



Introduction à la modélisation de la distribution des espèces

Agathe Leriche
agathe.leriche@univ-cezanne.fr

Crédits : « *Species distribution modeling for conservation educators and practitioners* » de R.G. Pearson et de nombreuses publications citées

Contenu

- ▶ Pourquoi ?
- ▶ Quoi ?
- ▶ Comment ?
 - ▶ Les modèles
 - ▶ Modèles empiriques
 - ▶ Modèles mécanistes
- ▶ Conclusions

Pourquoi ?

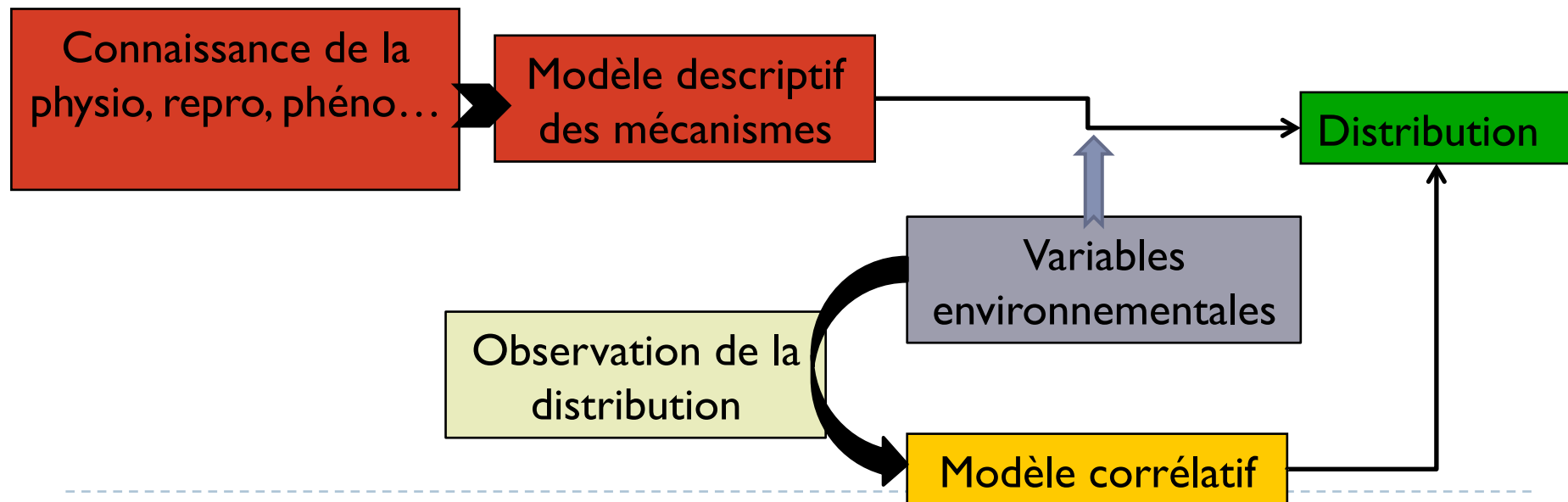
- ▶ Applications de la modélisation de la distribution des espèces :
 - ▶ Gestion et conservation
 - ▶ Guider l'échantillonnage pour détecter une espèce connue
 - ▶ Guider l'échantillonnage pour accélérer la détection de nouvelles espèces
 - ▶ Définir zone de protection
 - ▶ Guider la réintroduction d'espèces menacées
 - ▶ Projection de l'impact des CG et des activités anthropiques
 - ▶ Biogéographie, paléodistribution
 - ▶ Explorer les mécanismes de spéciation, phylogéographie
 - ▶ Risques biologiques, espèces invasives

Quoi ?

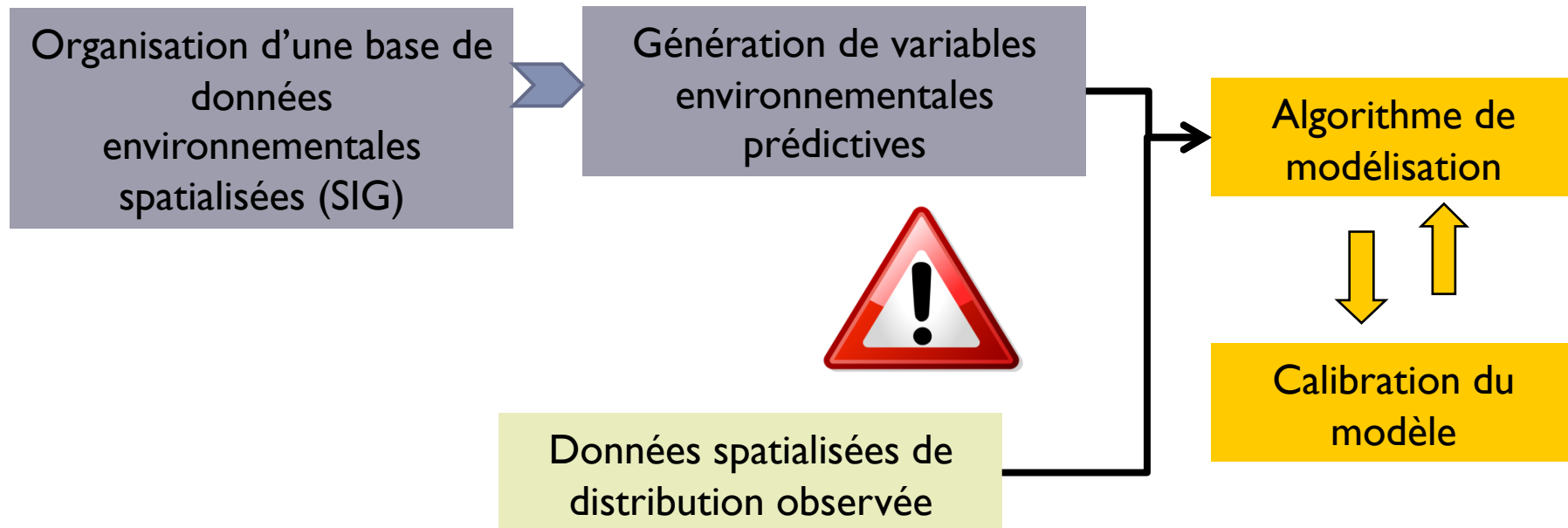
- ▶ **Important : attentions aux termes**
 - ▶ Modèle de distribution des espèces
 - ▶ Modèle de la niche écologique des espèces
 - ▶ Modèle de la niche environnementale des espèces
 - ▶ Modèle de l'habitat des espèces
 - ▶ Modèle de niche bioclimatique
 - ▶ ...
- ▶ **On ne modélise pas :**
 - ▶ la distribution des espèces
 - ▶ La niche écologique des espèces
- ▶ **On modélise:**
 - ▶ L'espace où les variables environnementales sont favorables à la persistance de l'espèce

Comment ?

- ▶ Utilisation de modèles
 - ▶ Représentation mathématique simplifiée d'un processus
- ▶ Les types de modèles
 - ▶ Modèles mécanistes
 - ▶ Modèles empiriques ou corrélatifs ou statistiques

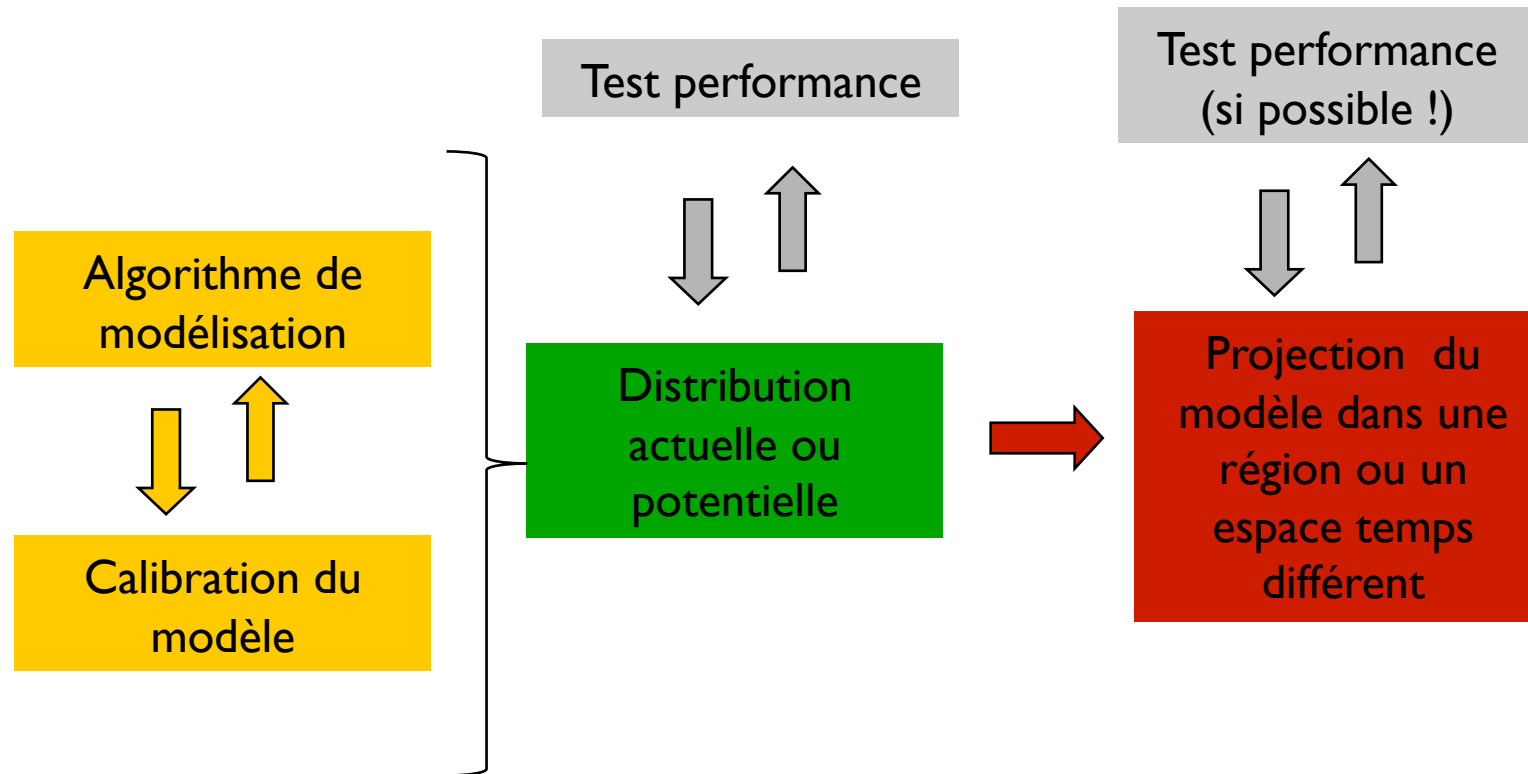


Les modèles EMPIRIQUES



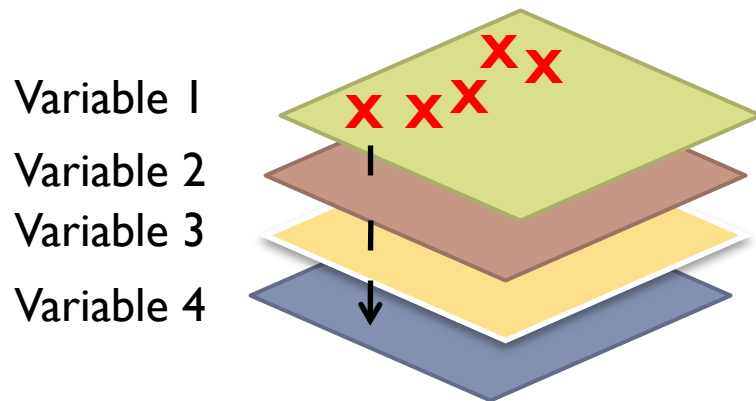
Exemple d'organigramme des étapes principales de génération d'un modèle empirique

Les modèles EMPIRIQUES



Exemple d'organigramme des étapes principales de génération d'un modèle empirique (suite)

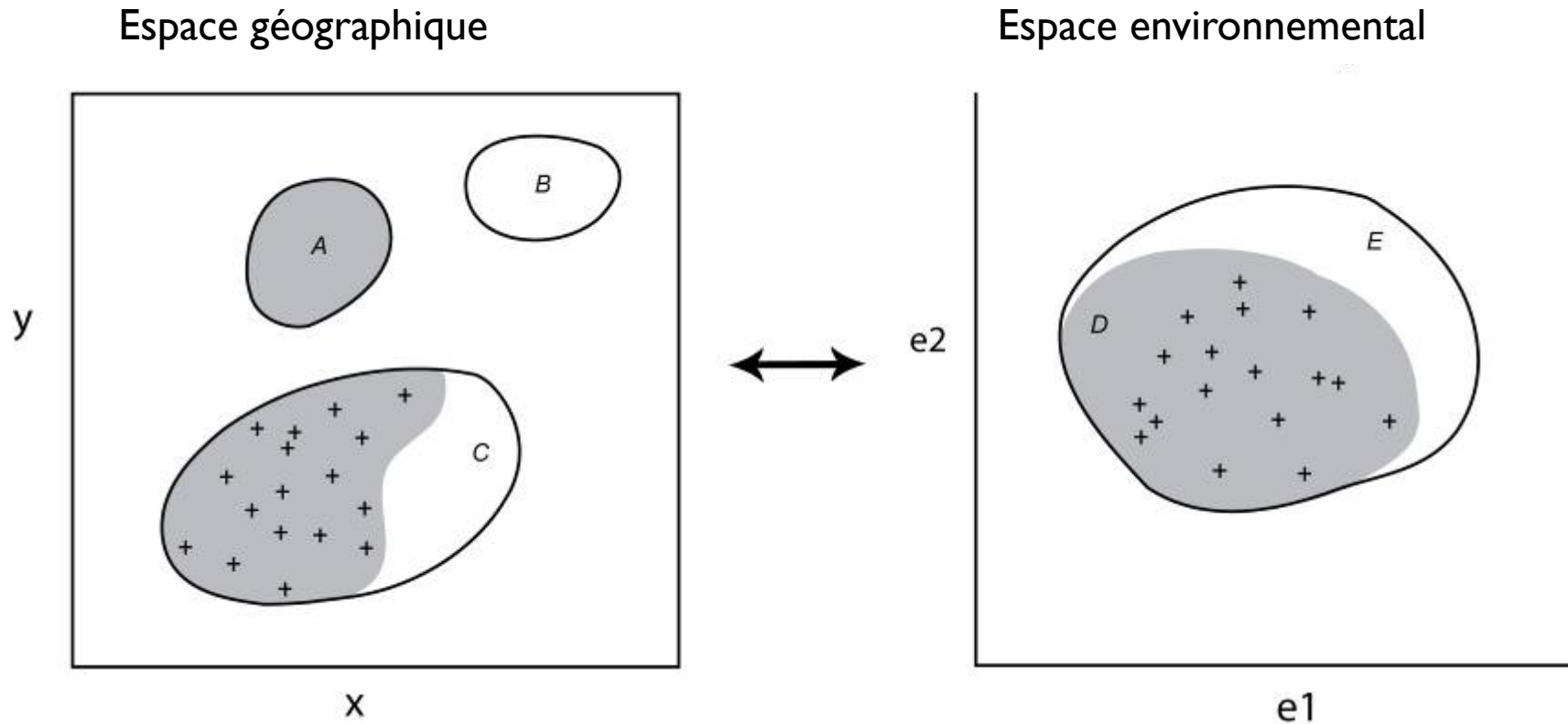
Les modèles EMPIRIQUES



	Pt1	Pt2	Pt3	Pt4	...
Var1	25	12	24	20	
Var2	125	151	171	123	
Var3	10	8	4	12	
Var4	456	587	125	502	

