



Les modèles d'IA générative peuvent-ils simuler des trajectoires d'oiseaux marins réalistes ?

Ecostat 2026 - Grenoble - 27/05/2026

Thomas BARBEDETTE-GERARD

Encadré par Sophie Lanco et Benoît Delahaye



Pourquoi étudier les oiseaux marins ?

De nombreuses espèces sont menacées (~31%)

- 56% des espèces sont en déclin
- **Nombreuses menaces:** changement climatique, prises accidentelles, surpêche, pollution, ...



Prises accidentelles: une des nombreuses menaces



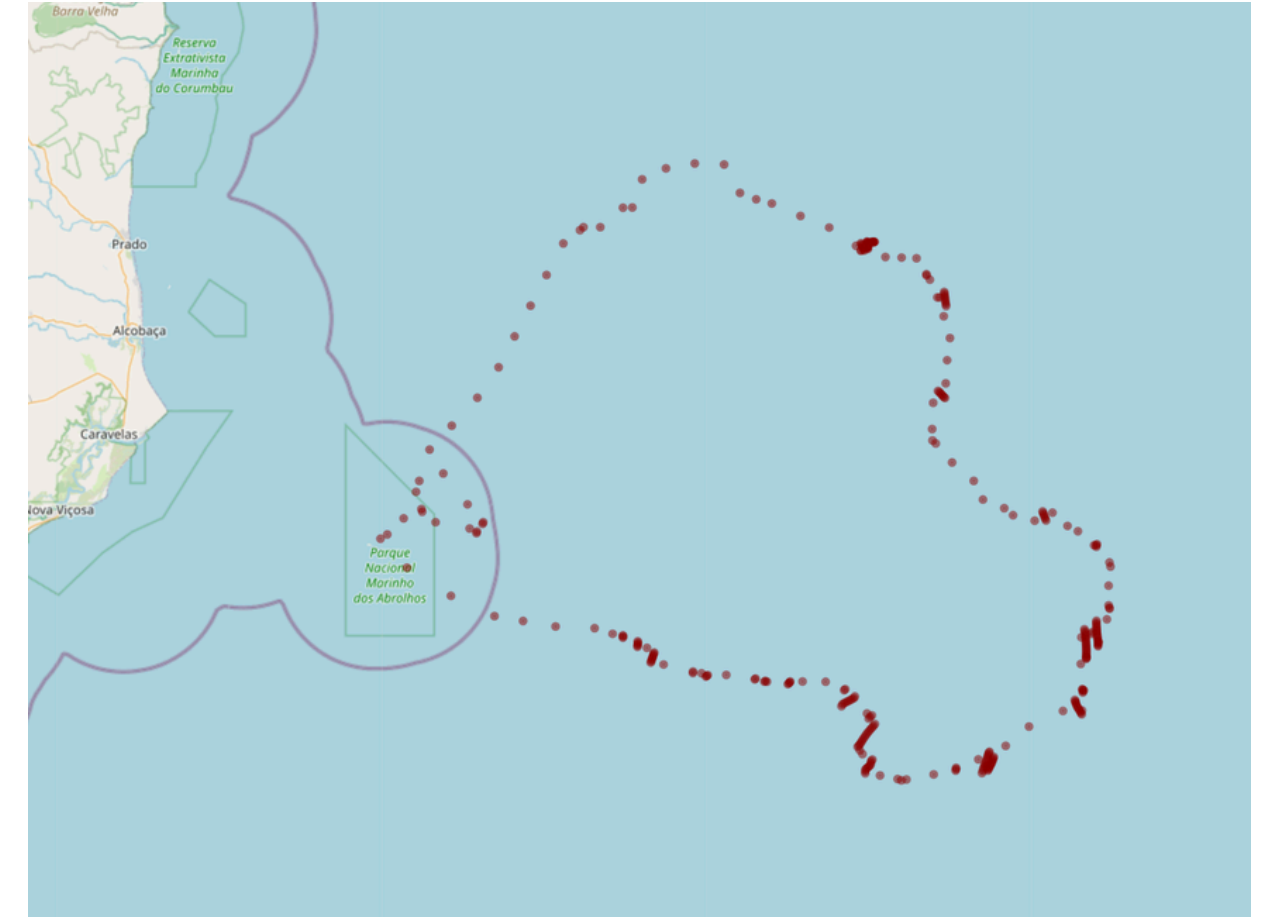
Macareux moine avec poissons

Des sentinelles des écosystèmes marins

- Longue durée de vie et distribution mondiale
- Plus accessibles que les espèces sous-marines
- Dépendent des ressources marines : enregistrent des changements dans les écosystèmes

Pourquoi étudier les oiseaux marins ?

