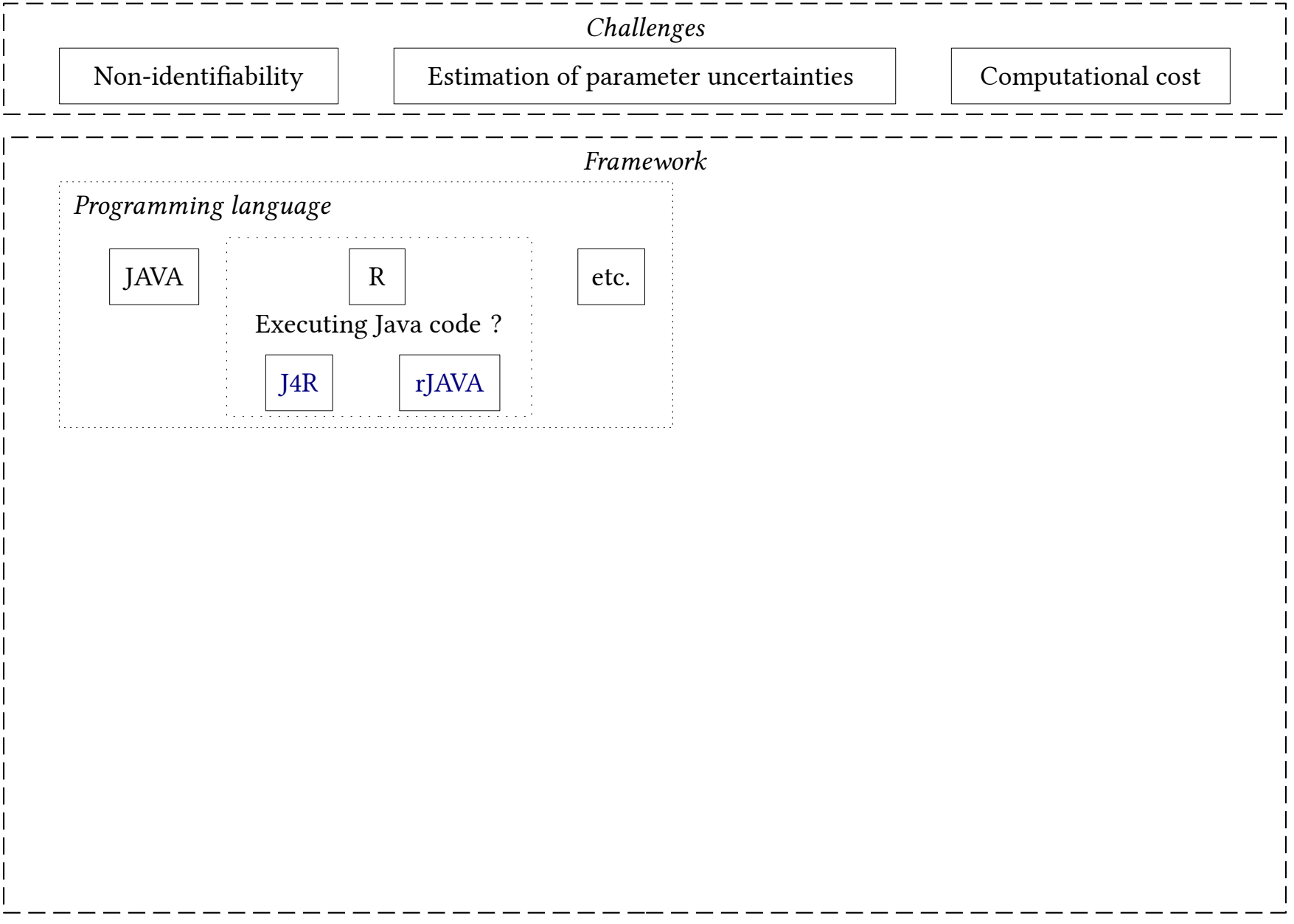

Models
Correlative
Common framework
Point-process model
Maxent
GAM
GLM
Random forest
Boosted regression tree
etc.
Ensemble model

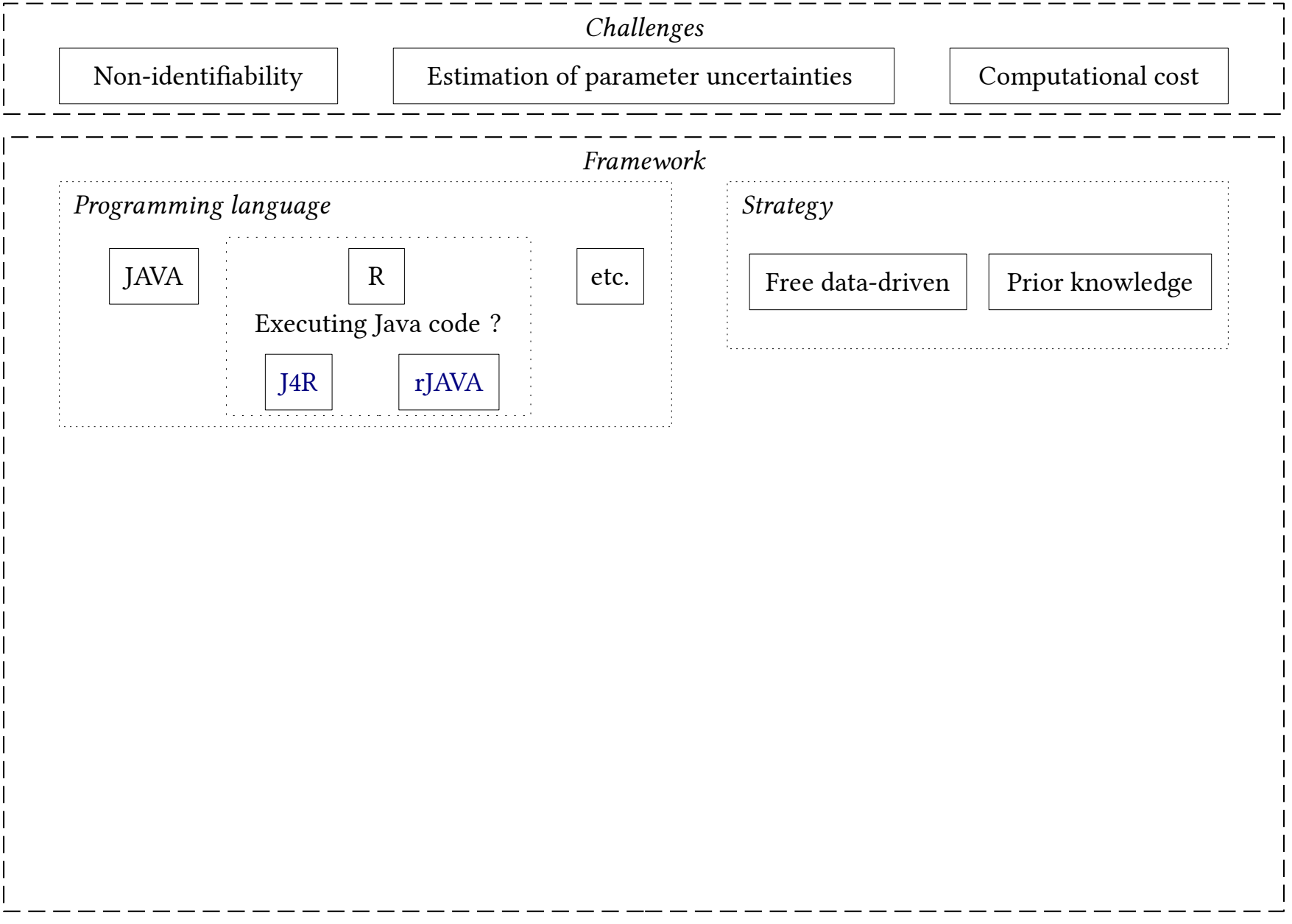

Models
Correlative
Common framework
Point-process model
Maxent
GAM
GLM
Random forest
Boosted regression tree
etc.
Ensemble model
Calibration
Which predictors ?
Which model validation technique?

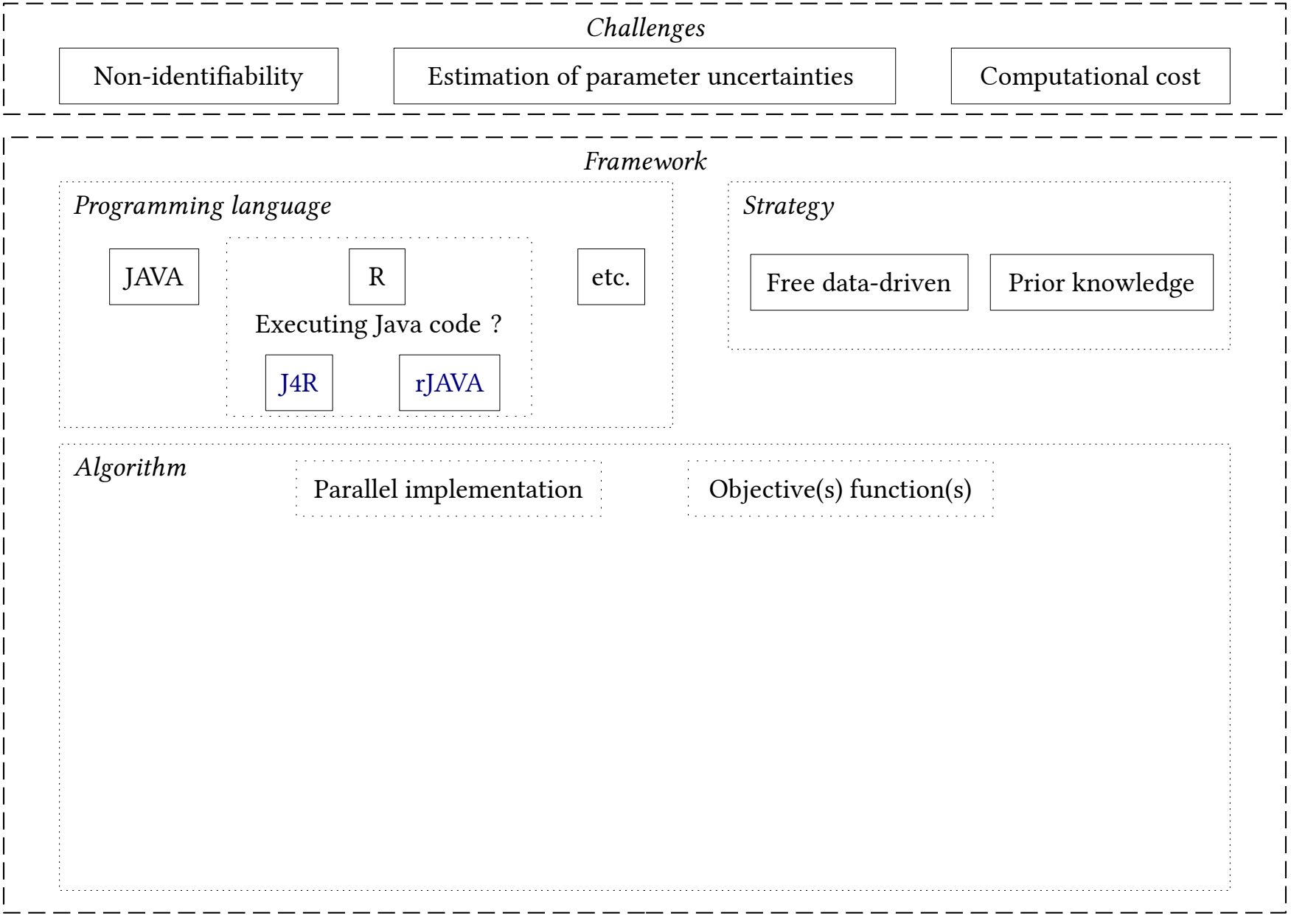

Models
Correlative
Common framework
Point-process model
Maxent
GAM
GLM
Random forest
Boosted regression tree
etc.
Ensemble model
Calibration
Which predictors ?
Which model validation technique?
Process-based
Versions of varying complexity
PHENOFIT 4
CASTANEA
PHENOFIT 5 ?
FORCEEPS ?