→ Modèle statistique

Relation entre la répartition de l'espèce dans l'espace géographique et dans l'espace environnemental

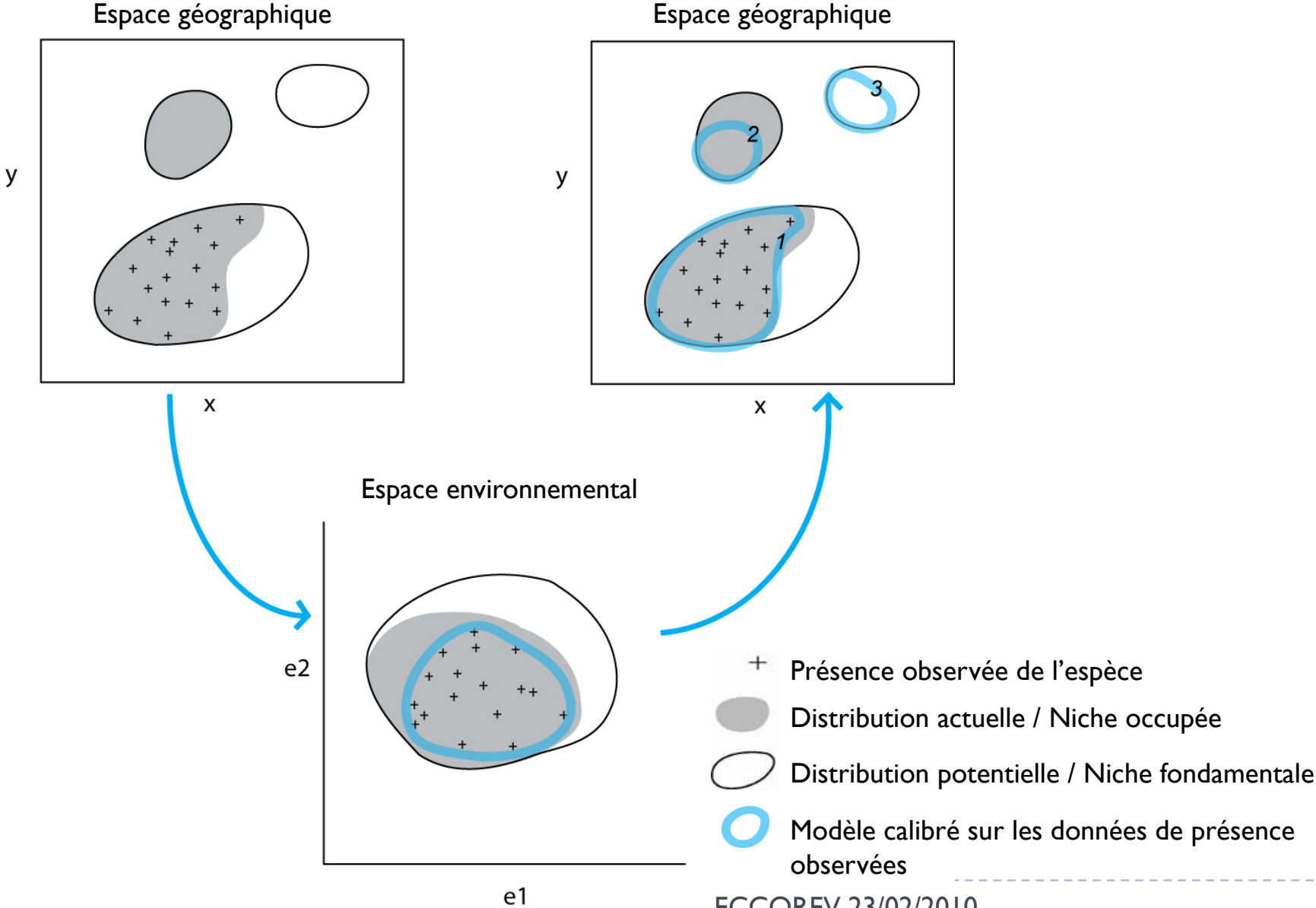


+ Présence observée de l'espèce

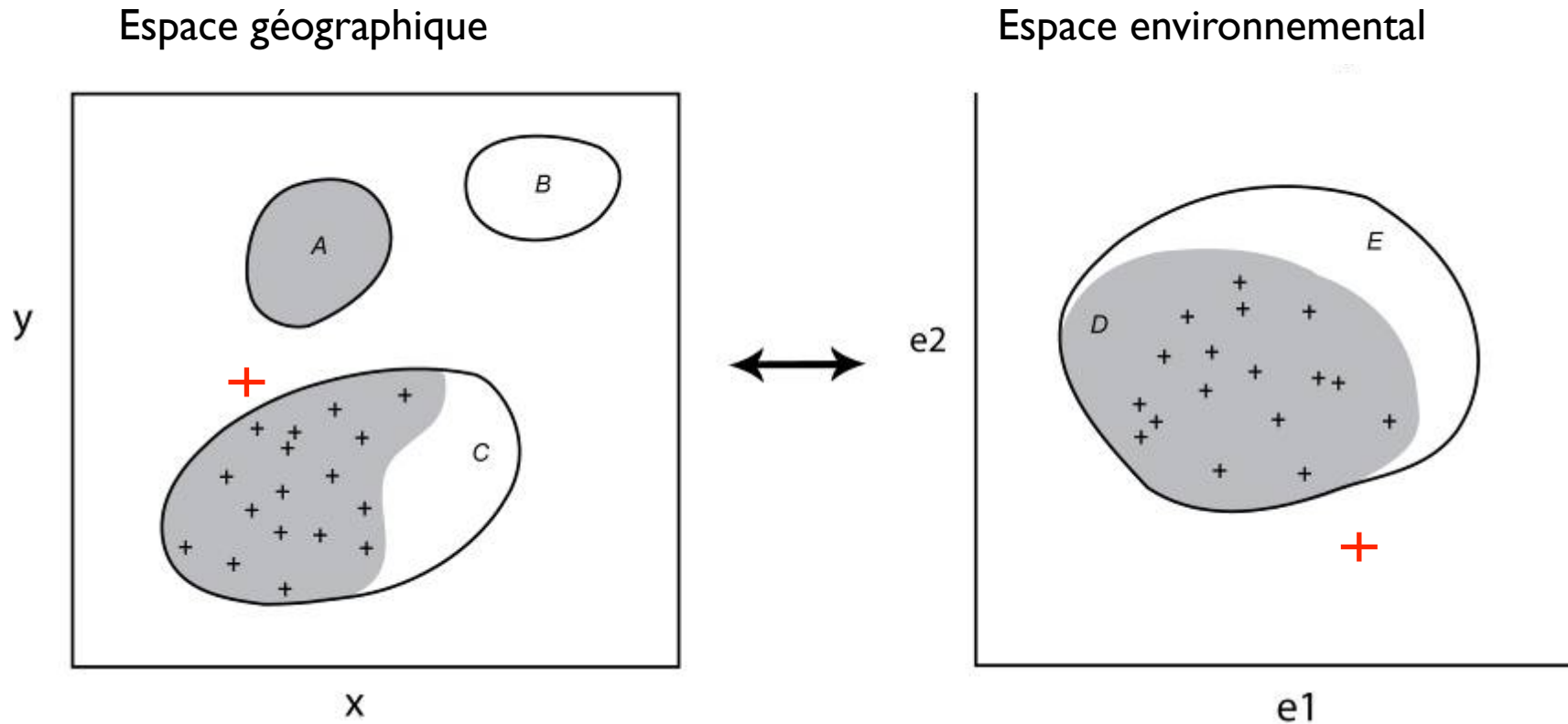
● Distribution actuelle / Niche réalisée

○ Distribution potentielle / Niche fondamentale

Illustration de l'approche de modélisation



Relation entre la répartition de l'espèce dans l'espace géographique et dans l'espace environnemental




+ Présence observée de l'espèce

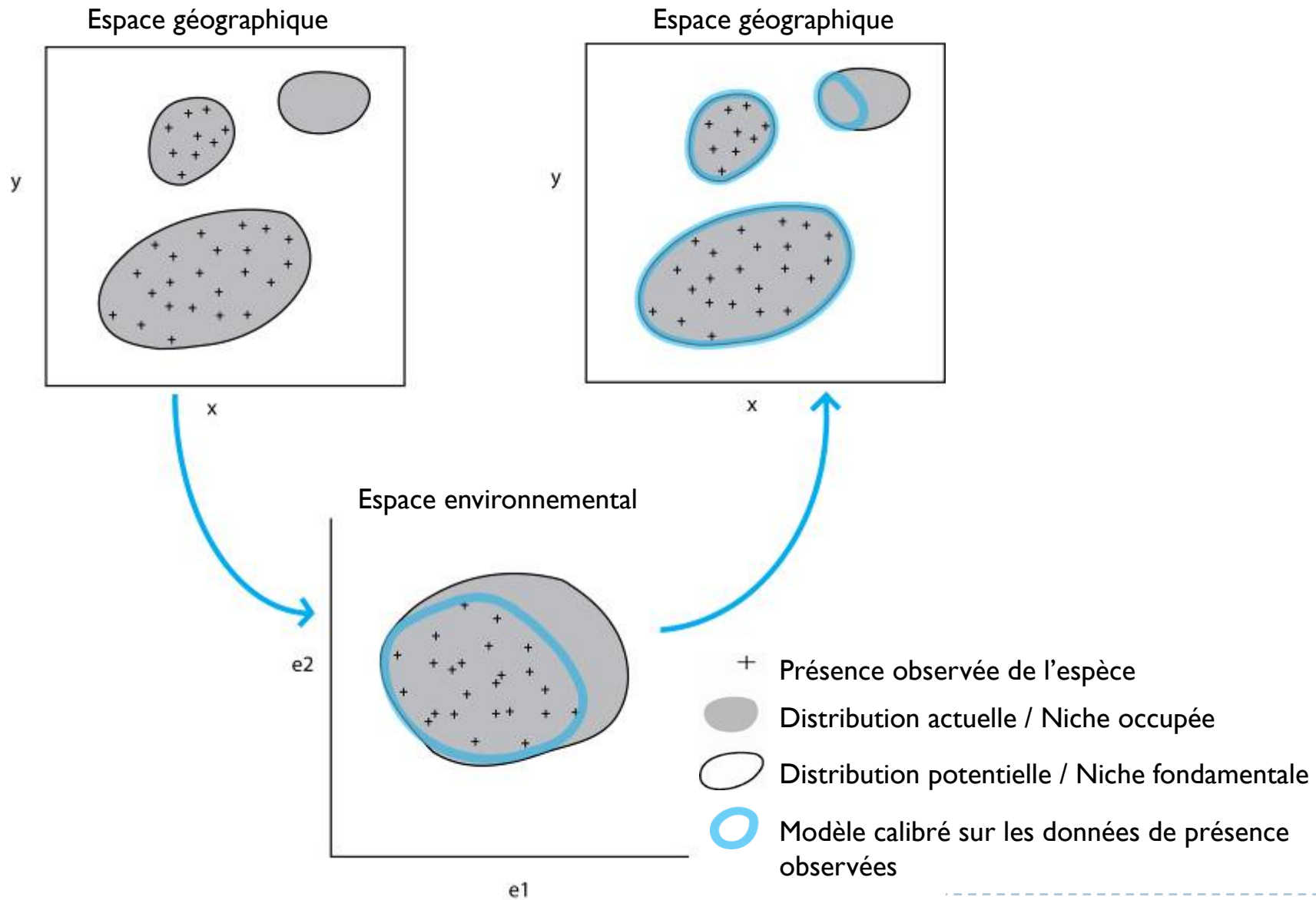
● Distribution actuelle / Niche réalisée