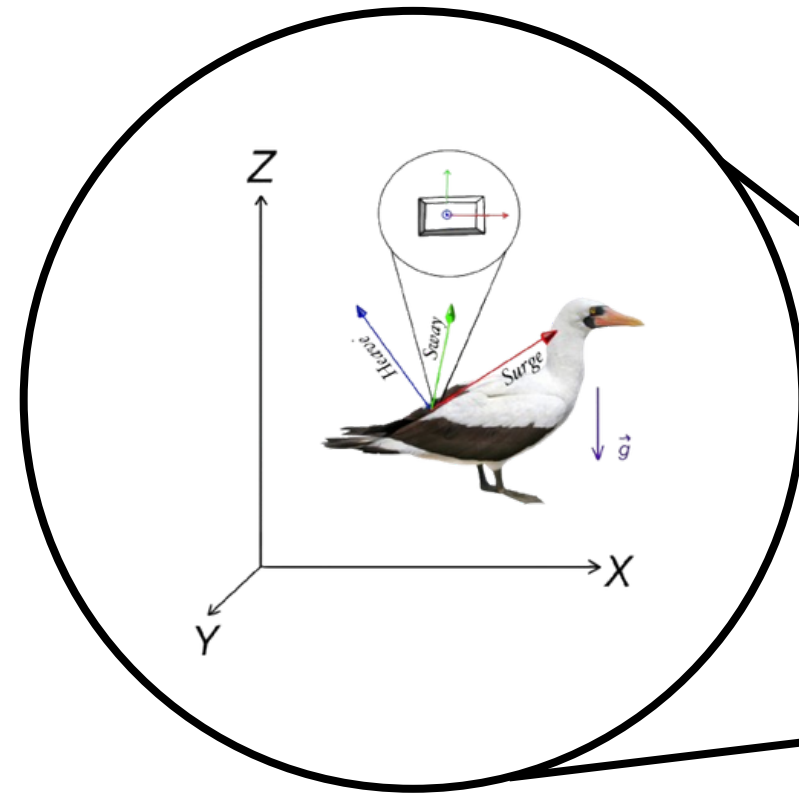
- Reproduction à terre.
Retour régulier obligatoire pour nourrir les poussins
- La période de reproductrice est critique pour la viabilité des colonies



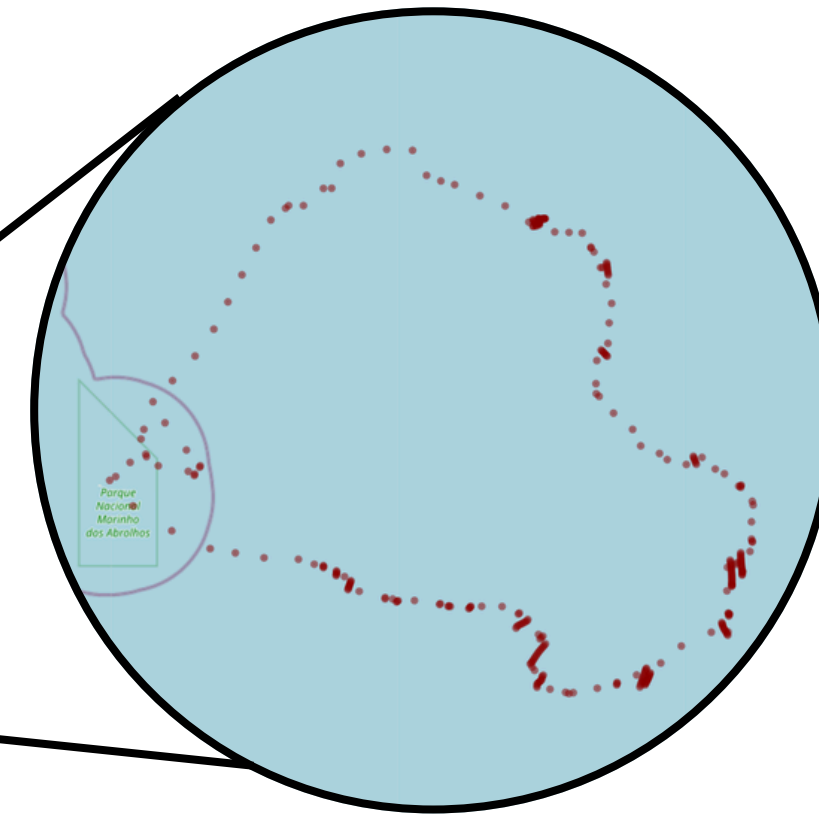
L'étude des déplacements des oiseaux marins est cruciale pour :