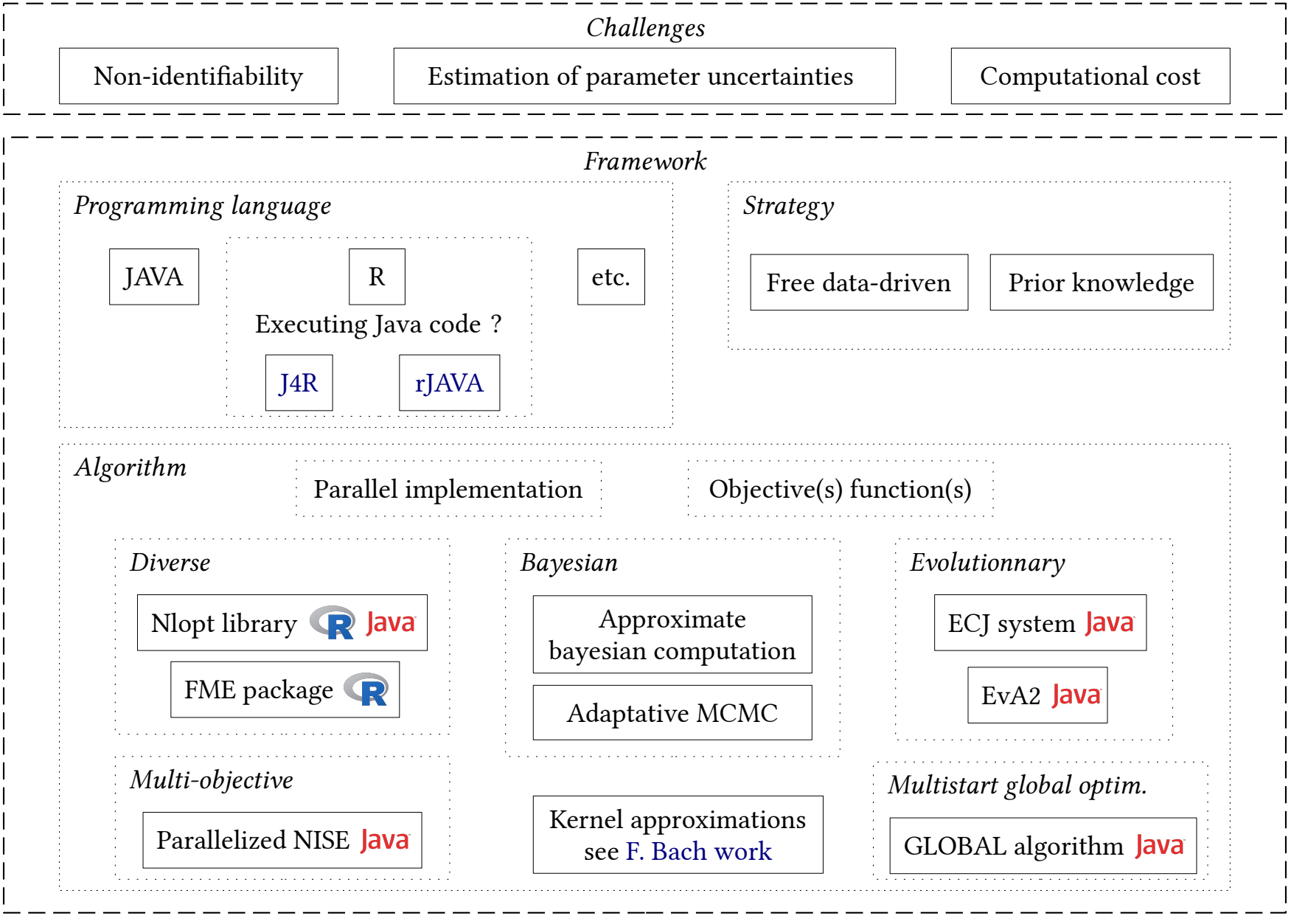

Models
Correlative
Common framework
Point-process model
Maxent
GAM
GLM
Random forest
Boosted regression tree
etc.
Ensemble model
Calibration
Which predictors ?
Which model validation technique?
Process-based
Versions of varying complexity
PHENOFIT 4
CASTANEA
PHENOFIT 5 ?
FORCEEPS ?
Calibration
Forward
Current settings ?
New
parameterisation ?
Backward
Which
optimisation
method ?

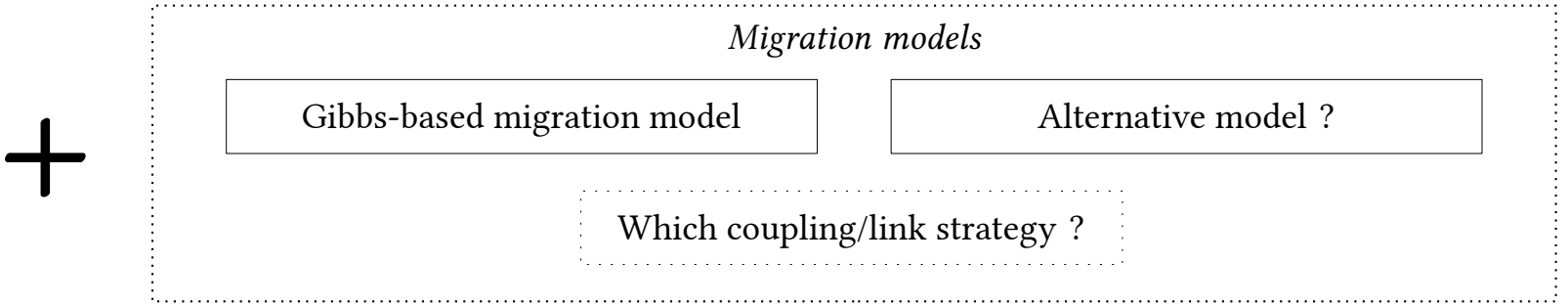

Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost


Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost
Framework
Programming language
JAVA
R
etc.
Executing Java code ?
J4R
rJAVA

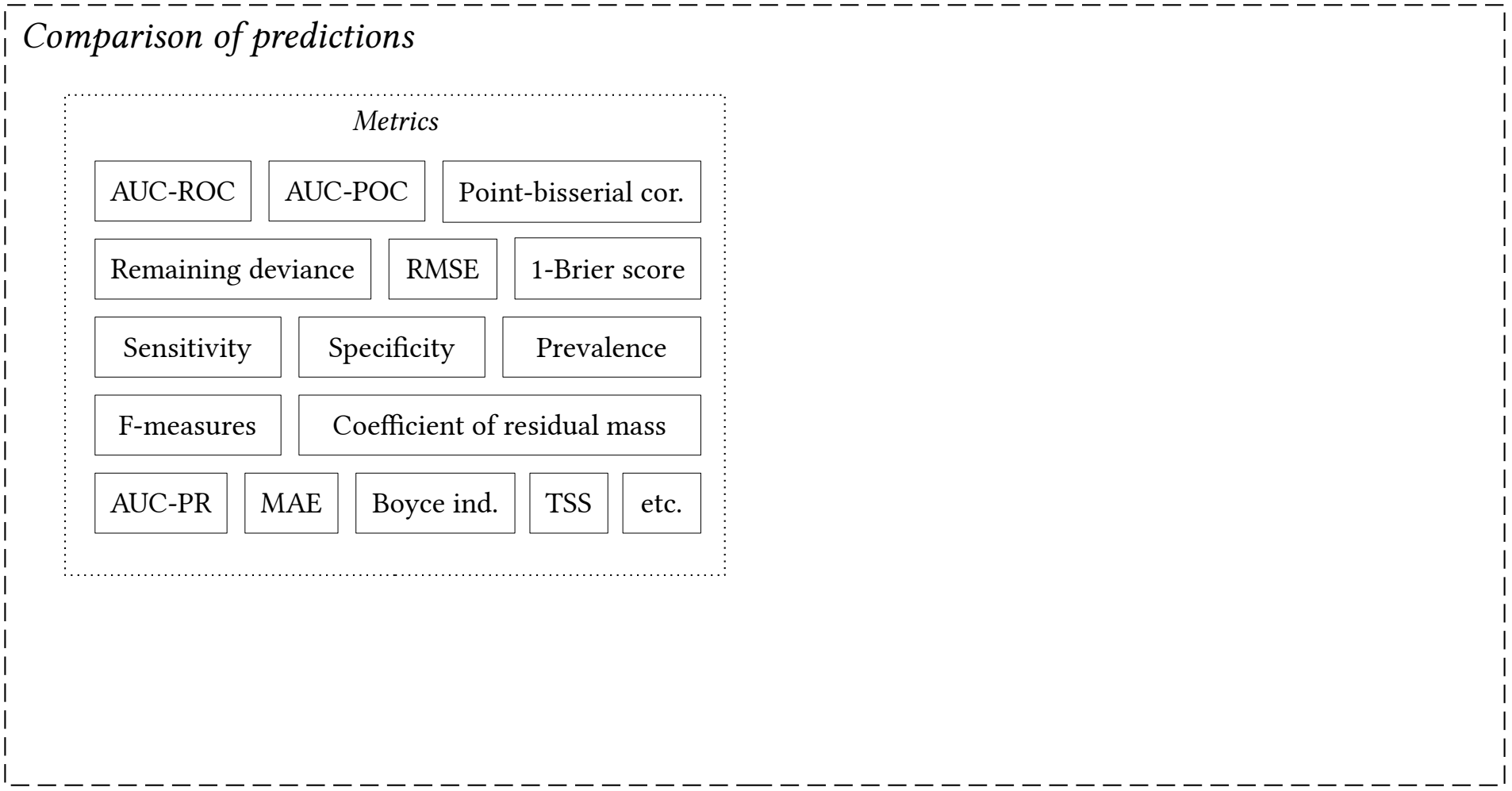

Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost
Framework
Programming language
JAVA
R
Executing Java code ?
J4R
rJAVA
etc.
Strategy
Free data-driven
Prior knowledge