○ Distribution potentielle / Niche fondamentale

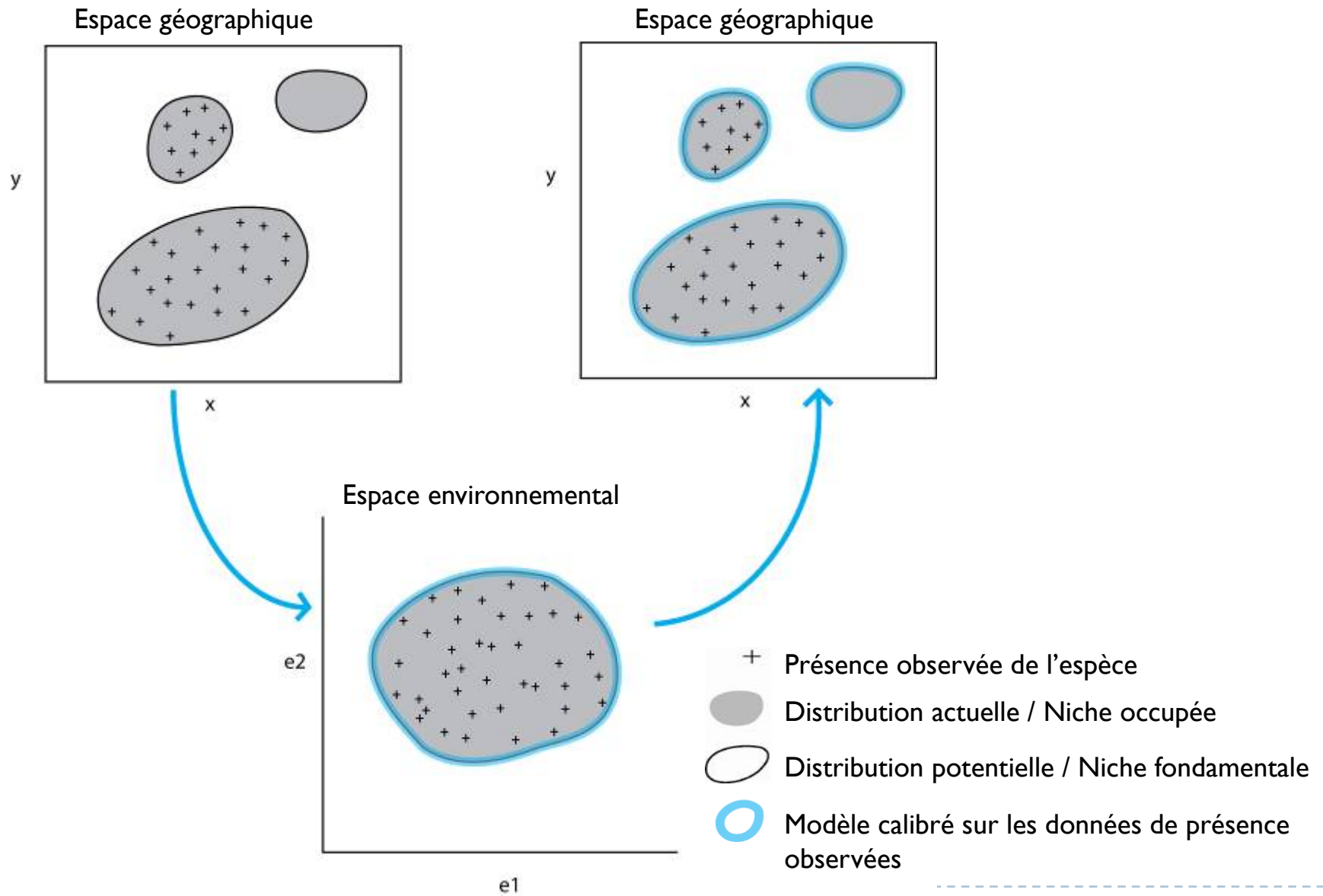
Les modèles EMPIRIQUES

- ▶ Deux facteurs déterminent le degré de pertinence d'un modèle de distribution
 - ▶ Le degré d'équilibre de l'espèce
 - ▶ % de la niche réalisée potentielle ?
 - ▶ La représentativité de l'échantillonnage
 - ▶ Etendue de l'espace environnemental représentée ?
 - ▶  pas de lien direct entre échantillonnage dans l'espace géographique et dans l'espace environnemental !
- ▶ Conclusion : bien analyser ses données avant toute modélisation !

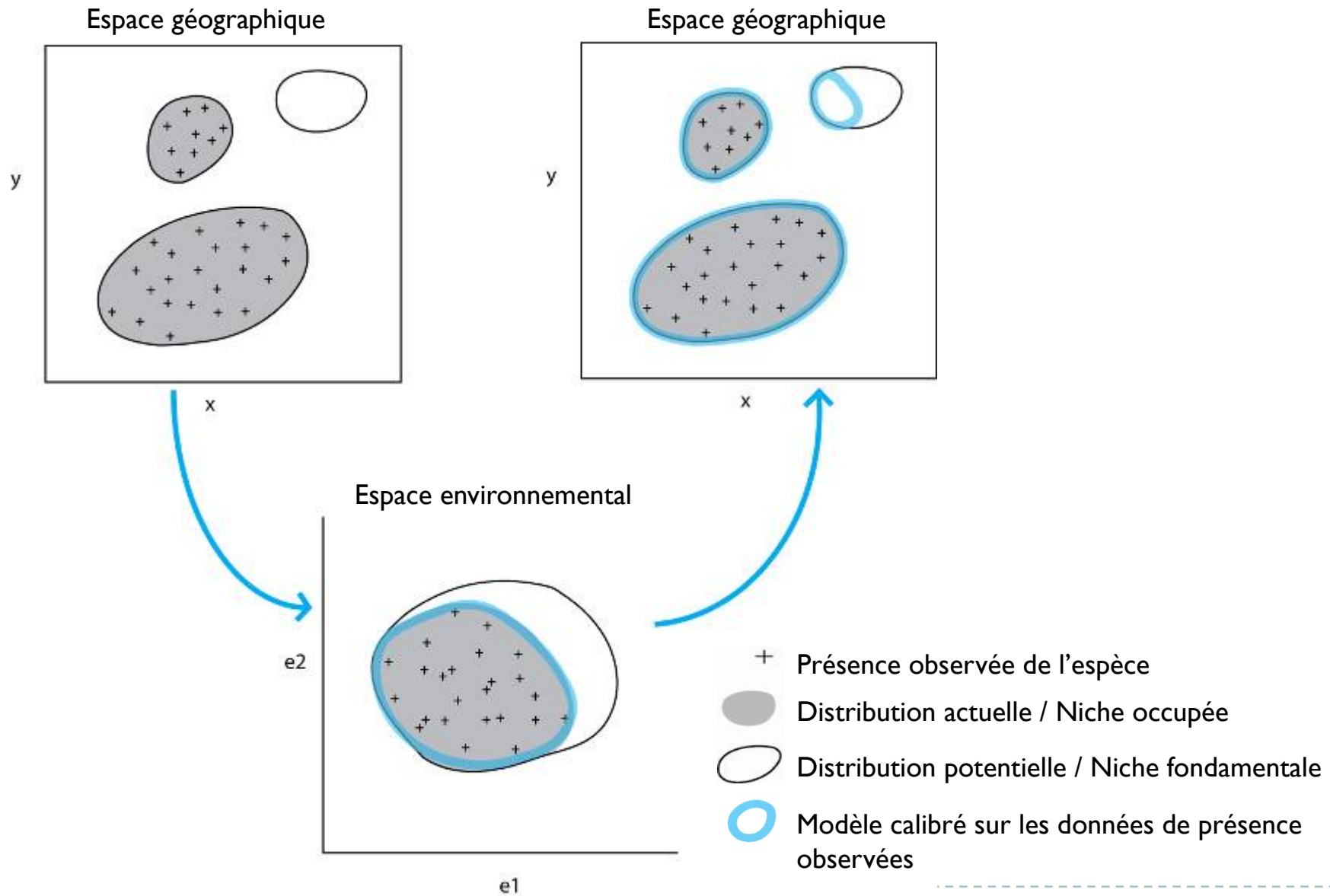
Fort équilibre mais faible échantillonnage (espace géo et environmental)



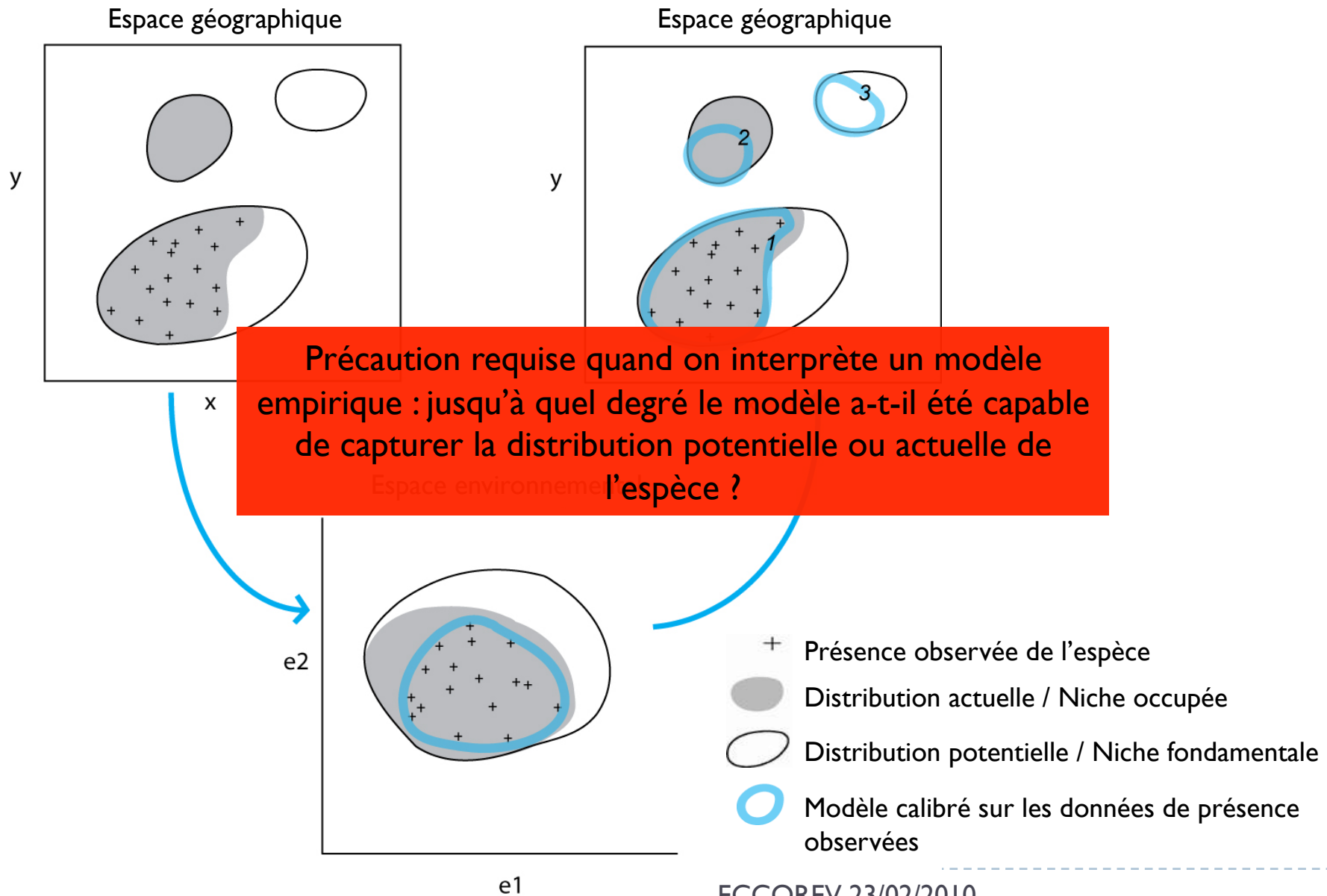
Fort équilibre, faible échantillonnage géographique mais bon échantillonnage dans l'espace environnemental



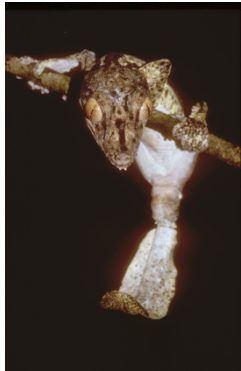
Faible équilibre mais bon échantillonnage



Exemple d'origine : combinaison d'équilibre moyen et d'échantillonnage incomplet



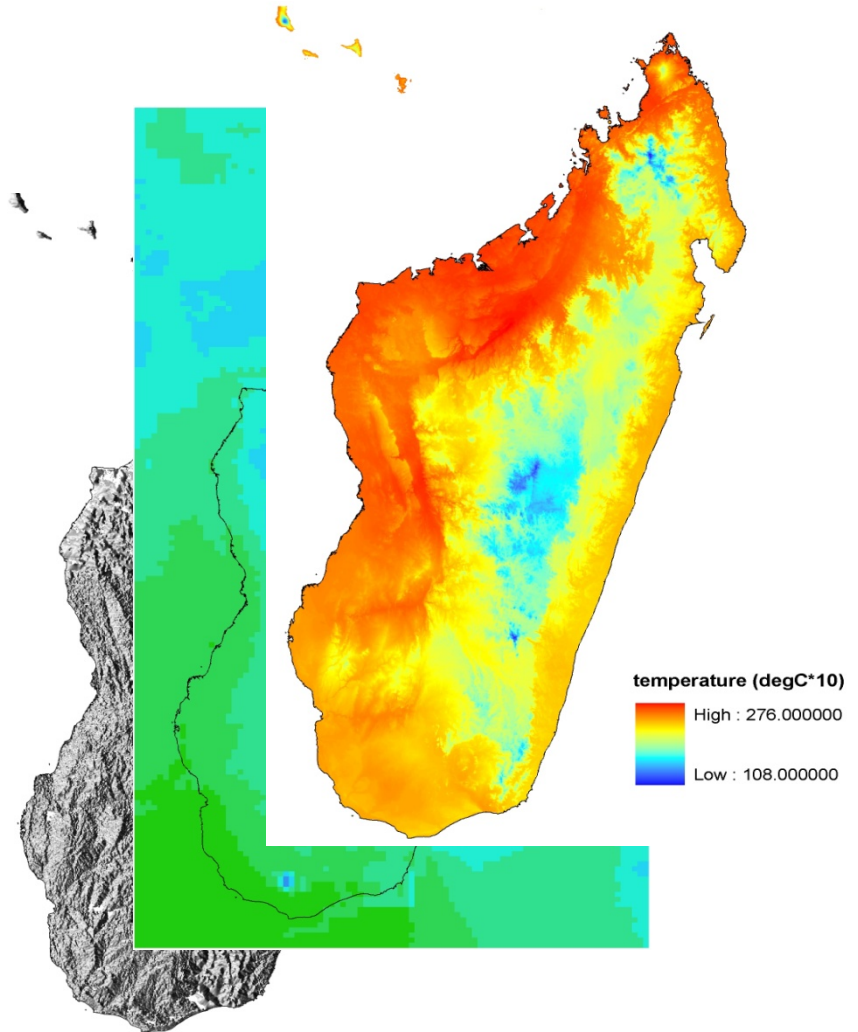
Les modèles EMPIRIQUES : les données



Uroplatus sp.
(leaf-tailed gecko)