1. Évaluer la **viabilité des colonies** dans un contexte de changement climatique et d'activités humaines.
2. Soutenir une **planification marine** plus éclairée.

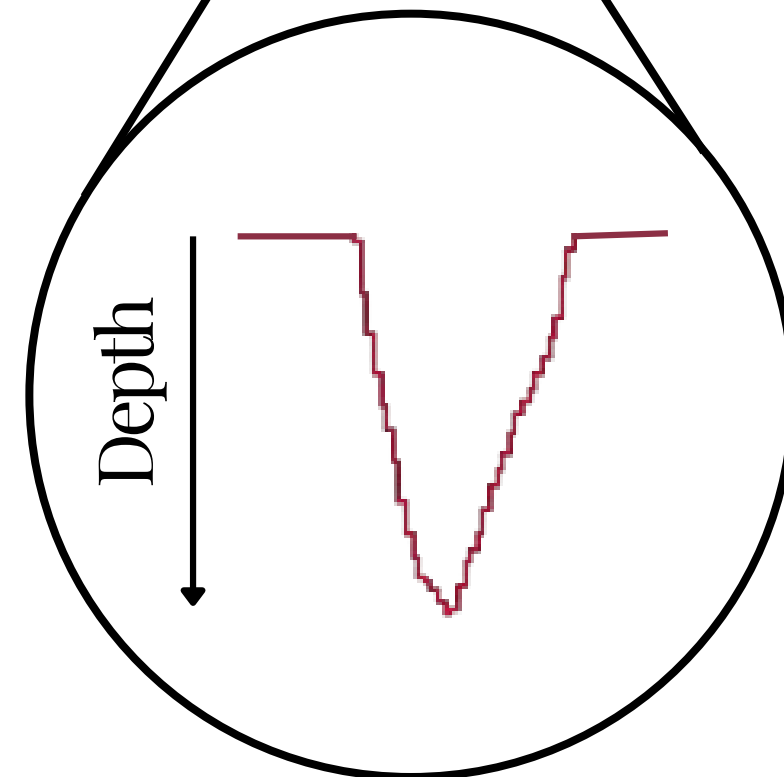
Biologging



Accéléromètre



GPS



Capteur de pression

Biologging d'oiseaux tropicaux



Sites étudiés dans les eaux tropicales Sud Américaines



Phaéton à bec rouge
Phaethon aethereus



Fou à pieds rouges
Sula sula



Fou varié
Sula sula



Fou masqué
Sula dactylatra



Fou brun
Sula leucogaster



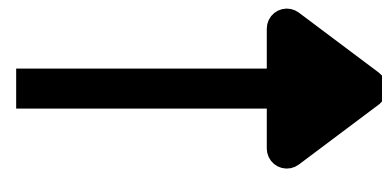
Cormorant de Bougainville
Leucocarbo bougainvillorum

Des individus aux populations

On instrumente quelques individus dans des colonies avec jusqu'à des centaines de milliers d'individus



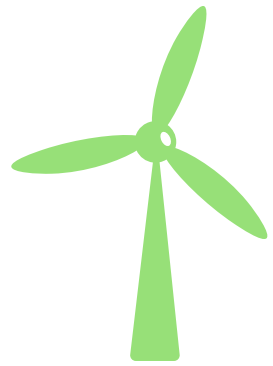
Fou varié, *Sula variegata*



Centaines de milliers de fous, cormorans et pélicans (Pérou)

Est-ce que les conclusions pour ces individus sont toujours valables pour des colonies/populations entières ?