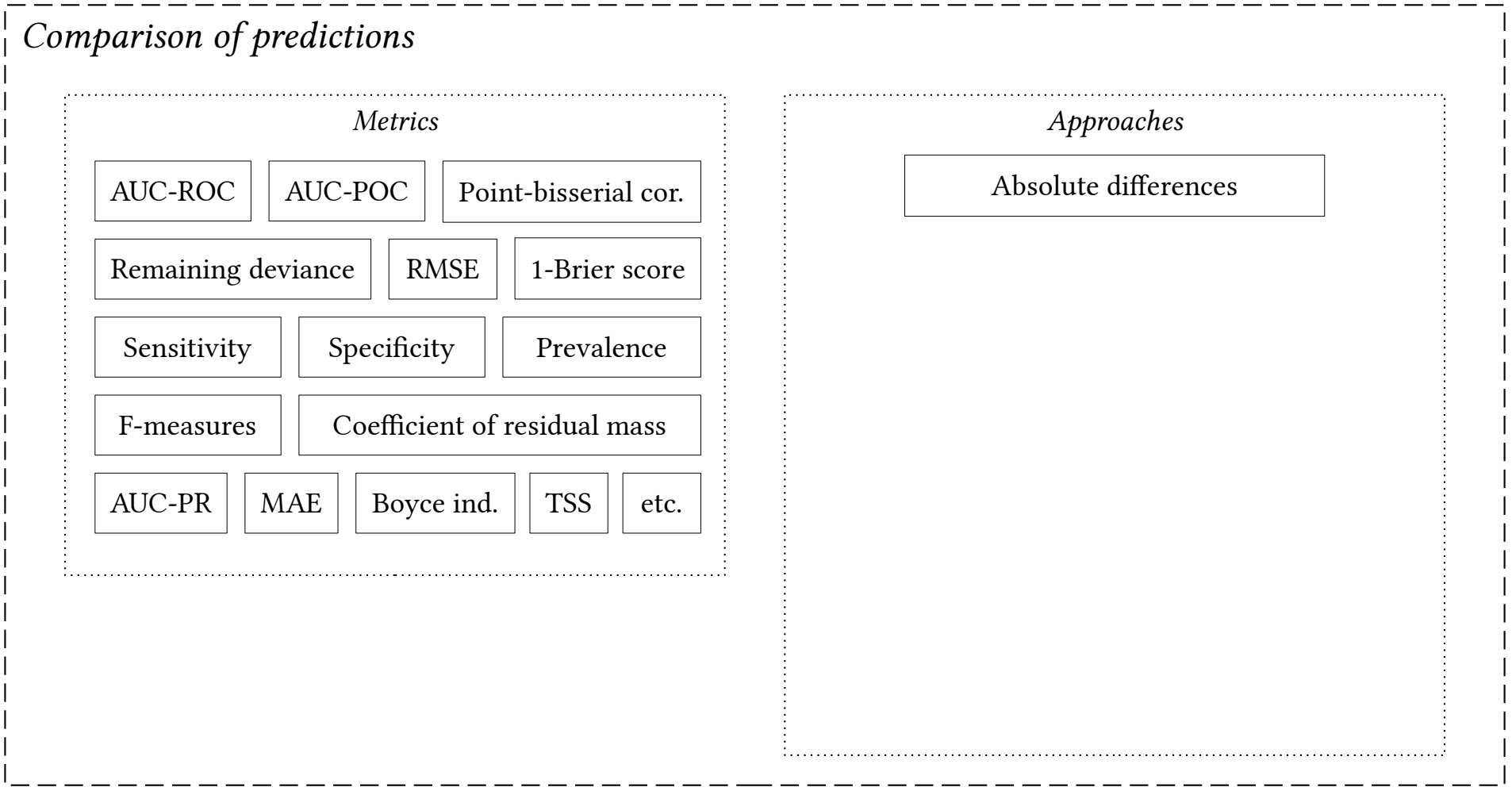

Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost
Framework
Programming language
JAVA
R
Executing Java code ?
J4R
rJAVA
etc.
Strategy
Free data-driven
Prior knowledge
Algorithm
Parallel implementation
Objective(s) function(s)

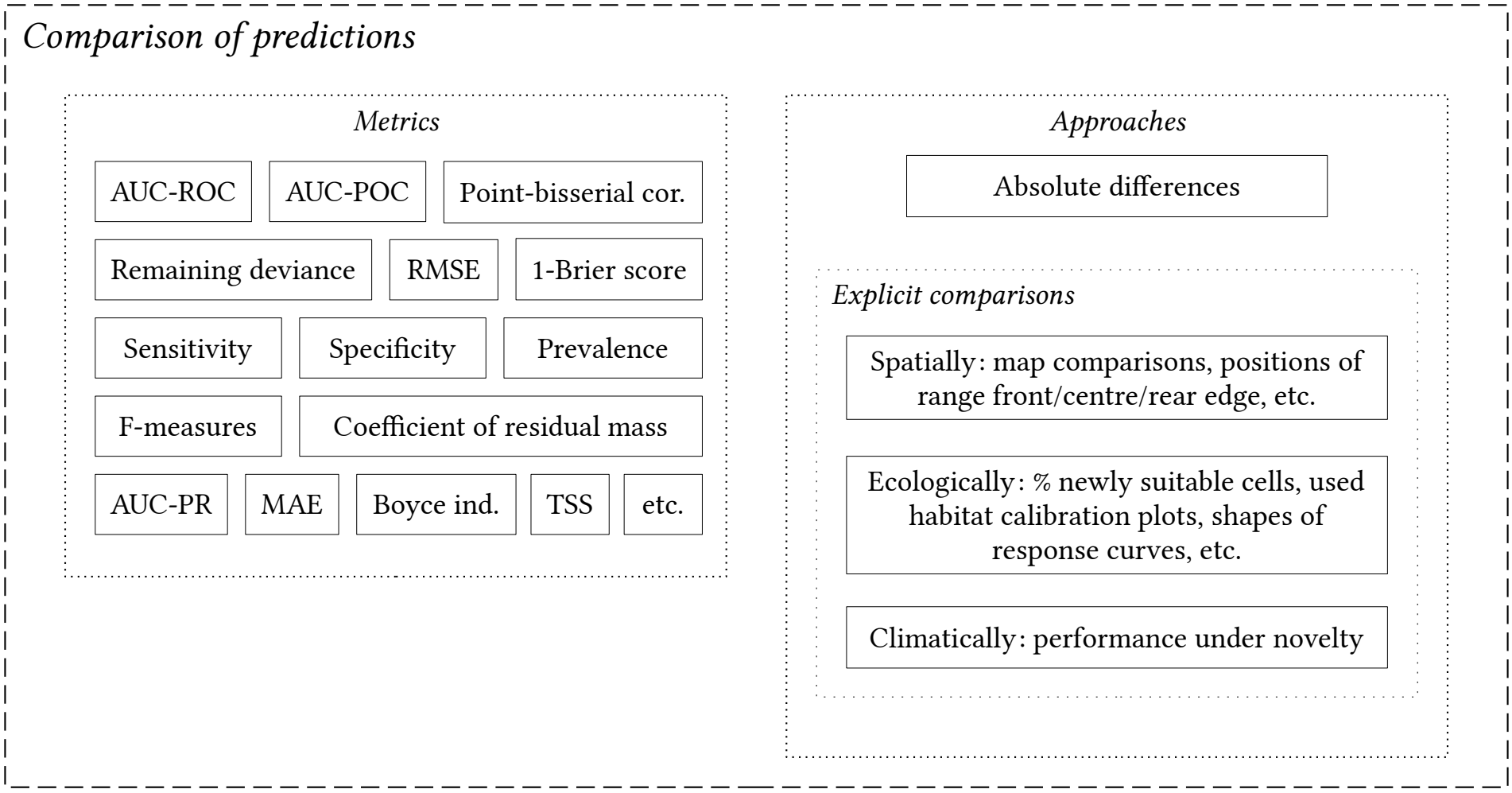

Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost
Framework
Programming language
JAVA
R
Executing Java code ?
J4R
rJAVA
etc.
Strategy
Free data-driven
Prior knowledge
Algorithm
Parallel implementation
Objective(s) function(s)
Diverse
Nlopt library
Java
FME package
Bayesian
Approximate bayesian computation
Adaptative MCMC
Evolutionnary
ECJ system
Java
EvA2
Java
Multi-objective
Parallelized NISE
Java
Kernel approximations see F. Bach work
Multistart global optim.
GLOBAL algorithm
Java

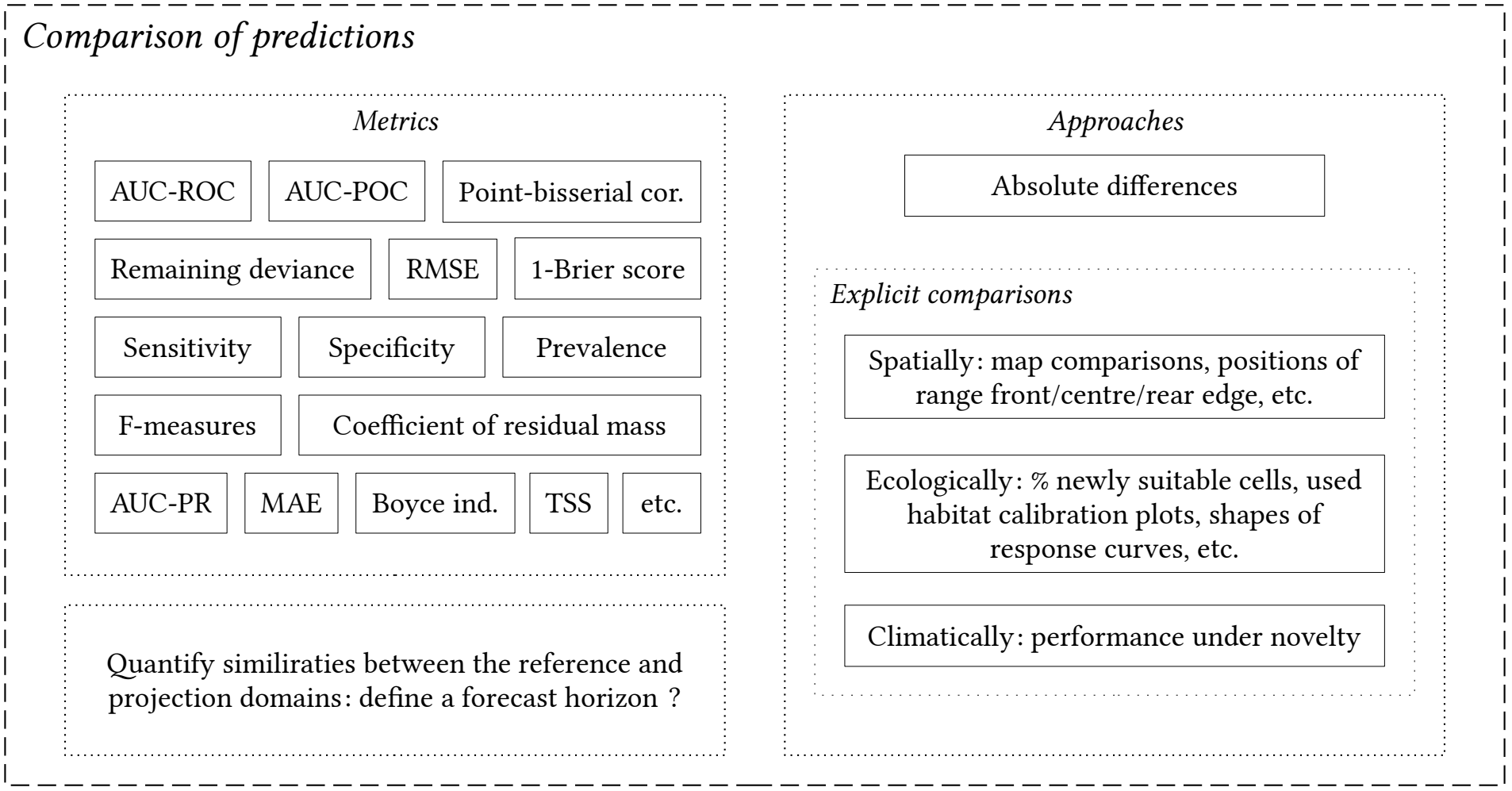

Migration models
Gibbs-based migration model
Alternative model ?
Which coupling/link strategy ?

*Comparison of predictions*


Comparison of predictions
Metrics
AUC-ROC
AUC-POC
Point-bisserial cor.
Remaining deviance
RMSE
1-Brier score
Sensitivity
Specificity
Prevalence
F-measures
Coefficient of residual mass
AUC-PR
MAE
Boyce ind.
TSS
etc.