Distribution connue de l'espèce

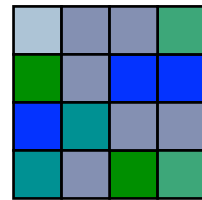


Données environnementales

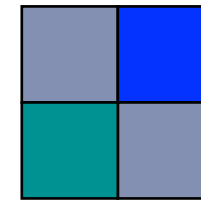
Les modèles EMPIRIQUES : les données

▶ Notion importante : échelle des données

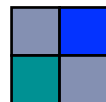
▶ Résolution



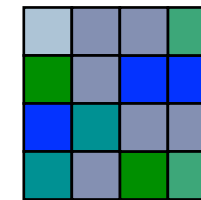
Diminution de la résolution



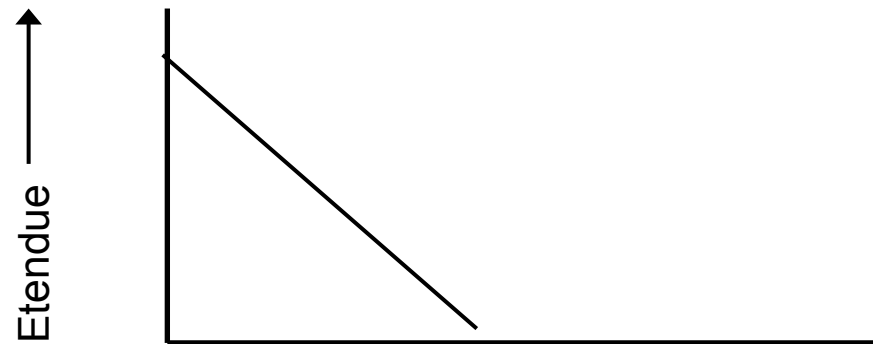
▶ Étendue



Accroissement de l'étendue

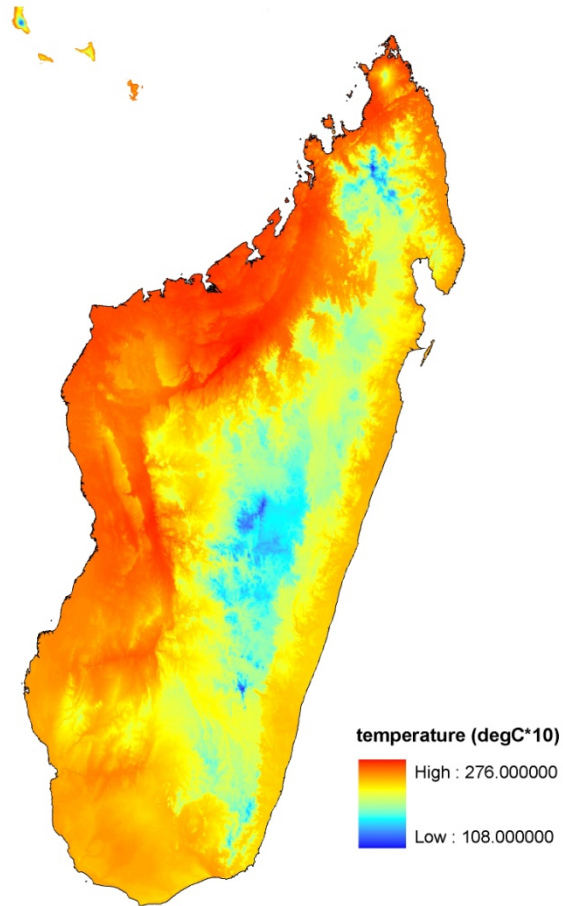


En pratique, la
résolution et l'étendue
tendent à être
inversement reliées

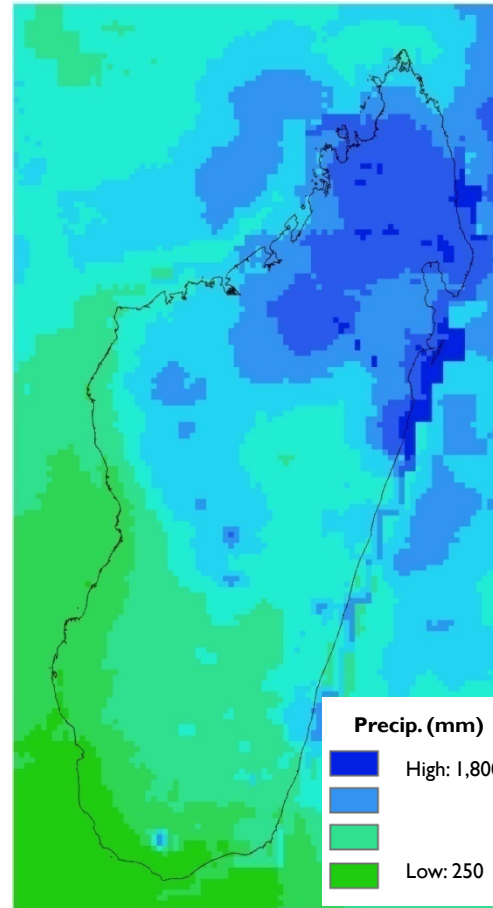


Résolution →
ECCOREV 23/02/2010

Les modèles EMPIRIQUES : les données



Température moyenne annuelle
Source: WorldClim

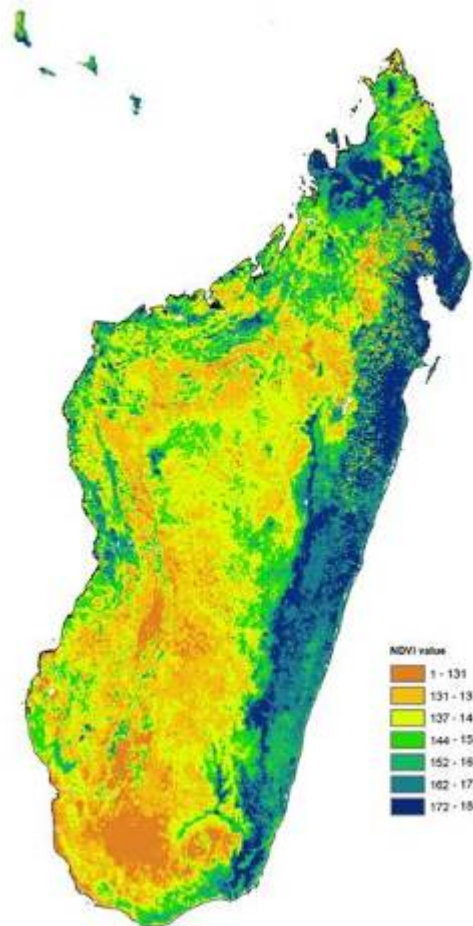


Précipitation moyenne annuelle
Source: NOAA FEWS

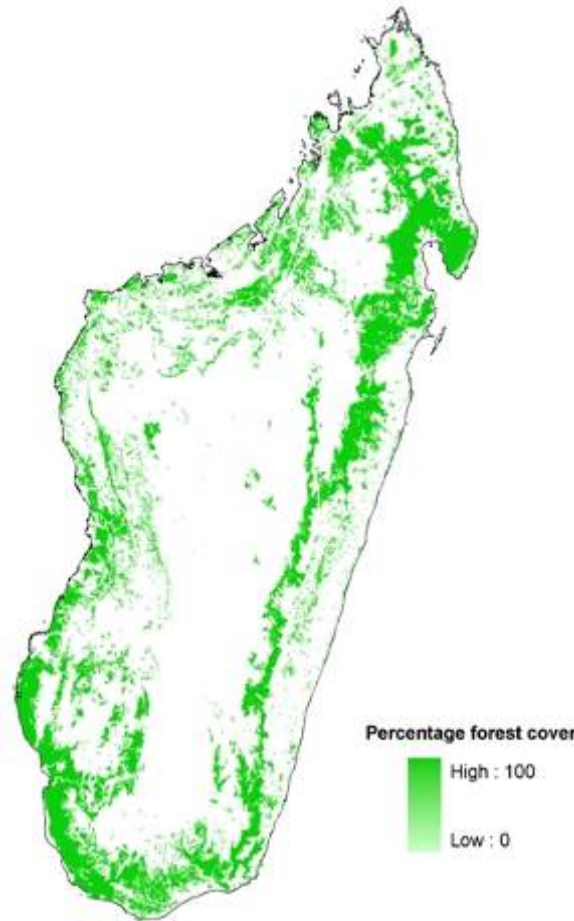


Exposition : East-West
Source: USGS Hydro 1k

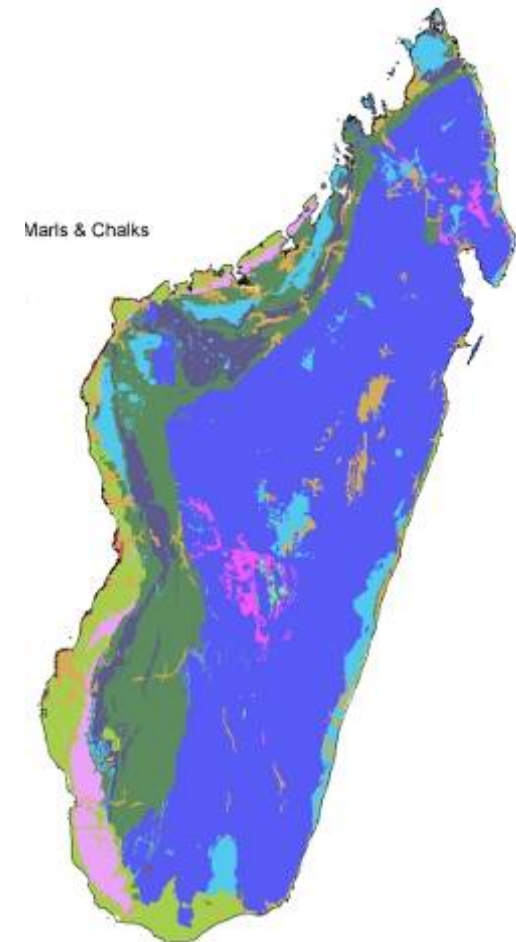
Les modèles EMPIRIQUES : les données



AVHRR NDVI
Source: NASA



% couverture forestière
Source: IEFN and CI



Geologie
Source: Kew gardens
(note: données catégorielles)

Les modèles EMPIRIQUES : les données

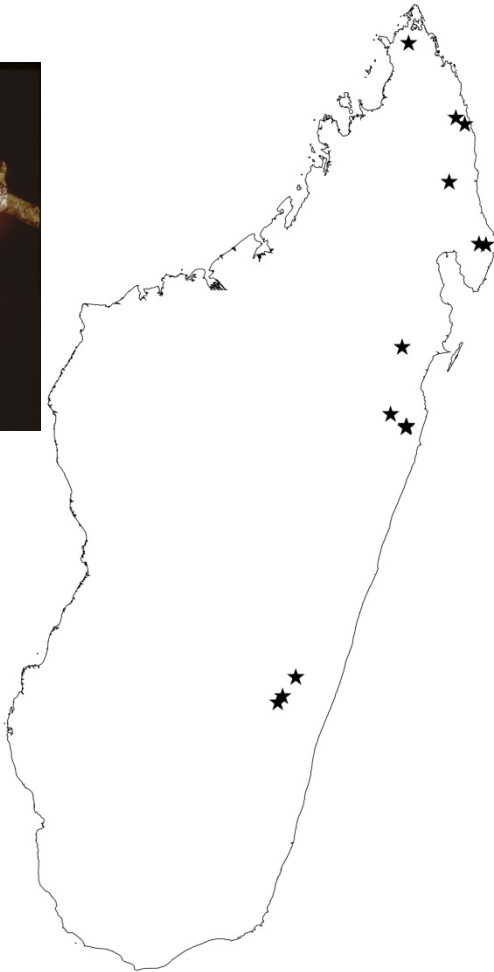
- ▶ **Extraction de variables directes...**
 - ▶ Préférable d'utiliser des variables qui ont un impact DIRECT physiologique (où sur le comportement etc.)
 - ▶ L'association de variables indirectes avec la distribution observée conduit à des erreurs lors de la projection

- ▶ **Exemple : altitude ou température + pression atmosphérique ?**

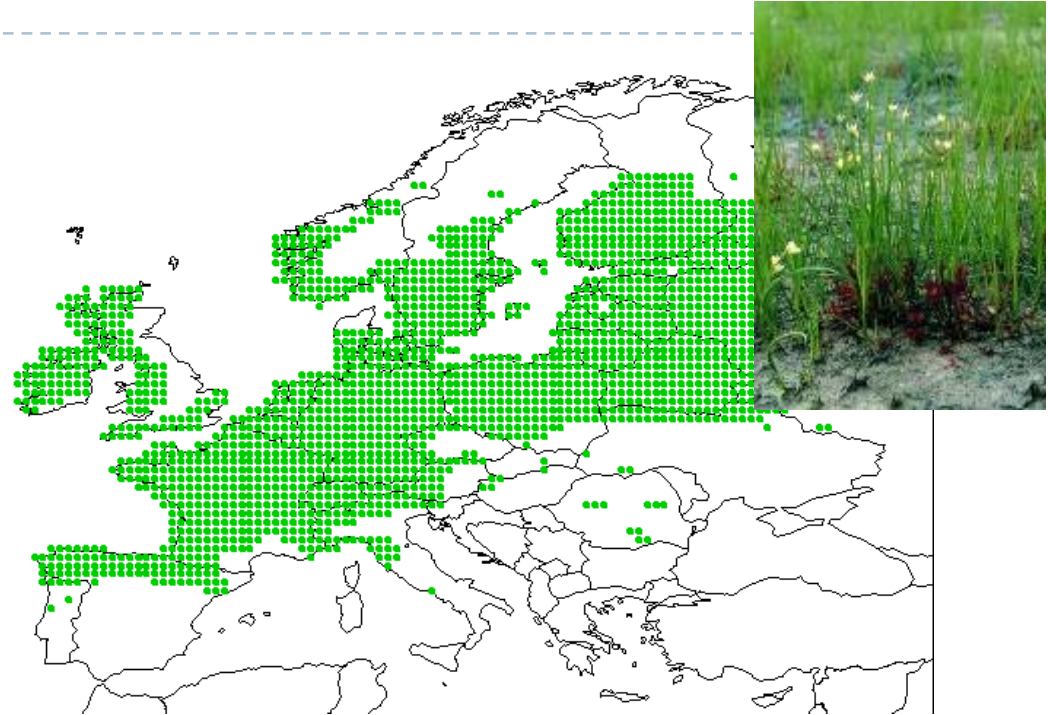
Les modèles EMPIRIQUES : les données

- ▶ **Sources des données de distribution**
 - ▶ Collecte « personnelle » : obtenue lors de campagne de terrain réalisée sur de petites zones dans le cadre d'études ciblées
 - ▶ Collecte à grande échelle : nombreuses personnes, grande zone couverte (exemple des atlas)
 - ▶ Collection des muséums
 - ▶ Ressources en ligne : nombreux organismes de recherches mettent en ligne leur données
- ▶ **Biais des données :**
 - ▶ Identification incorrecte
 - ▶ Mauvais référencement spatial
 - ▶ Centroïde de la région au lieu de la localisation précise

Les modèles EMPIRIQUES : les données



Données de présence seules
d'un gecko à Madagascar



Données de présence et absence présumées
pour une plante en Europe

**Données de distribution d'une espèce : présence
seule ou présence/absence?**

Les modèles EMPIRIQUES : les données

Question : est- ce qu'une absence observée correspond réellement à une véritable absence de conditions favorables ?