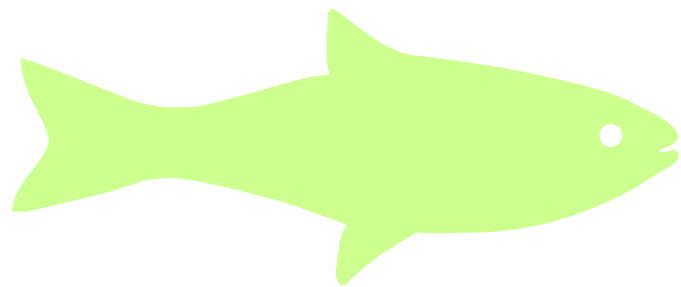
Variables environnementales



Activités
anthropiques

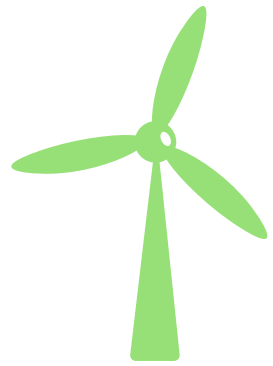


Conditions
climatiques



Distribution des proies

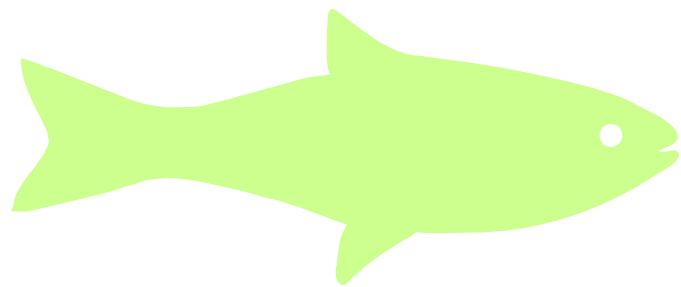
Variables environnementales



Activités
anthropiques



Conditions
climatiques



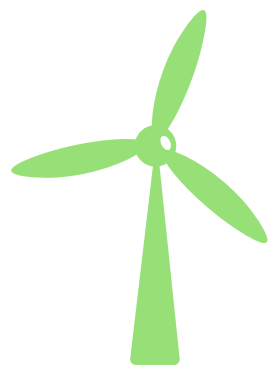
Distribution des proies

Variabilité inter-individuelle



- Espèce
- Sexe
- Age
- Personnalité
- ...

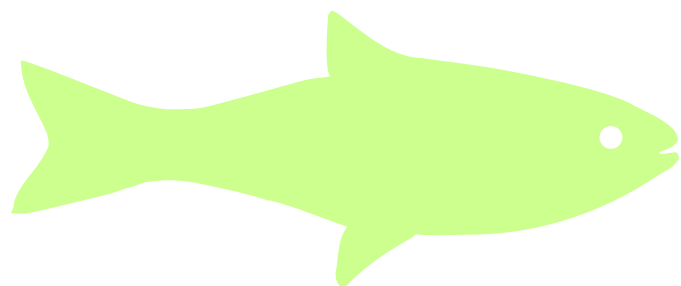
Variables environnementales



Activités
anthropiques



Conditions
climatiques



Distribution des proies

Variabilité inter-individuelle

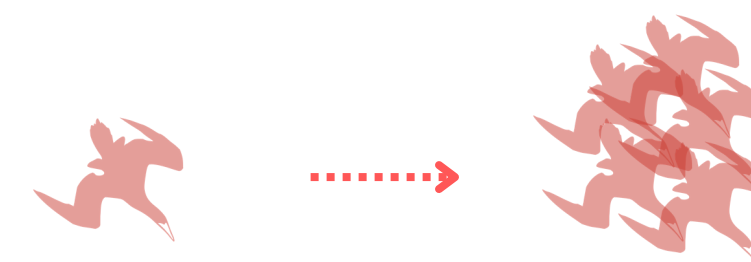


- Espèce
- Sexe
- Age
- Personnalité
- ...