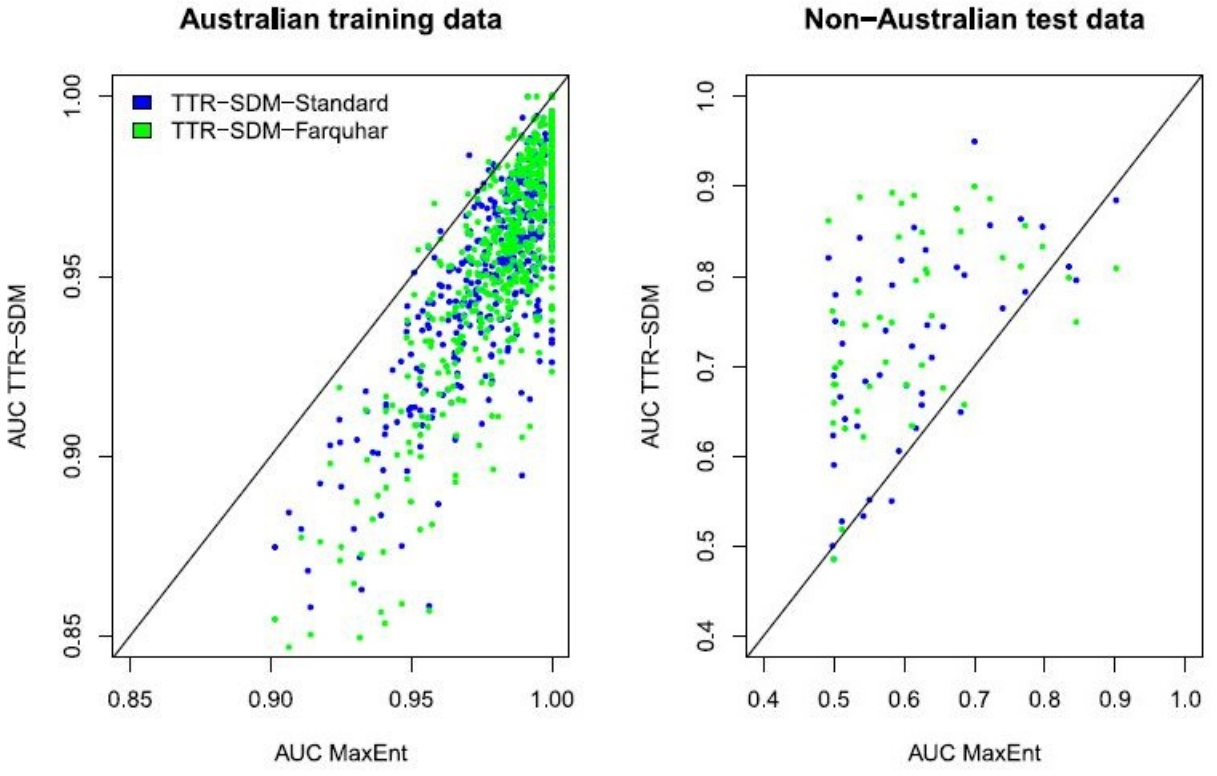

Comparison of predictions
Metrics
AUC-ROC
AUC-POC
Point-bisserial cor.
Remaining deviance
RMSE
1-Brier score
Sensitivity
Specificity
Prevalence
F-measures
Coefficient of residual mass
AUC-PR
MAE
Boyce ind.
TSS
etc.
Approaches
Absolute differences


Comparison of predictions
Metrics
AUC-ROC
AUC-POC
Point-bisserial cor.
Remaining deviance
RMSE
1-Brier score
Sensitivity
Specificity
Prevalence
F-measures
Coefficient of residual mass
AUC-PR
MAE
Boyce ind.
TSS
etc.
Approaches
Absolute differences
Explicit comparisons
Spatially: map comparisons, positions of range front/centre/rear edge, etc.
Ecologically: % newly suitable cells, used habitat calibration plots, shapes of response curves, etc.
Climatically: performance under novelty


Comparison of predictions
Metrics
AUC-ROC
AUC-POC
Point-bisserial cor.
Remaining deviance
RMSE
1-Brier score
Sensitivity
Specificity
Prevalence
F-measures
Coefficient of residual mass
AUC-PR
MAE
Boyce ind.
TSS
etc.
Quantify similiraties between the reference and projection domains: define a forecast horizon ?
Approaches
Absolute differences
Explicit comparisons
Spatially: map comparisons, positions of range front/centre/rear edge, etc.
Ecologically: % newly suitable cells, used habitat calibration plots, shapes of response curves, etc.
Climatically: performance under novelty

# Absolute differences

*Higgins et al. (2020)*

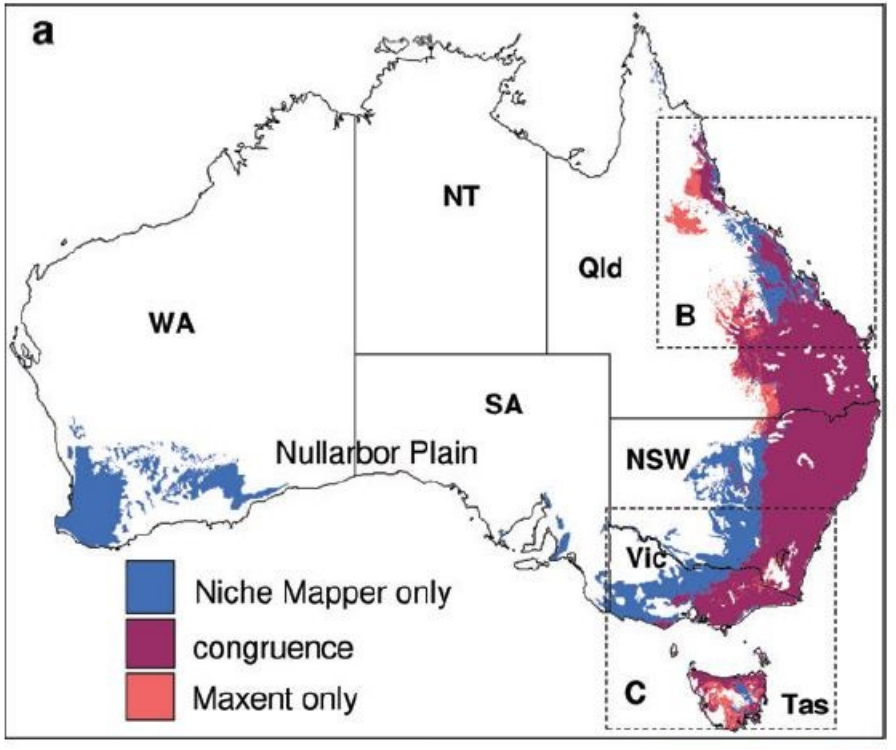



interpolation

extrapolation

## Spatially explicit comparisons

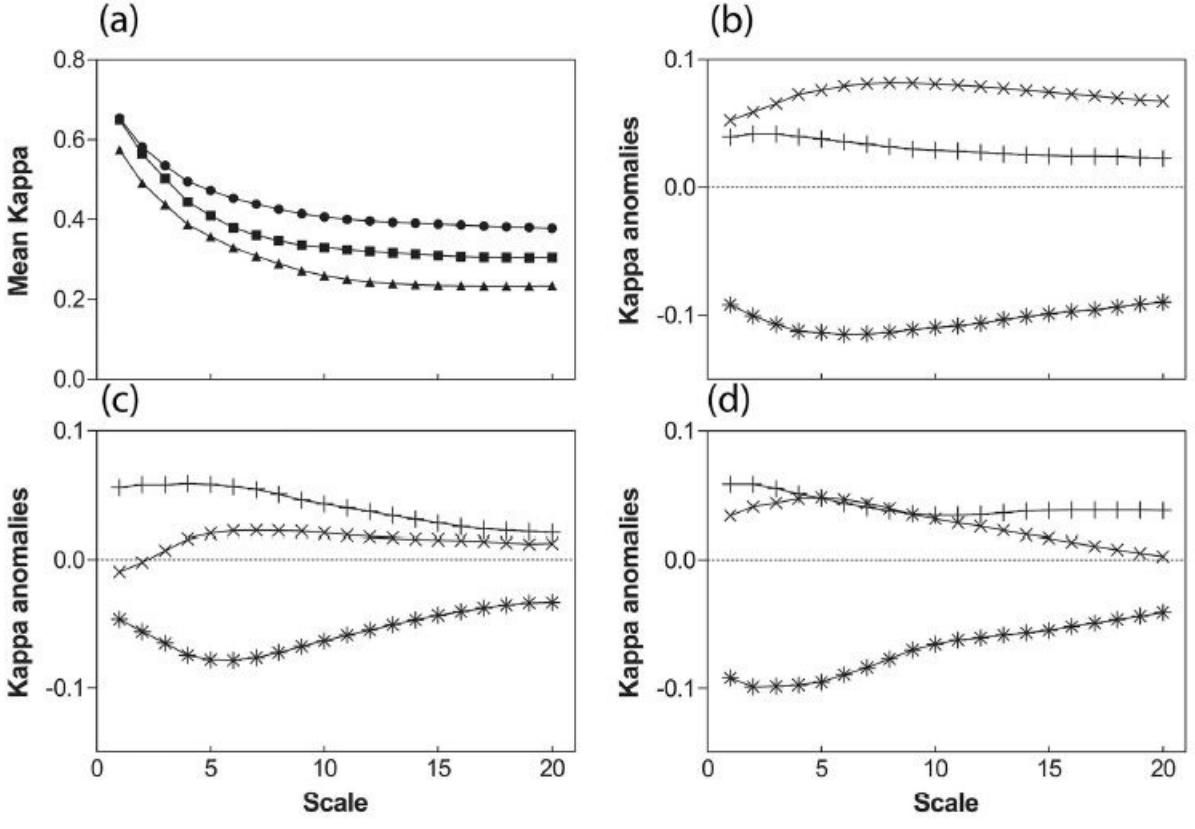
*Kearney et al. (2010)*




Congruence between model predictions

# Spatially explicit comparisons

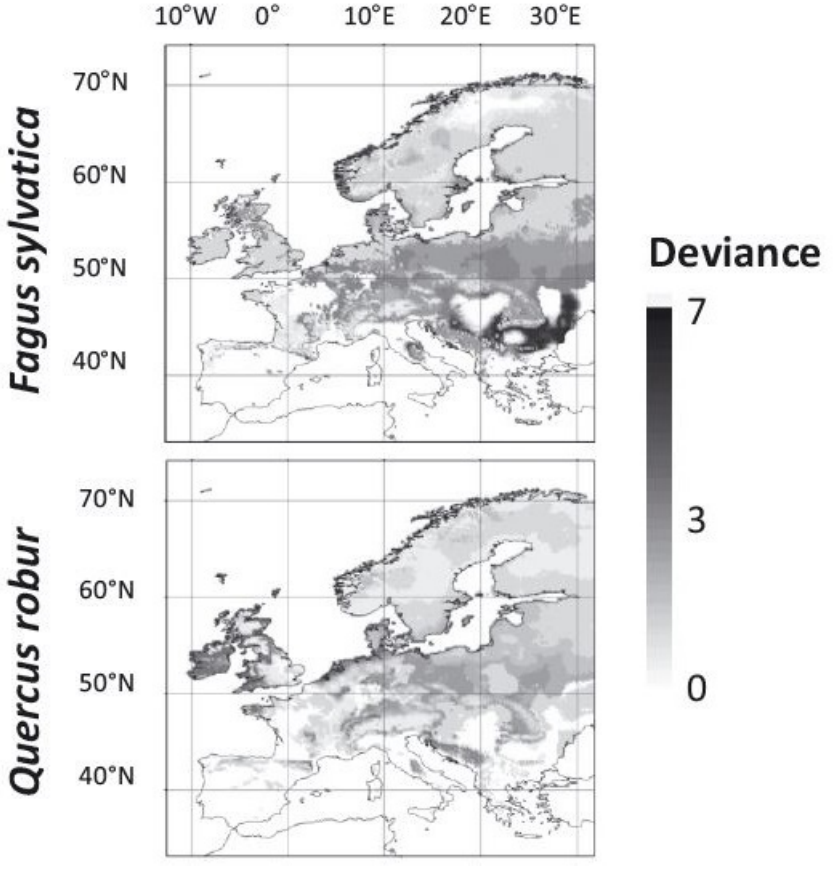
*Gritti et al. (2013a)*




Scale-dependent comparisons (CMP method)