1. L'espèce pouvait être présente, mais non détectée
2. L'espèce pouvait être réellement absente malgré un environnement favorable (e.g. limitations de dispersion, dynamique des pop)
3. L'environnement est réellement non favorable

'Fausse absence'

'Fausse absence'

'Absence vraie'

Bien réfléchir à l'utilisation de donnée "d'absence"



Les modèles EMPIRIQUES : Algorithmes

- ▶ **Algorithme :**
 - ▶ Probabilité de présence = f (variables explicatives)
 - ▶ Identifier les relations potentiellement complexes et non-linéaires dans un environnement multi-dimensionnel
- ▶ **Le meilleur modèle**
 - ▶ Maximise le nombre de points ajustés
 - ▶ Minimise l'ajustement de zones sans observation
- ▶ = « Modéliser tout ce que l'on sait sans modéliser ce que l'on ne sait pas »

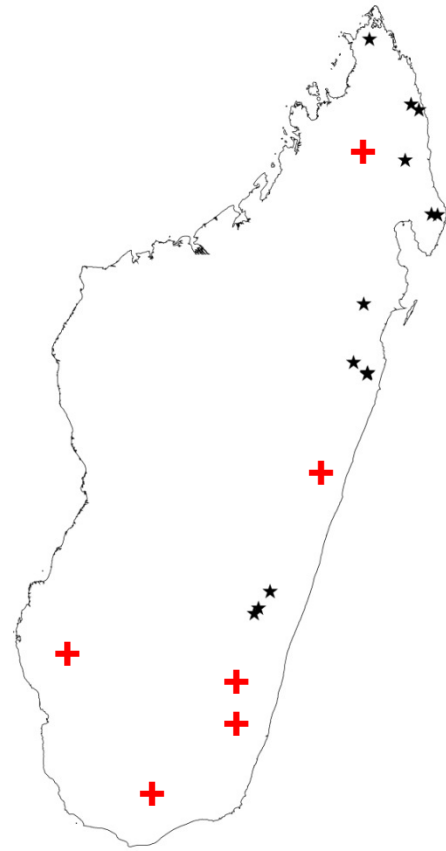
Les modèles EMPIRIQUES : Algorithmes

- ▶ **Nombreux algorithmes :**
 - ▶ Gower metric
 - ▶ Ecological Niche Factor Analysis (ENFA)
 - ▶ Maximum Entropy
 - ▶ Genetic Algorithm (GA)
 - ▶ Artificial Neural Network (ANN)
 - ▶ Algorithmes statistiques
 - ▶ Generalized Linear Model (GLM)
 - ▶ Generalized Additive Model (GAM)
 - ▶ Boosted Regression Tree (BRT)
 - ▶ Multivariate Adaptive Regression Spline (MARS)
 - ▶*Liste non exhaustive*.....

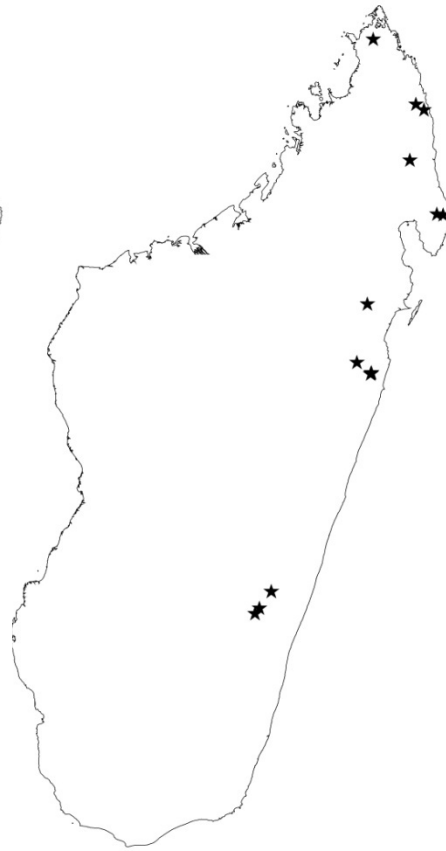
Les modèles EMPIRIQUES : Algorithmes

- ▶ Les algorithmes différent
 - ▶ De part les données de distribution requises
 - ▶ Présence et absence
 - ▶ Présence seulement
 - ▶ Présence et pseudo-absence
 - ▶ Présence et « fond » (background)

Alternative methods use species' distribution data differently:



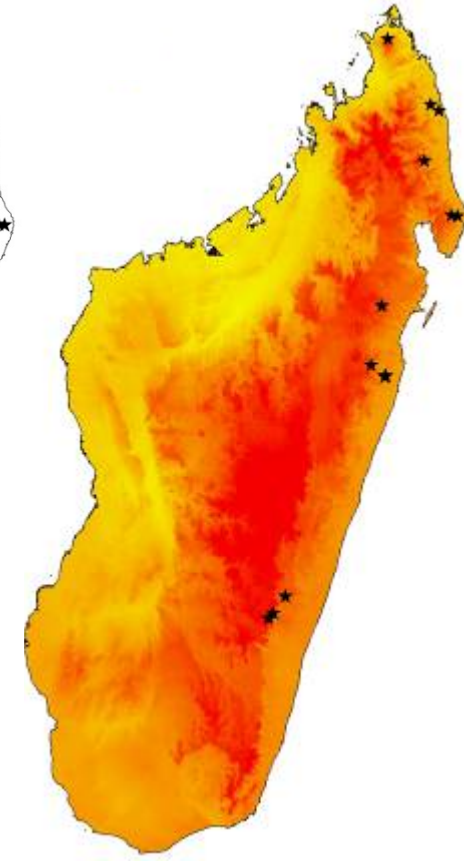
Presence/absence



Presence-only



Presence/pseudo-absence



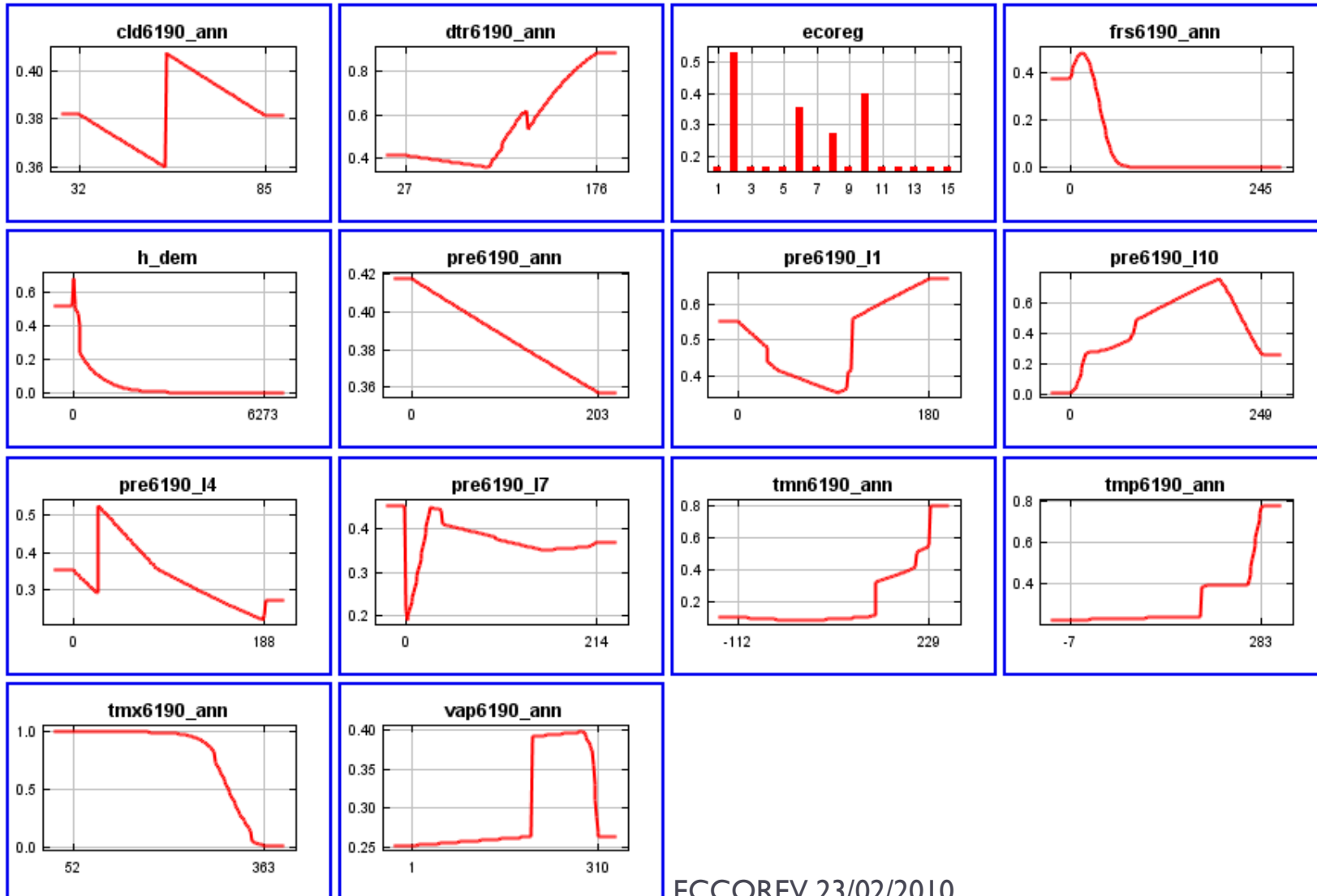
Presence/background absence



Les modèles EMPIRIQUES : Algorithmes

- ▶ On doit prendre en considération :
 - ▶ Les données de distribution requises par les modèles (et disponibles !)
 - ▶ Présence et absence
 - ▶ Présence seulement
 - ▶ Présence et pseudo-absence
 - ▶ Présence et « fond » (background)
 - ▶ La capacité des modèles à intégrer des données catégorielles
 - ▶ Le type de données en sortie
 - ▶ La capacité des modèles à déterminer l'influence des var. explicatives

Les modèles EMPIRIQUES : Algorithmes

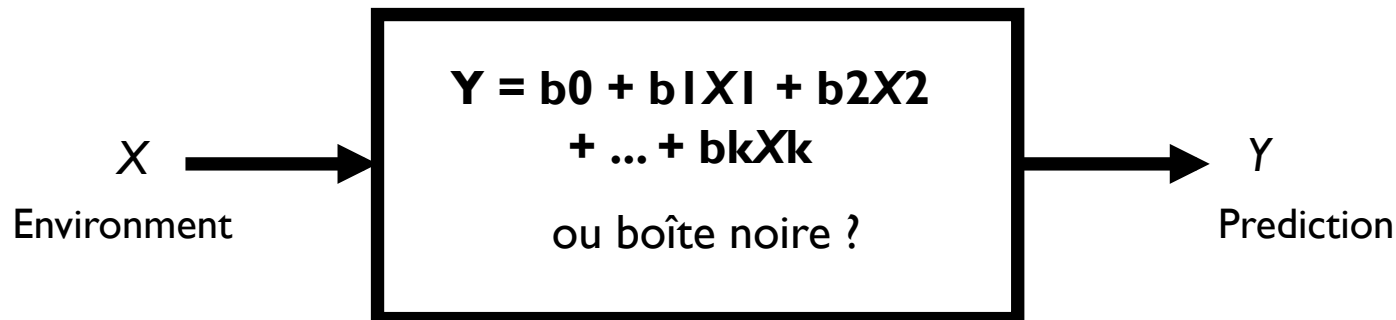


Les modèles EMPIRIQUES : Algorithmes

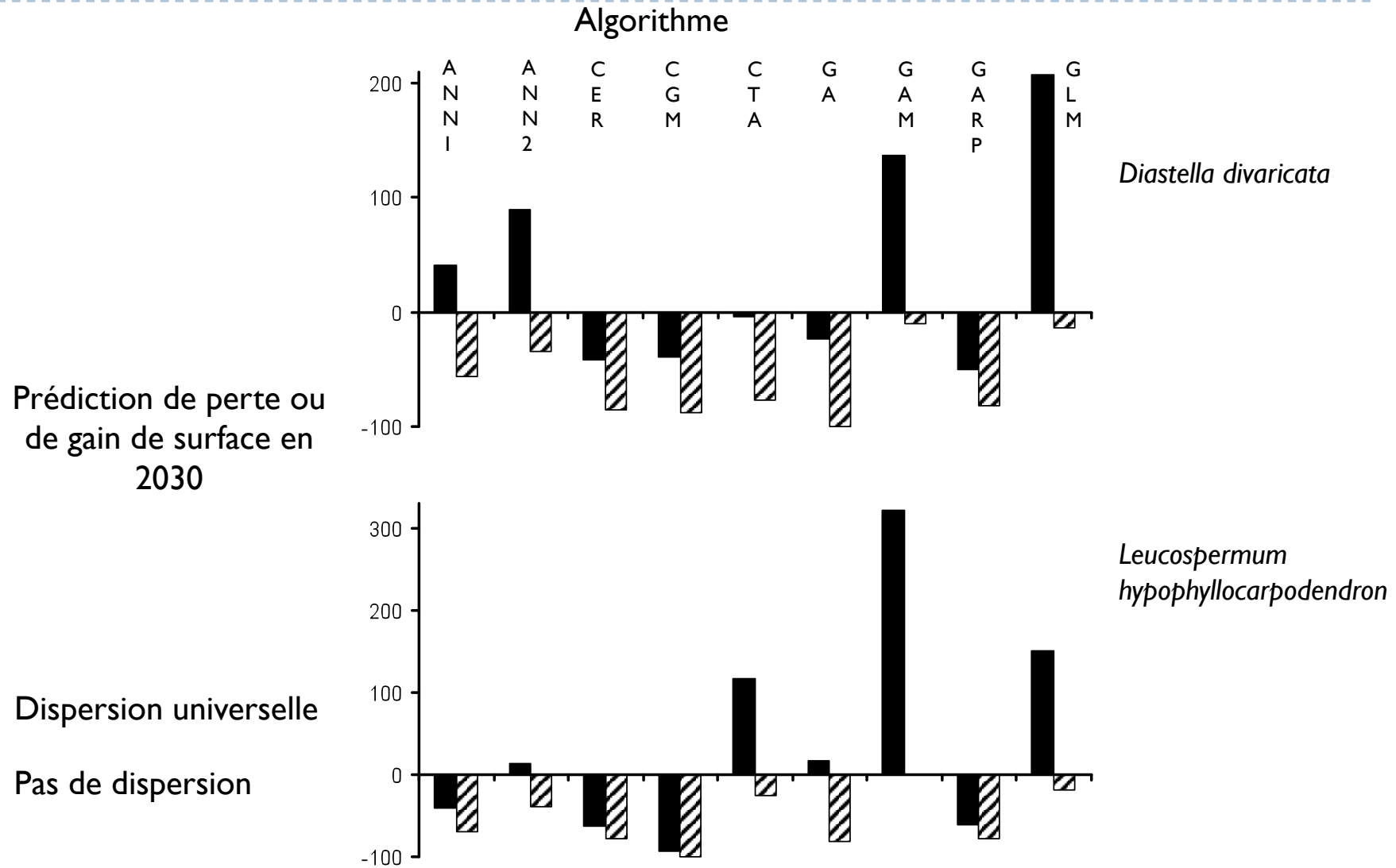
- ▶ **Modèle pour expliquer, comprendre les processus écologiques/physiologiques/biogéographiques...**
 - ▶ La prédiction la plus simple est préférée
 - ▶ Accès aux courbes de réponses
 - ▶ Accès à l'influence des variables