## Spatially explicit comparisons

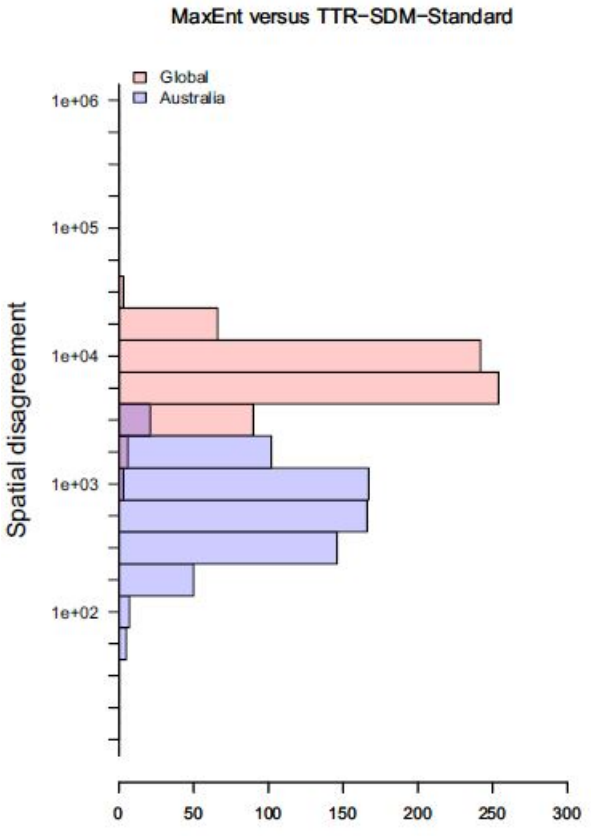
*Gritti et al. (2013b)*




Deviance between models

# Spatially explicit comparisons

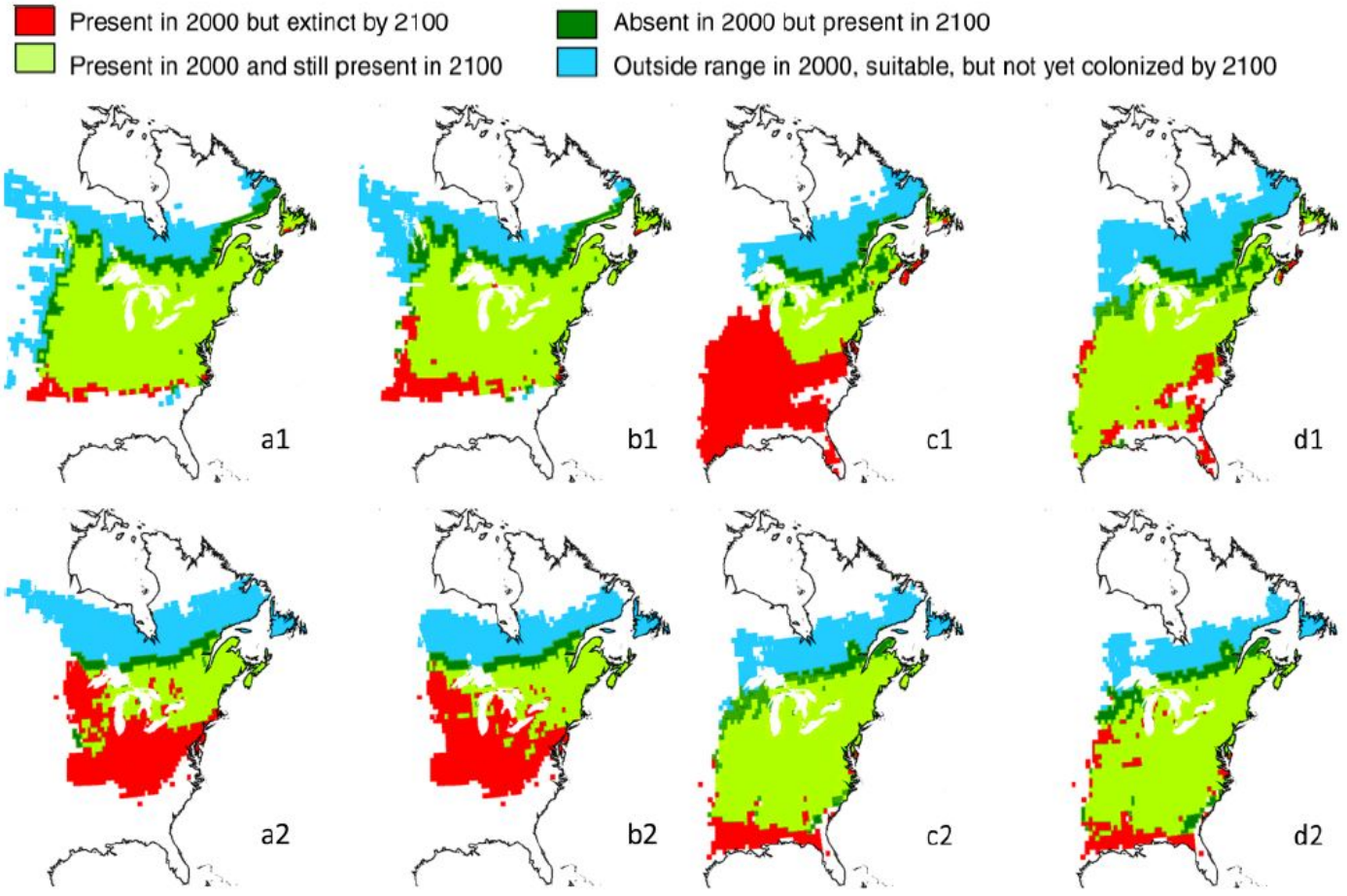
*Higgins et al. (2020)*




Spatial disagreement

# Ecologically explicit comparisons

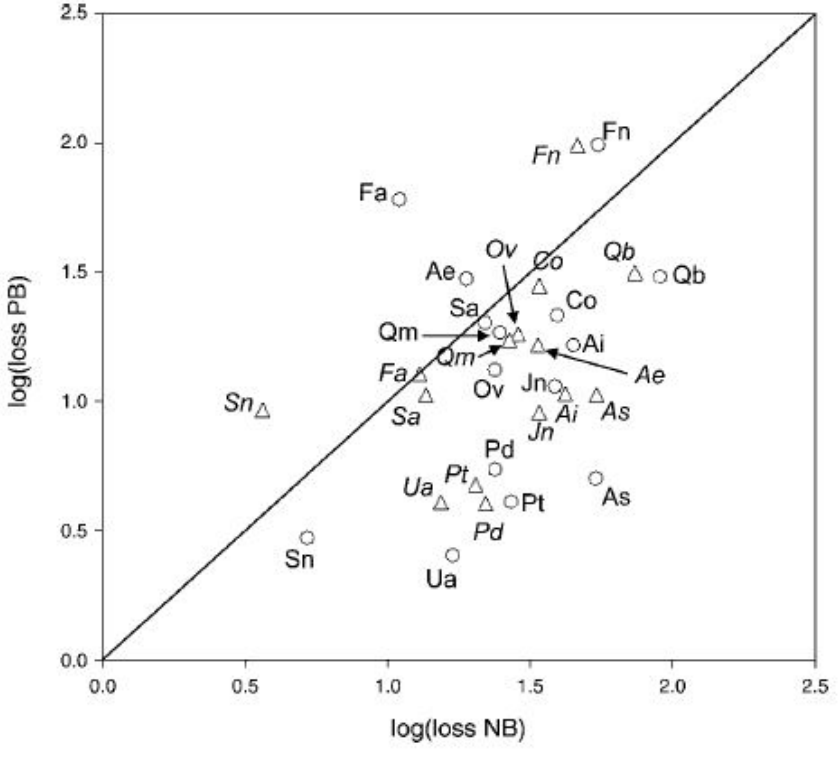
*Morrin et Thuiller (2009)*




Extinction, colonization

# Ecologically explicit comparisons

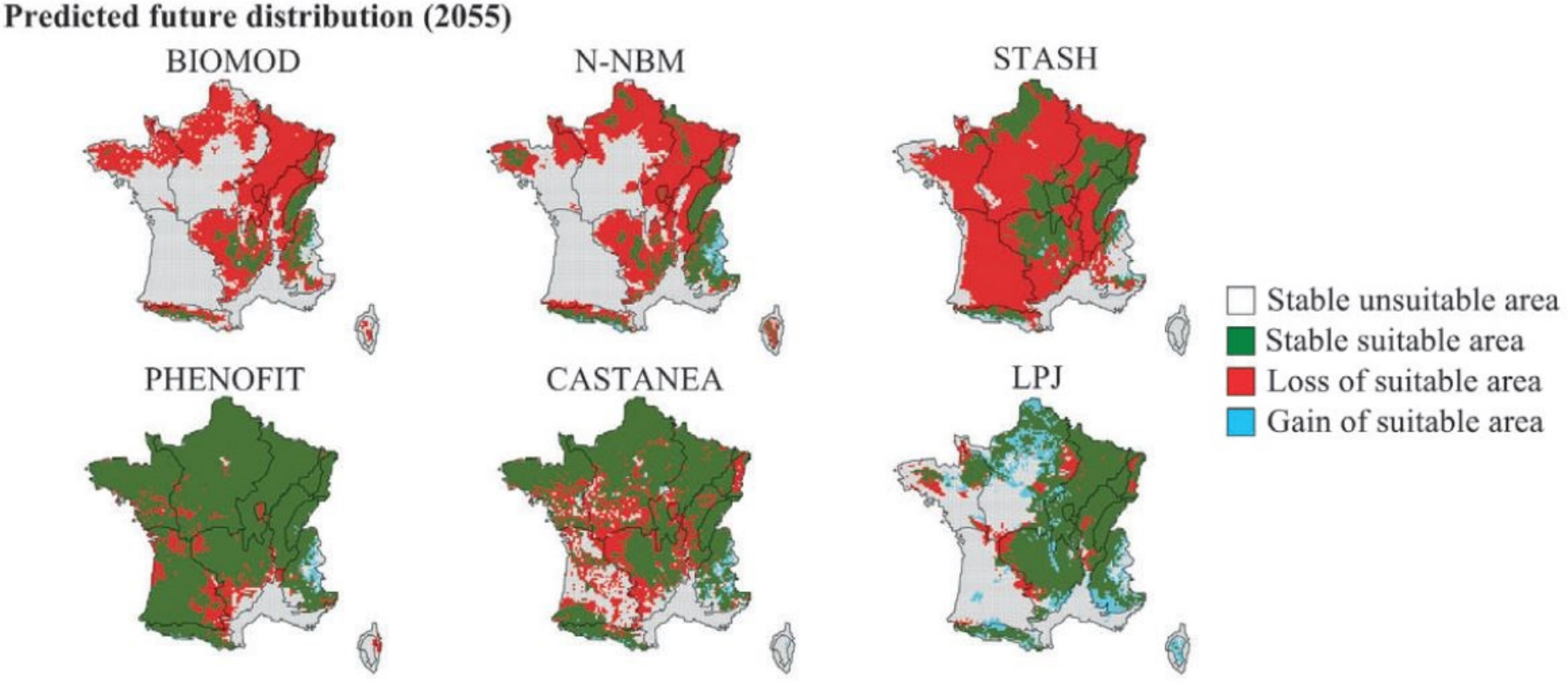
*Morrin et Thuiller (2009)*




Percentage of extinction

# Ecologically explicit comparisons

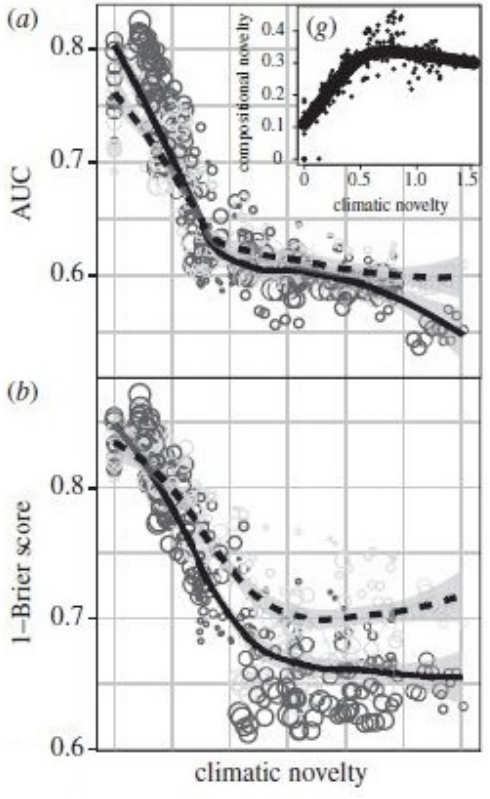
*Cheaib et al. (2012)*



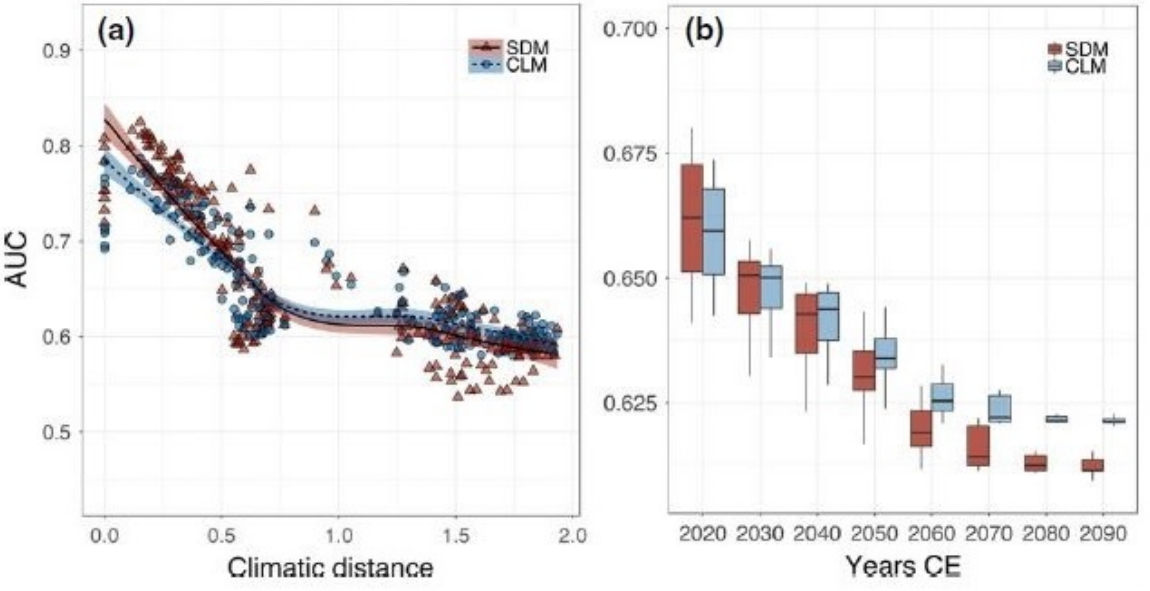

Suitable, unsuitable

# Climatically explicit comparisons

*Maguire et al. (2016)*

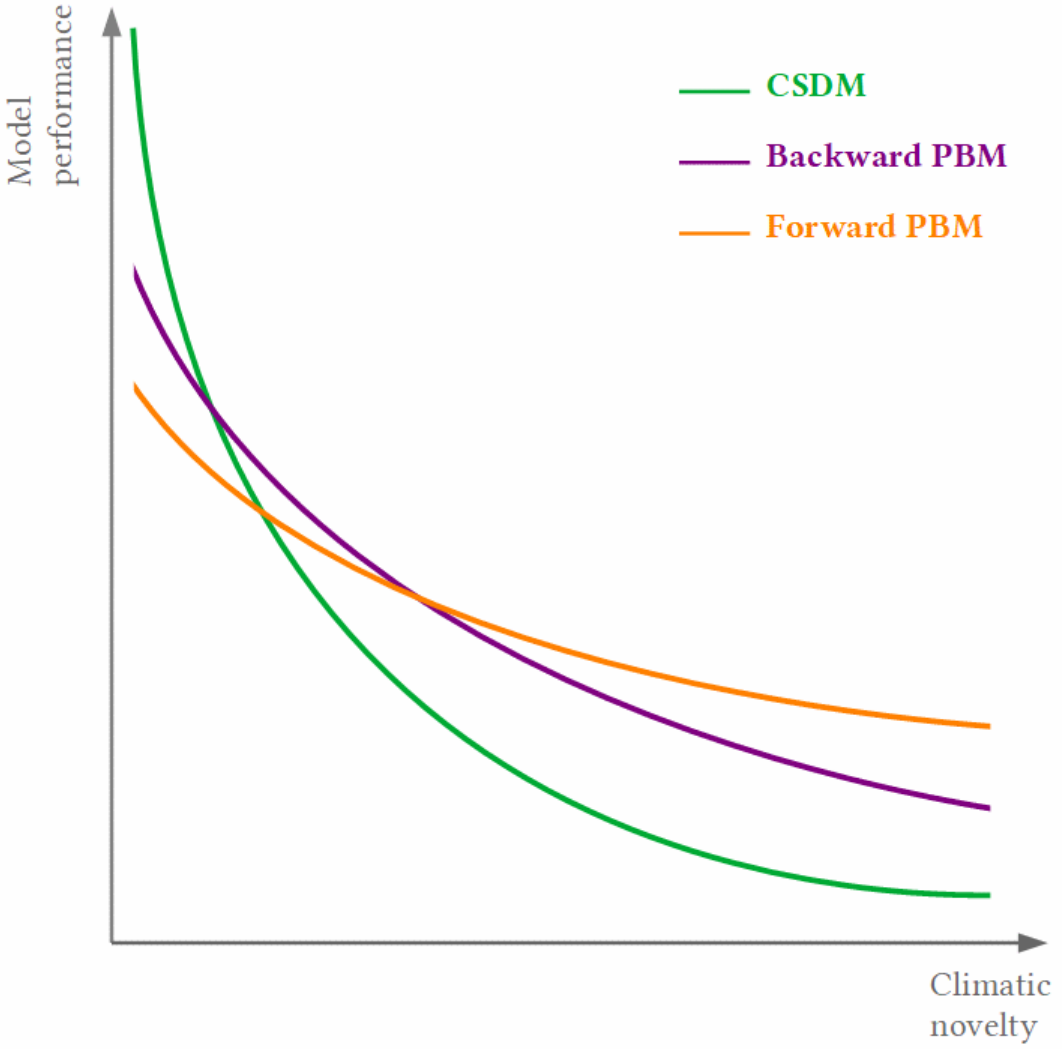



*Fitzpatrick et al. (2018)*

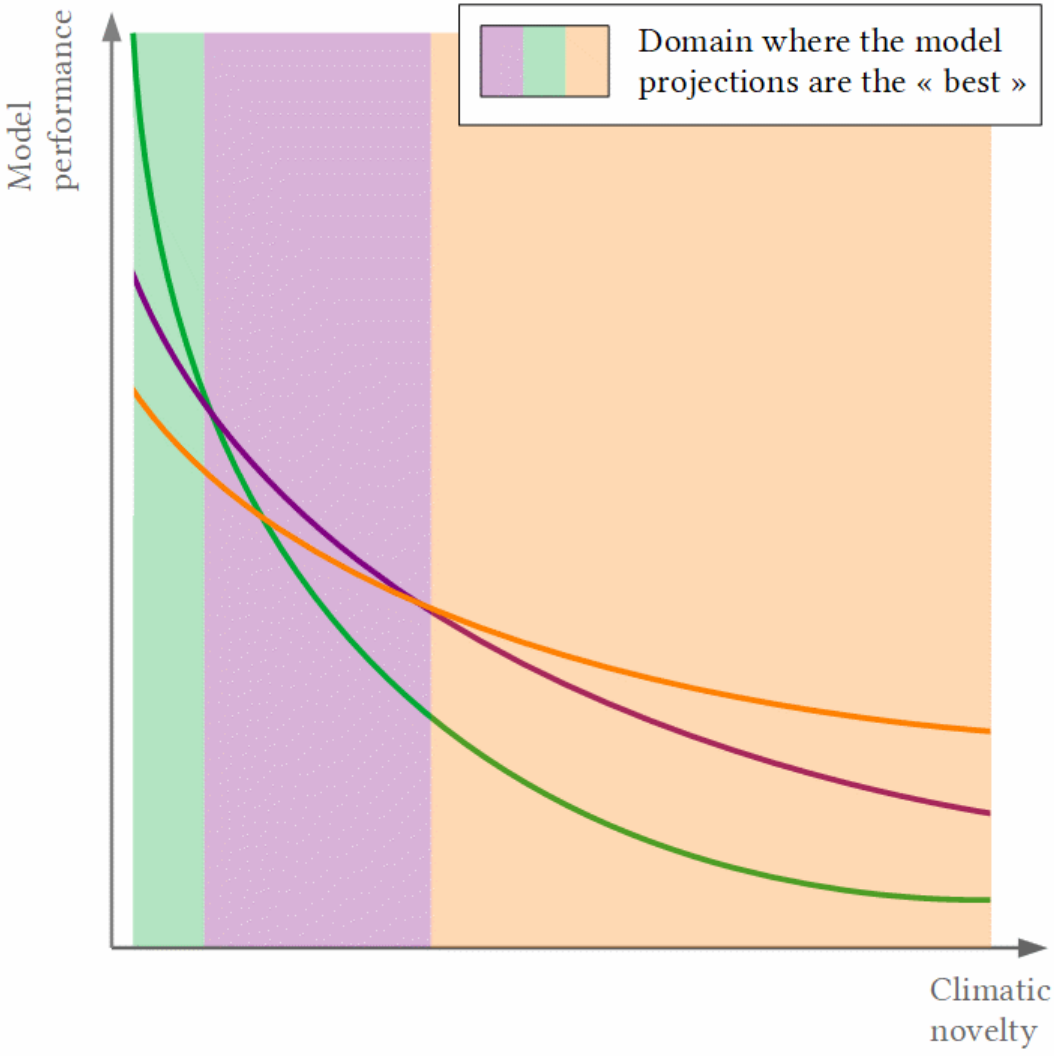



Performance under climatic novelty