Les modèles EMPIRIQUES : Algorithmes

- ▶ **Modèle pour expliquer, comprendre les processus écologiques/physiologiques/biogéographiques...**
 - ▶ La prédiction la plus simple est préférée
 - ▶ Accès aux courbes de réponses
 - ▶ Accès à l'influence des variables
- ▶ **Modèle pour prédire....**
 - ▶ La meilleure prédiction possible
 - ▶ Ne permet pas la compréhension du système

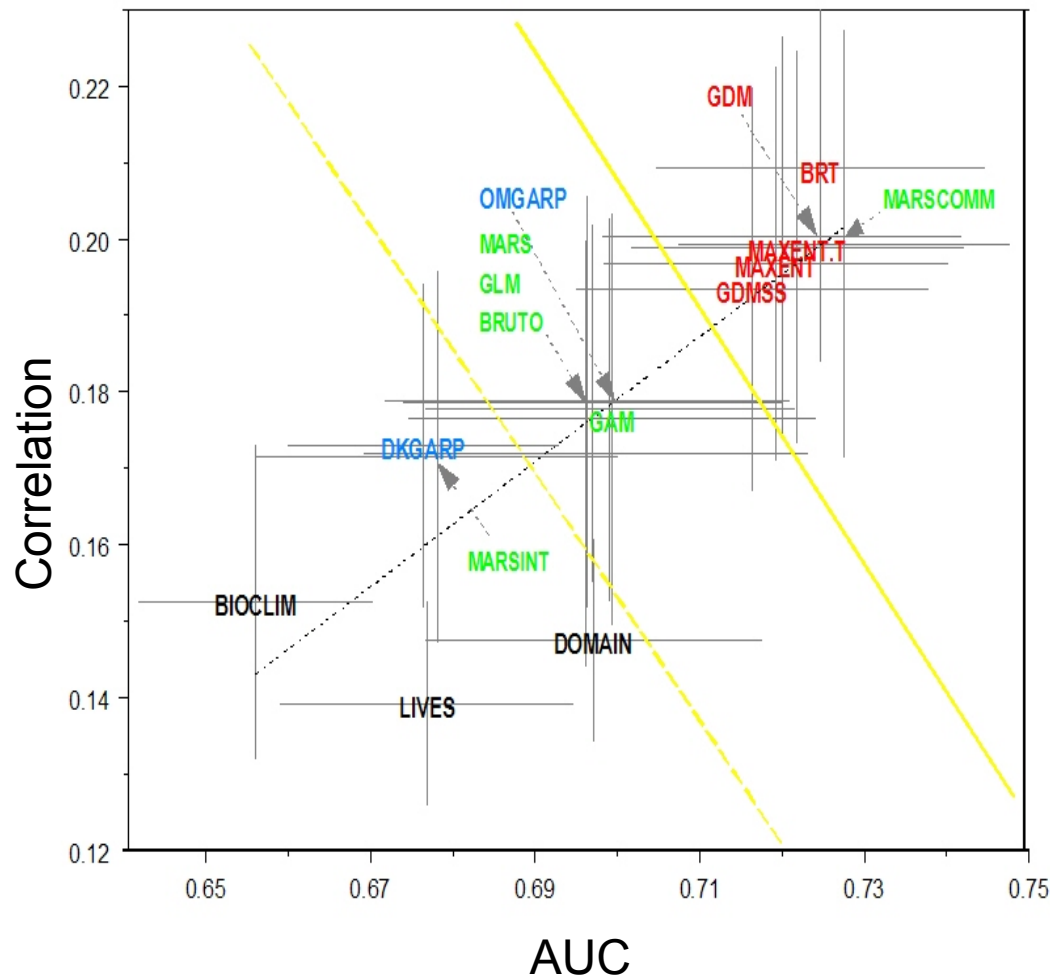


Les modèles EMPIRIQUES : Algorithmes



(crédits : Pearson et al., *Biogeography*, 2006) ECCOREY 23/03/2010

Les modèles EMPIRIQUES : Algorithmes



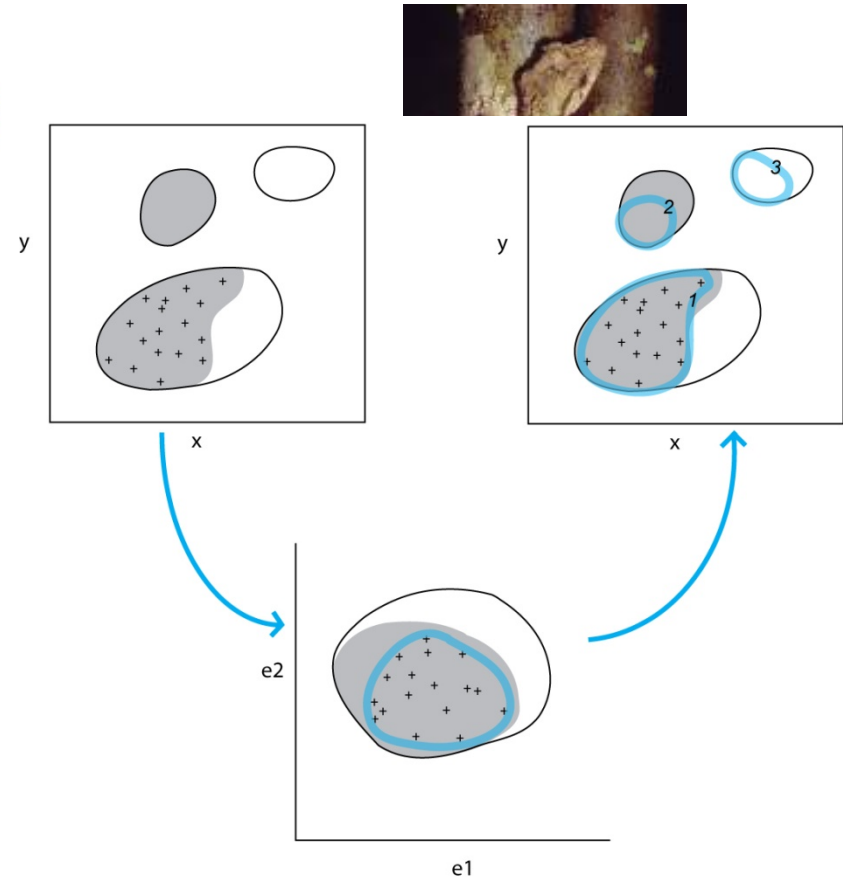
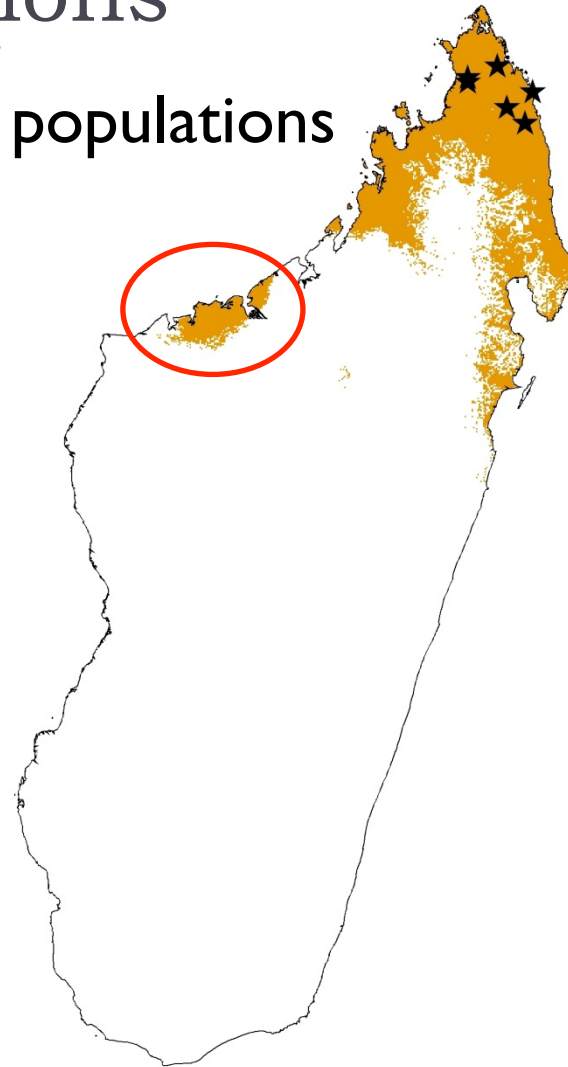
Différences liées :

- à la capacité à identifier relation inter variables
- à les approches d'interpolation
- à l'utilisation de Présence/Absence ou présence seule
- à l'utilisation de Pseudo-absence ou background
- au type paramétrique ou non paramétrique

(crédits : Elith *et al.*, *Ecography* 2006)

Applications

► Nouvelles populations

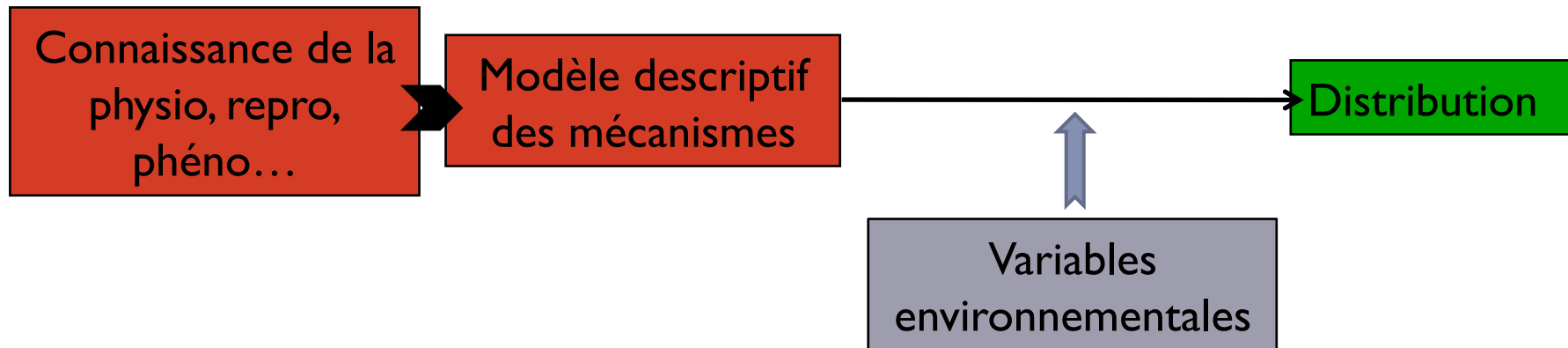


(Pearson et al. 2007 *J. Biogeog.*; Raxworthy et al. 2003 *Nature*)

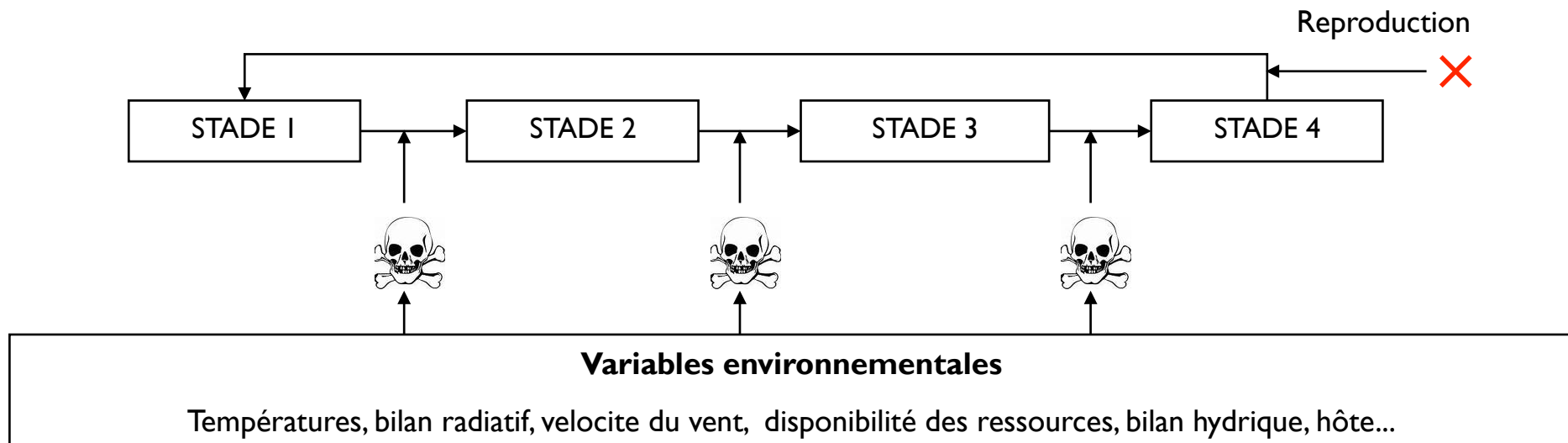
Les modèles EMPIRIQUES : Algorithmes

- ▶ En pratique...
 - ▶ Choix de l'algorithme :
 - ▶ Type de données de distribution disponible ?
 - Présence ou présence/absence ?
 - ▶ Type de données environnementale ?
 - Données catégorielles ou non ?
 - ▶ Utilité de connaître l'influence des variables ?
 - ▶ Application de plusieurs algorithmes
 - ▶ Définir une fourchette de prédictions
 - ▶ Quantifier l'incertitude liée au choix de l'algorithme

Les modèles MECANISTES : principes



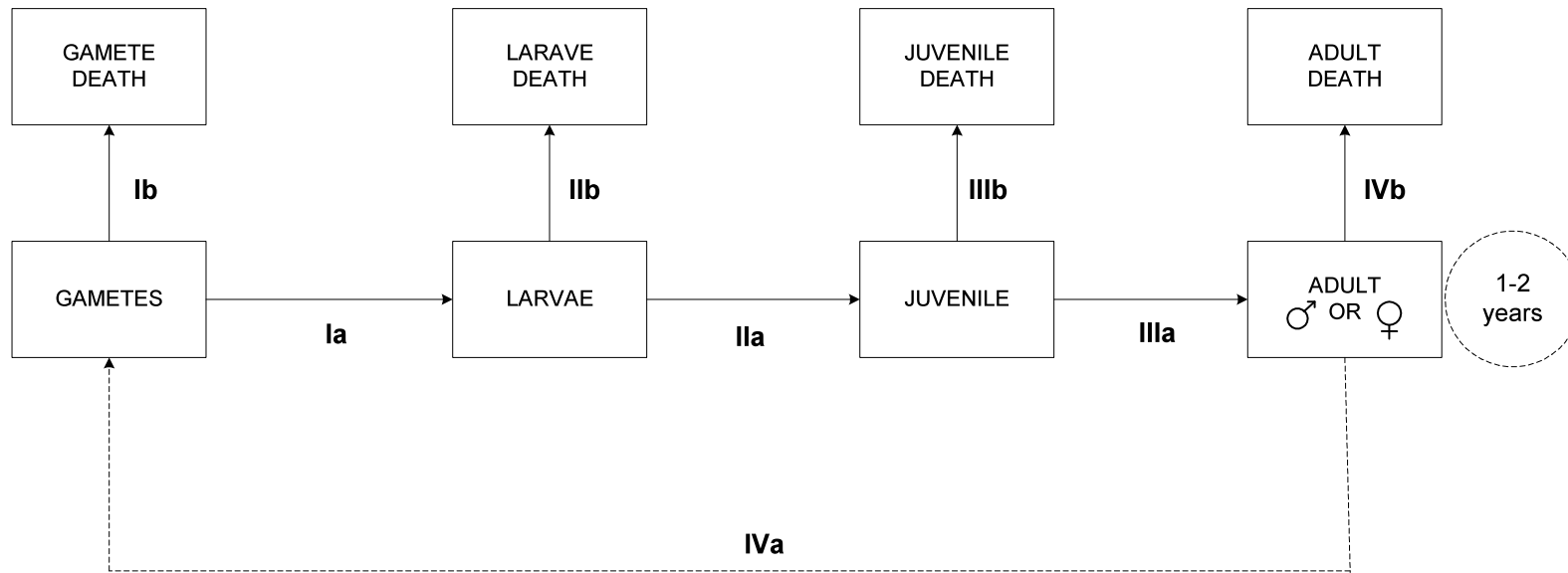
Les modèles MECANISTES : principes



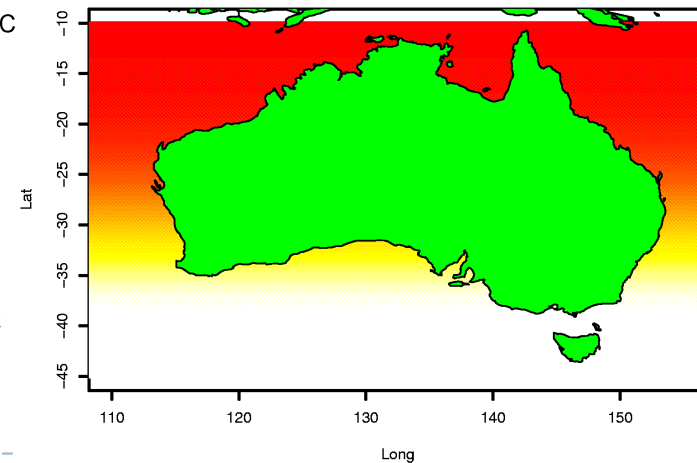
Physiologie (intervalles de tolérance, valence écologique)

Phénologie

Reproduction...

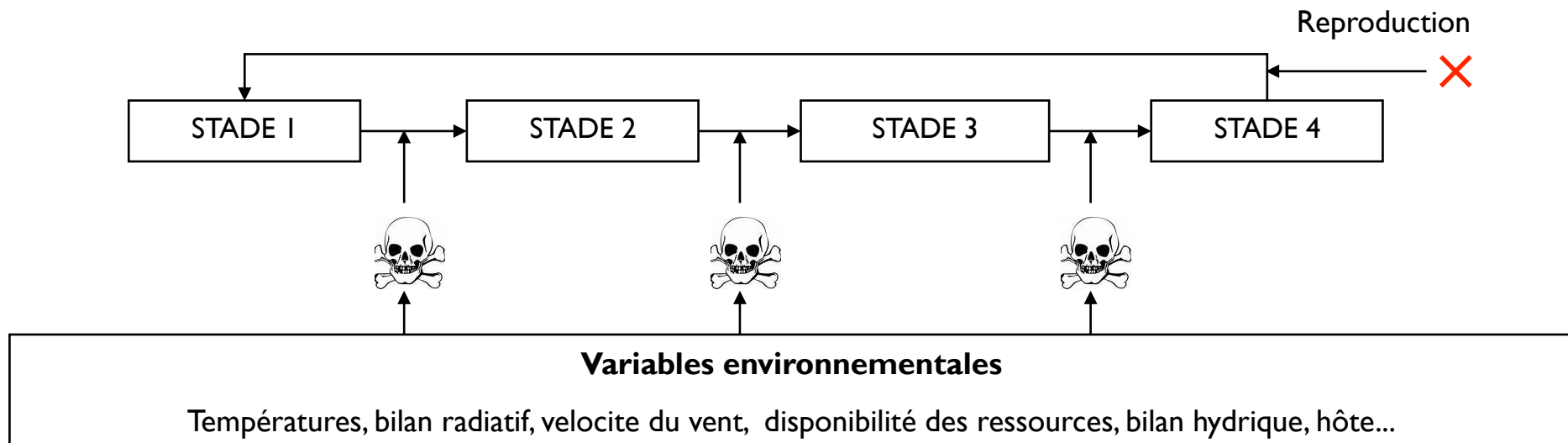


<u>Transition</u>	<u>Duration</u>	<u>Model</u>	<u>Temperature range</u>	<u>Temperature model</u>
Ia	Few hours	< 1 day	< 15 to >20 deg C	Min ~ Uniform [10, 13]; Max ~ Uniform [25, 30]
Ib	Few hours	<1 day	no data to > 30 deg C	Min > T > Max
IIa	2 to 4 days	Uniform [2, 4]	no data to >31 deg C	
IIb	Few hours	<1 day	T > 31 deg C	
IIIa	20 to 30 days	Uniform [20, 30]	no data	
IIIb	Few hours	<1 day	no data	
IVa	4 to 14 hours	< 1 day	no data to 32 deg C	
IVb	Few hours	< 1 day	5 > T > 45 deg C	



*Minimum inferred from adult data
 **Inferred from adult data
 ***Minimum inferred from gamete data

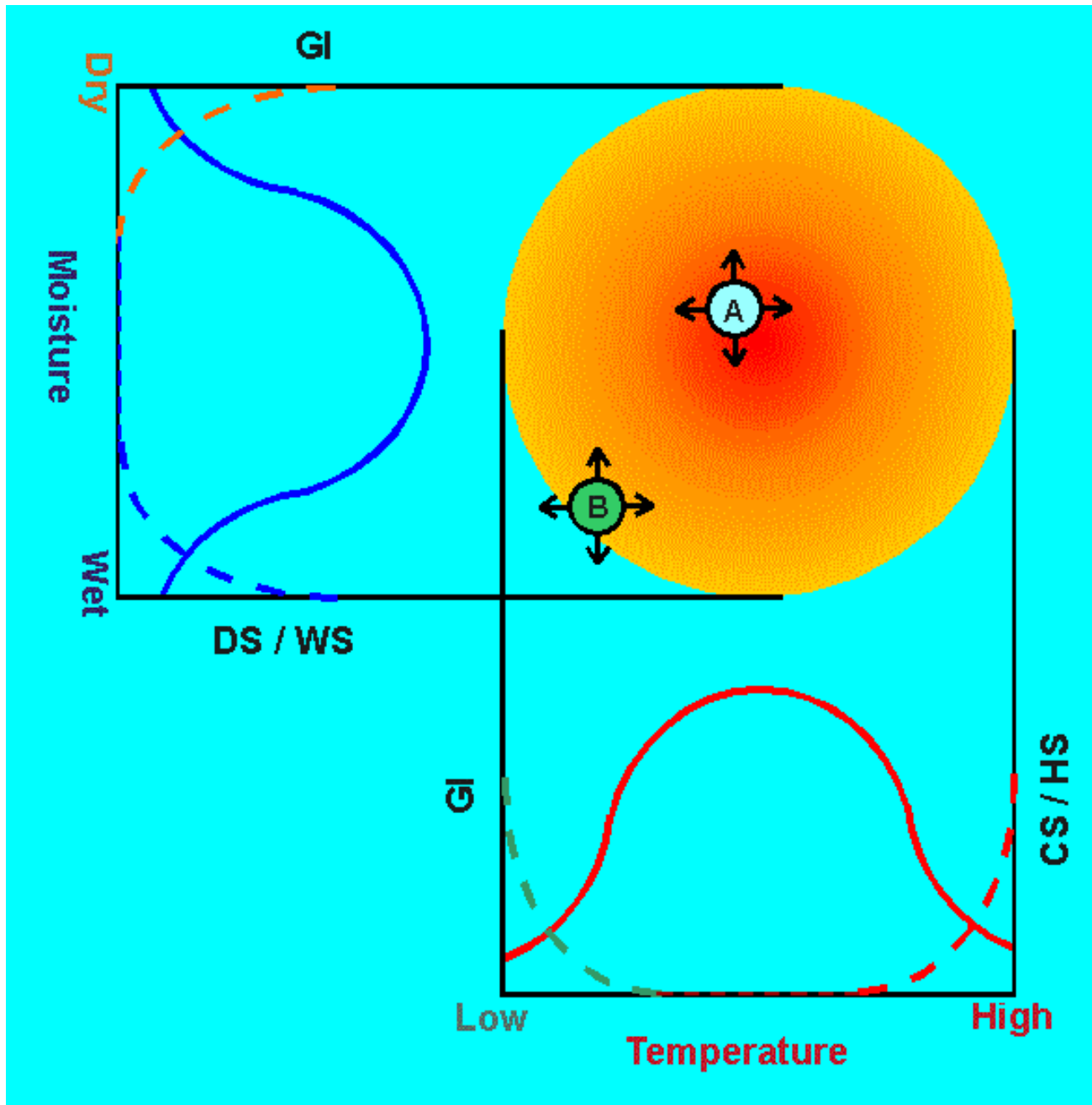
Les modèles MECANISTES : principes



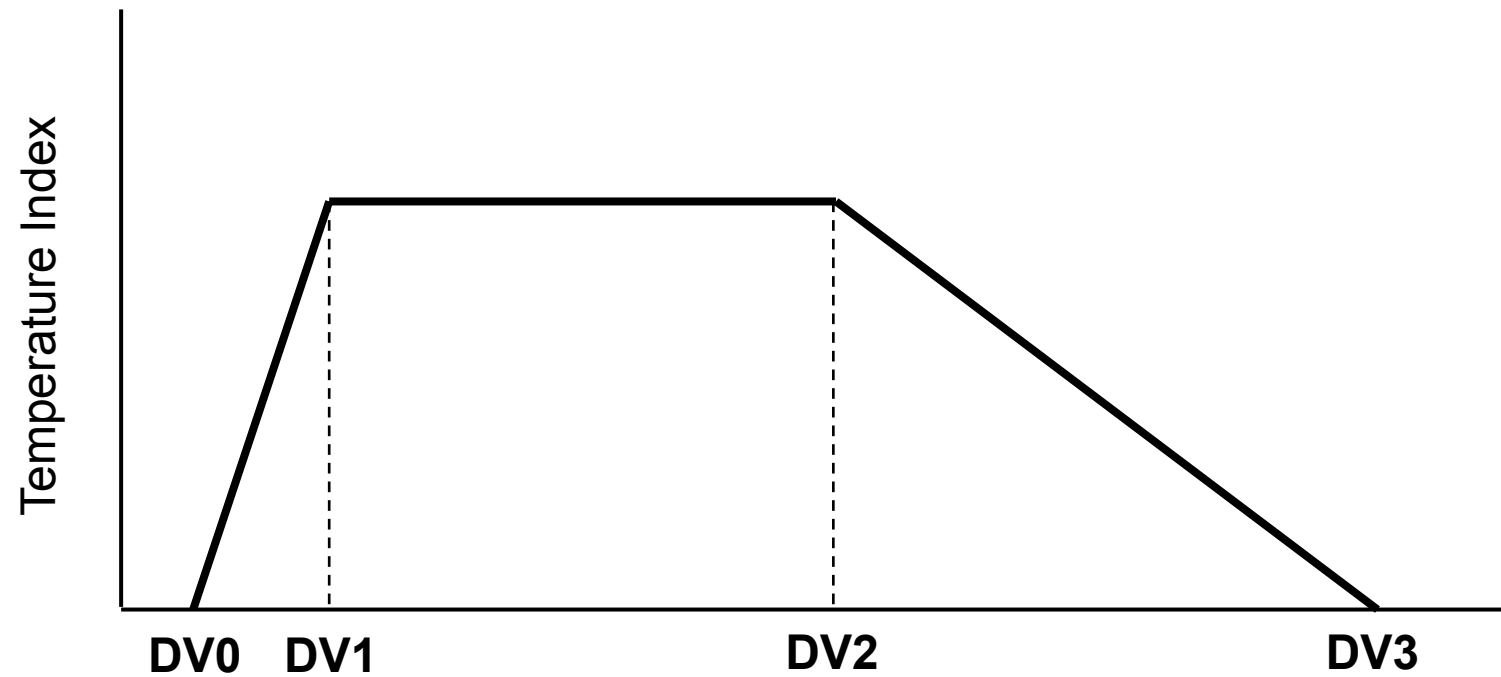
Physiologie (intervalles de tolérance, valence écologique)