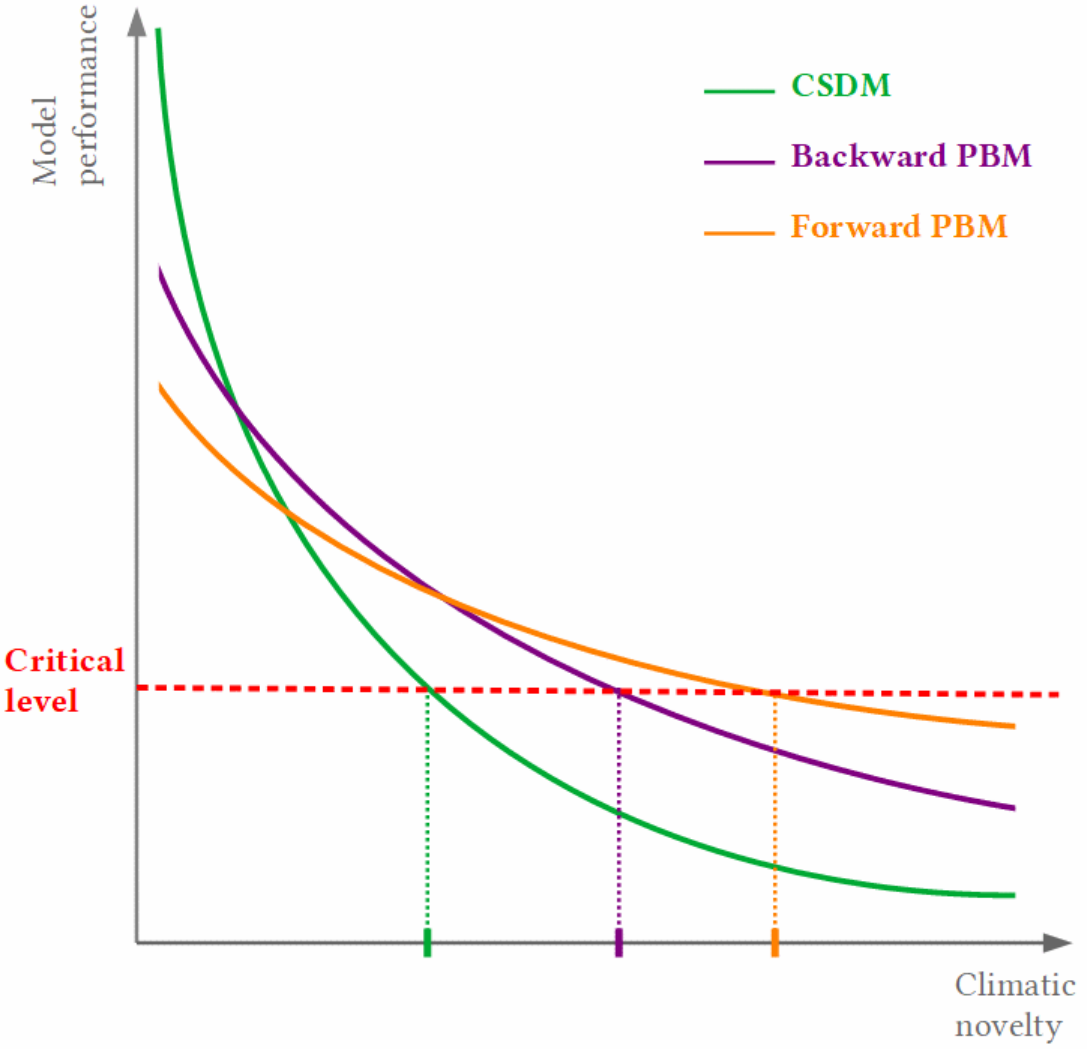
# A possible hypothesis ?



# A possible hypothesis ?

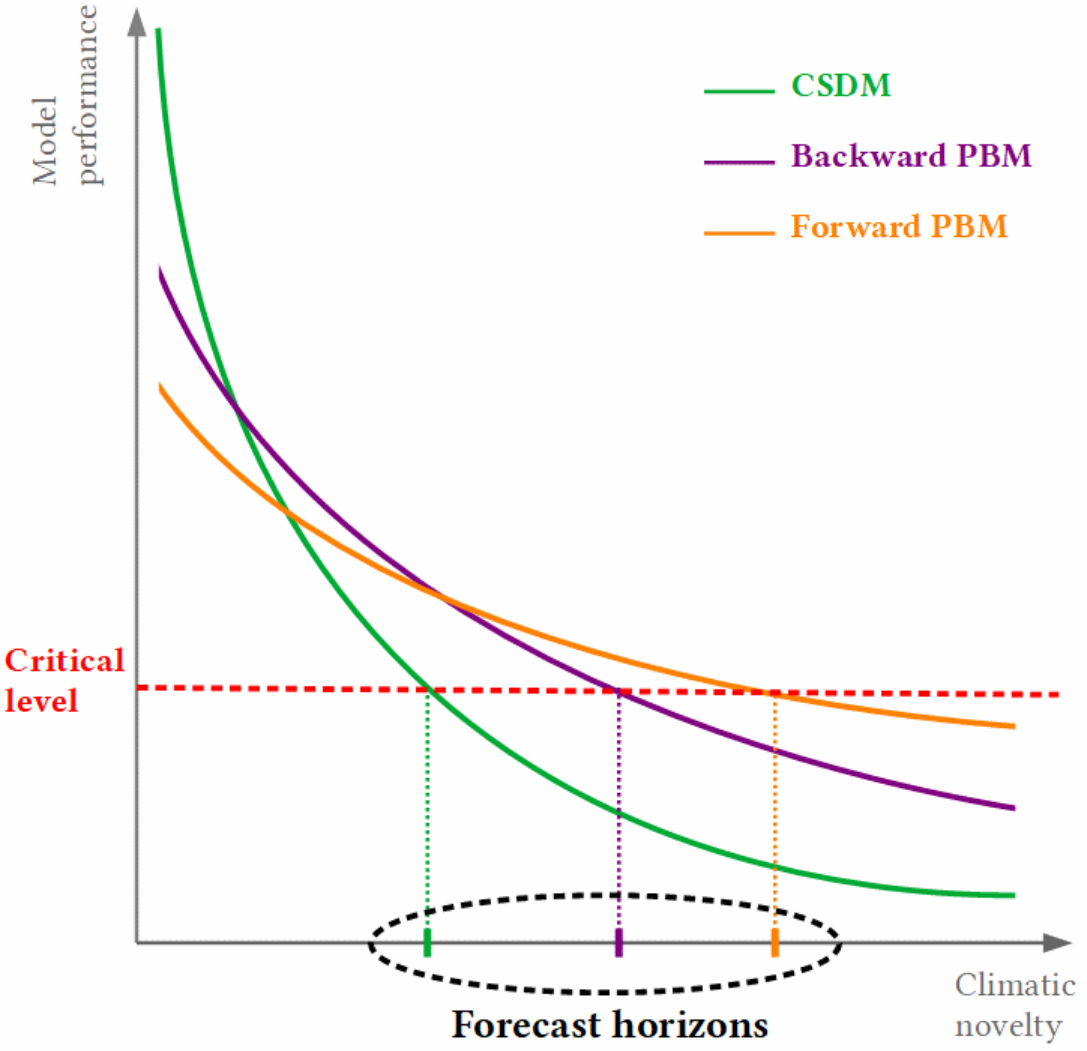


# A possible hypothesis ?

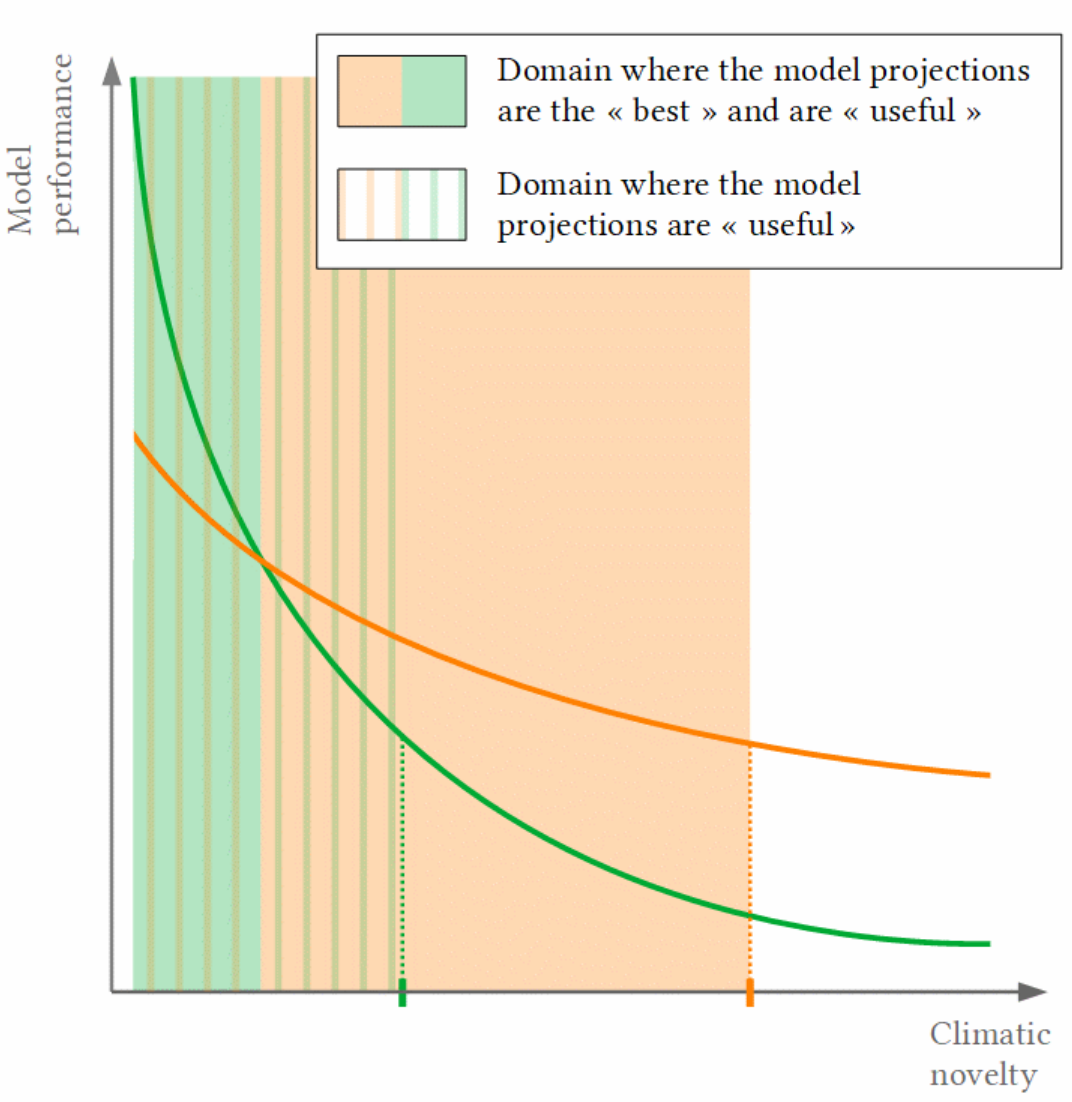


# A possible hypothesis ?

Define a "**forecast horizon**", beyond which sufficiently useful predictions can no longer be made (Petchey et al, 2015)

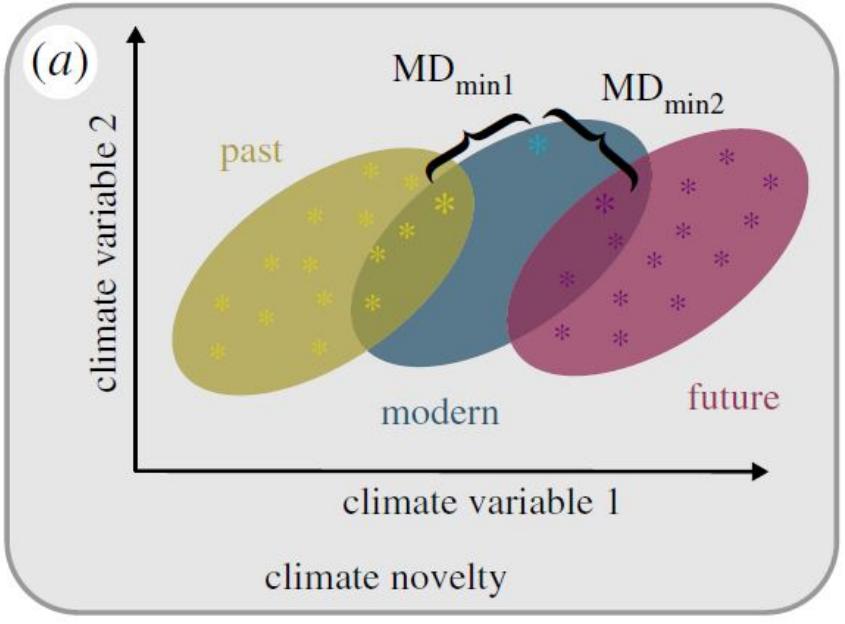


# A possible hypothesis ?



## Others climatic mechanisms ?

Climate novelty...