Phénologie

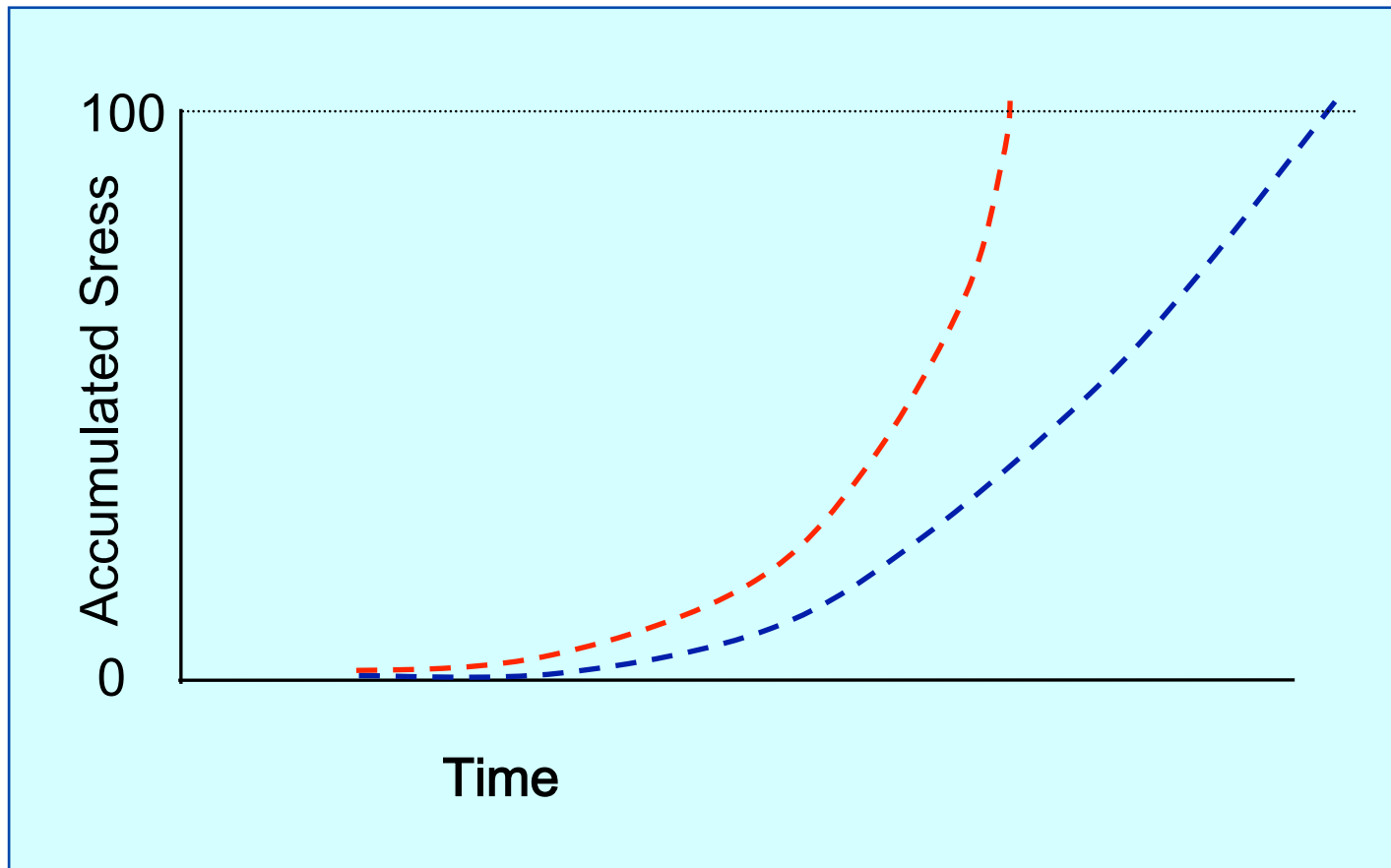
Reproduction...



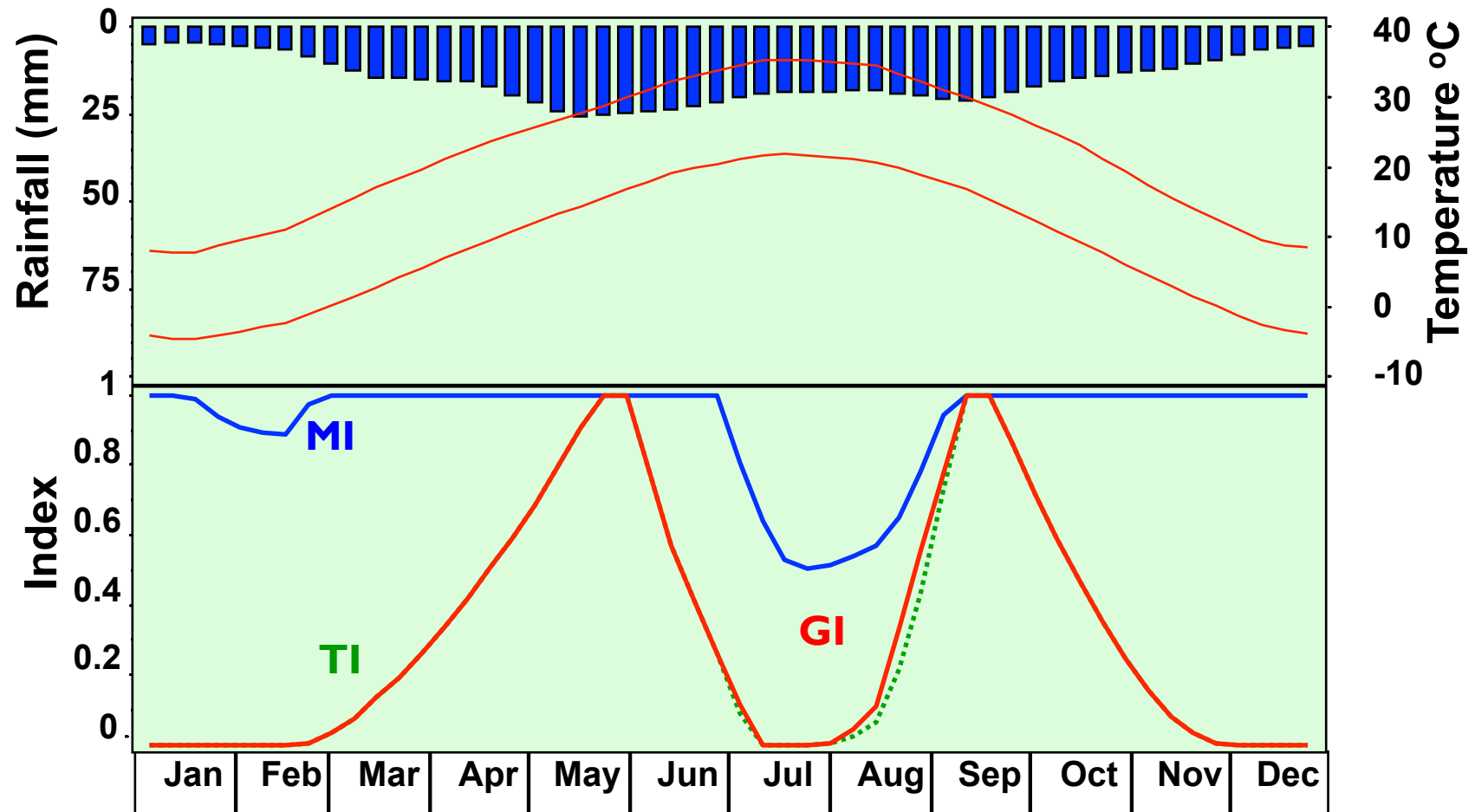
Les modèles MECANISTES : principes



Les modèles MECANISTES : principes



Les modèles MECANISTES : principes



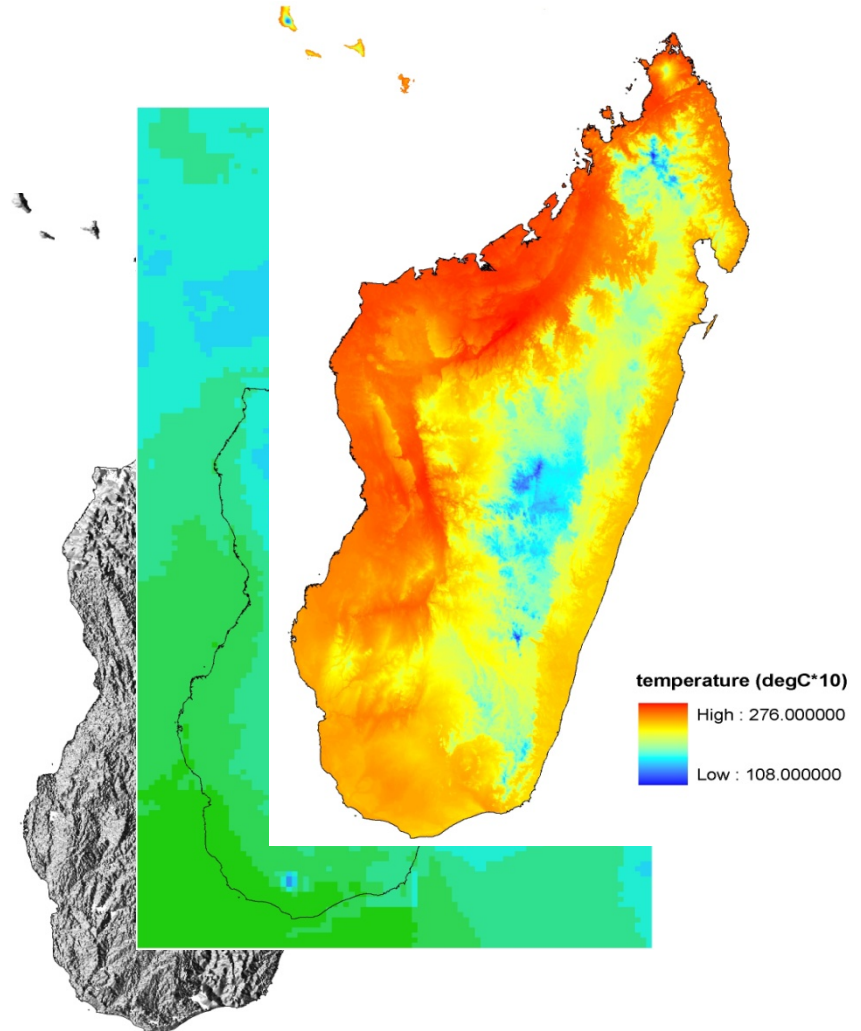
Medfly - Oklahoma

ECCOREV 23/02/2010

Les modèles MECANISTES : les données

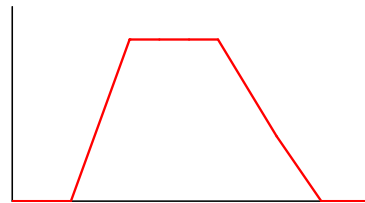
Tolérance de l'espèce :
expérimentations
→ Pur modèle mécaniste

Tolérance de l'espèce : couplage
expérimentations /
distribution
→ Modèle orienté
mécaniste



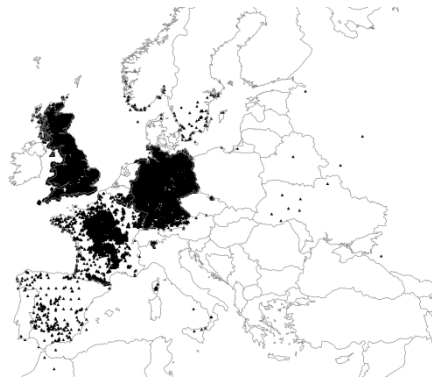
Données environnementales

Physiologie

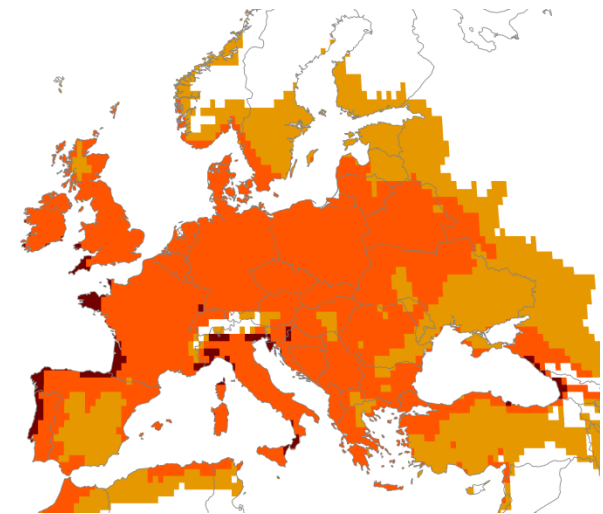


Cytisus scoparius

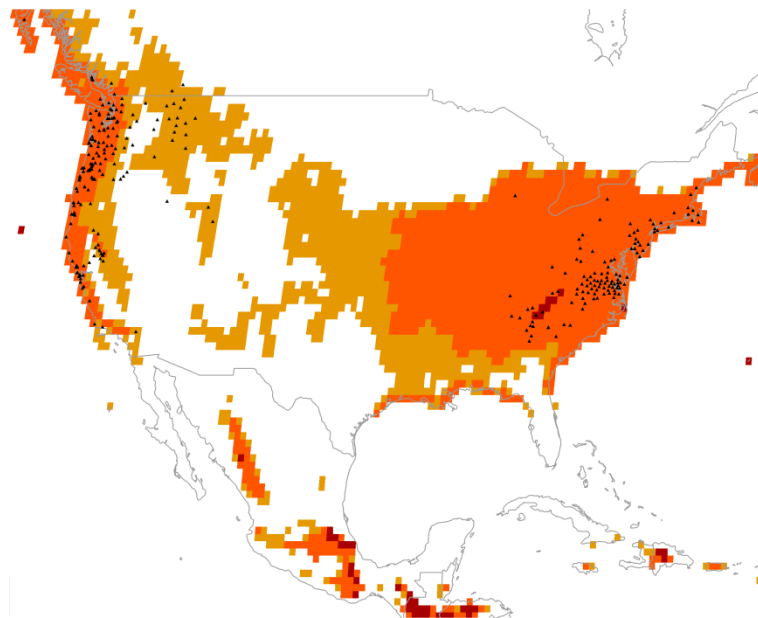
Distribution



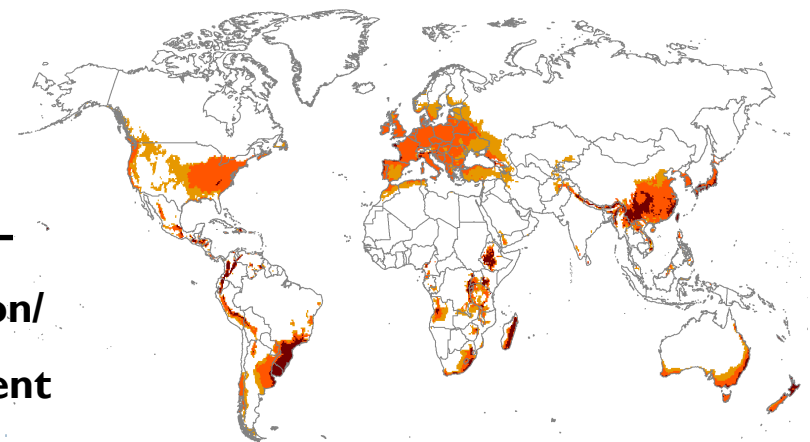
Paramétrisation



Projection



Evaluation/
Ajustement



ECCOREV 23/02/2010

Conclusions

	Mécaniste	Empirique
Besoin en données	Fort	Faible
Temps de de calibration et de traitement	Long	Court
Nombre d'espèces traitées	Faible	Elevé
Puissance explicative	Forte	Faible
Projection future	Possible	Erronée
Validation	Indépendante	Croisée
Approximation de la niche...	Fondamentale	Réalisée



Merci de votre attention

Agathe Leriche

agathe.leriche@univ-cezanne.fr

Europôle méditerranéen de l'Arbois, Aix-en-Provence

Tel. : 04.42.90.84.55

ECCOREV 23/02/2010