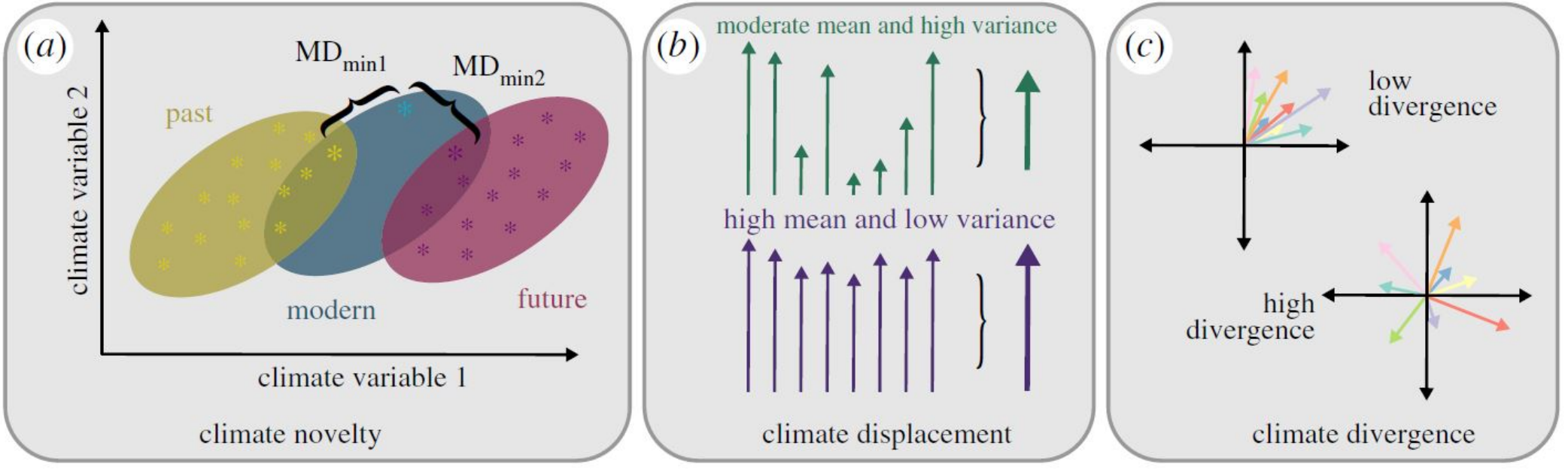



*Burke et al. (2019), adapted from Ordonnez et al. (2016)*

## Others climatic mechanisms ?

Climate displacement

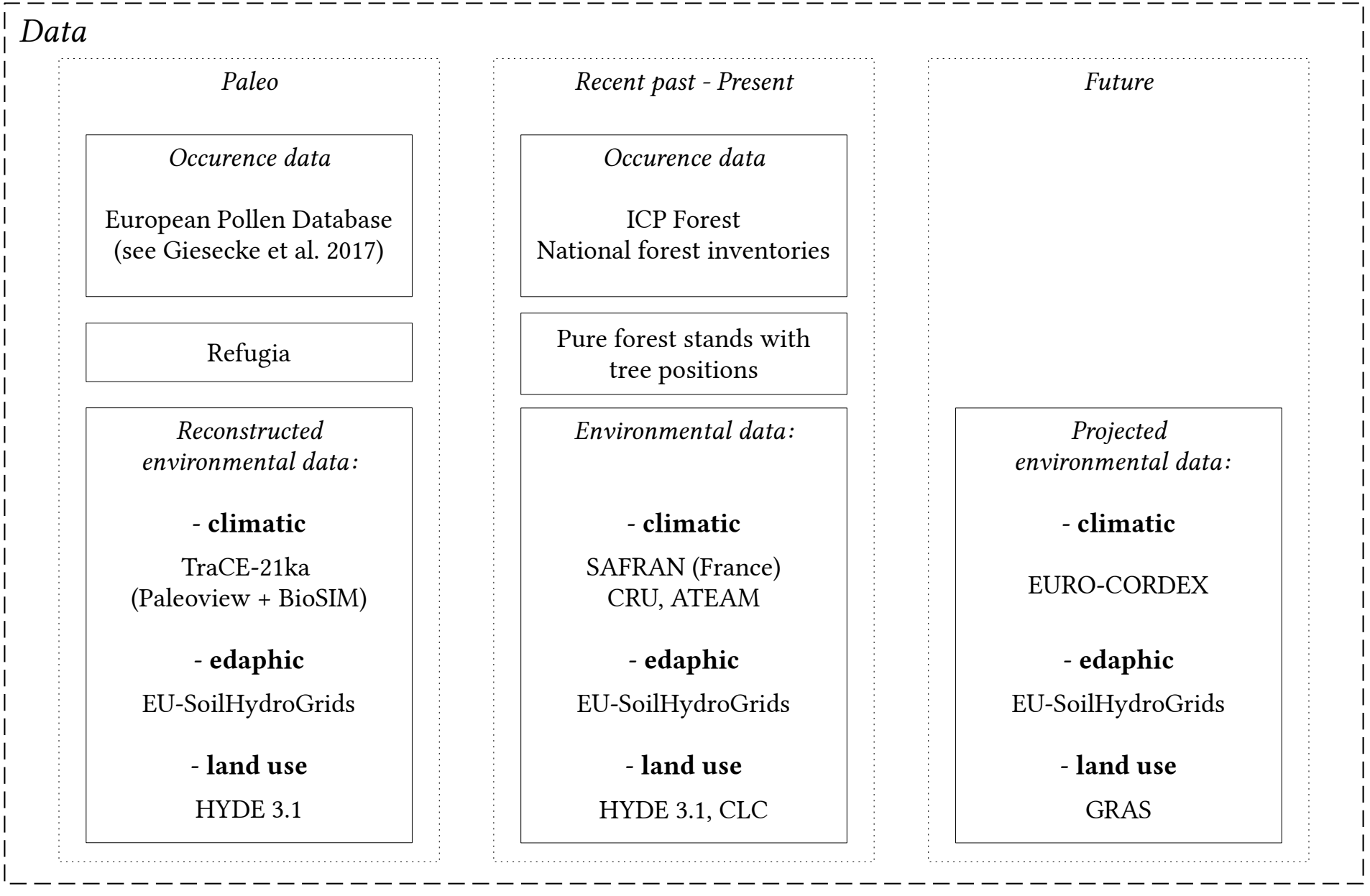
Climate divergence

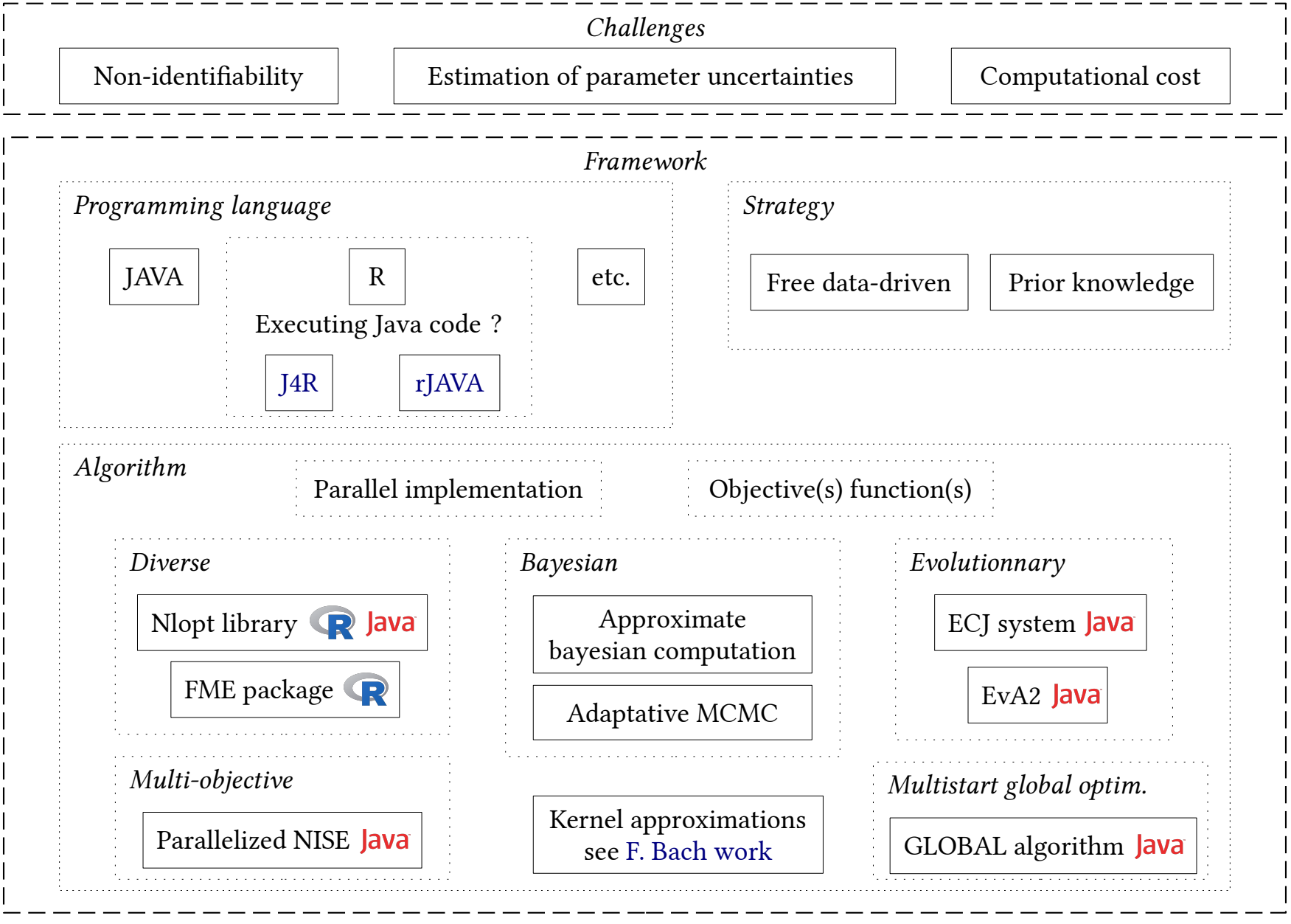



*Burke et al. (2019), adapted from Ordonnez et al. (2016)*

## Points à aborder

- Données
- Période de terrain
- Stratégie pour l'*inverse modelling*
- Interactions avec l'équipe en fonction du calendrier (fréquence, déplacements, outils de travail)
- Publications


Data
Paleo
Occurence data
European Pollen Database
(see Giesecke et al. 2017)
Refugia
Reconstructed environmental data:
- climatic
TraCE-21ka
(Paleoview + BioSIM)
- edaphic
EU-SoilHydroGrids
- land use
HYDE 3.1
Recent past - Present
Occurence data
ICP Forest
National forest inventories
Pure forest stands with tree positions
Environmental data:
- climatic
SAFRAN (France)
CRU, ATEAM
- edaphic
EU-SoilHydroGrids
- land use
HYDE 3.1, CLC
Future
Projected environmental data:
- climatic
EURO-CORDEX
- edaphic
EU-SoilHydroGrids
- land use
GRAS


Challenges
Non-identifiability
Estimation of parameter uncertainties
Computational cost
Framework
Programming language
JAVA
R
Executing Java code ?
J4R
rJAVA
etc.
Strategy
Free data-driven
Prior knowledge
Algorithm
Parallel implementation
Objective(s) function(s)
Diverse
Nlopt library
Java
FME package
Bayesian
Approximate bayesian computation
Adaptative MCMC
Evolutionnary
ECJ system
Java
EvA2
Java
Multi-objective
Parallelized NISE
Java
Kernel approximations see F. Bach work
Multistart global optim.
GLOBAL algorithm
Java

## Publications potentielles

- (Article 1 : méthode pour la calibration *backward*)
- Article 2 : Paléo, passé lointain
- Article 3 : Passé récent
- Article 4 : Futur
- (Article 5 : incertitudes des modèles, ensemble des périodes)