Interactions



Transfert
d'information

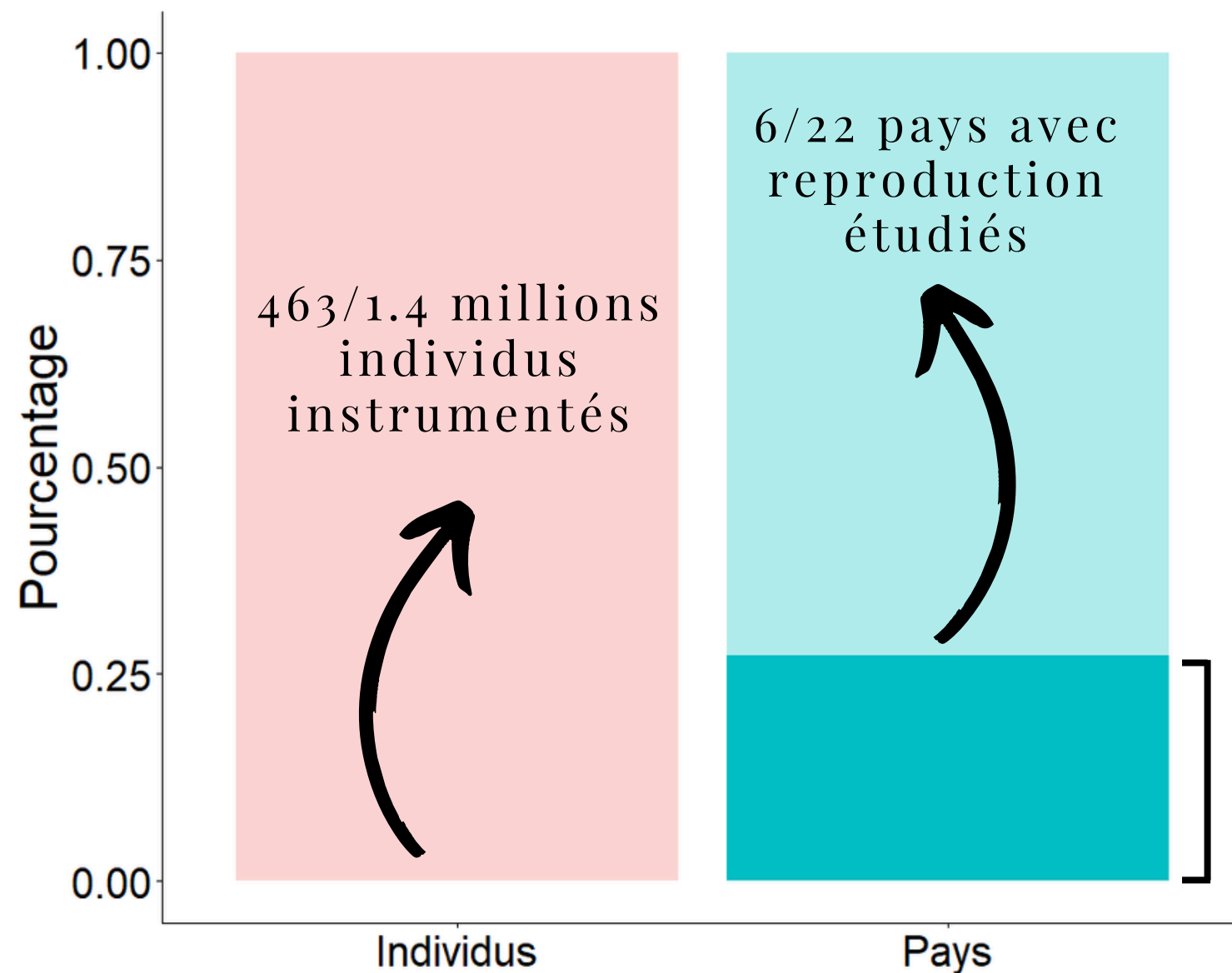


Effet de la taille de
la colonie

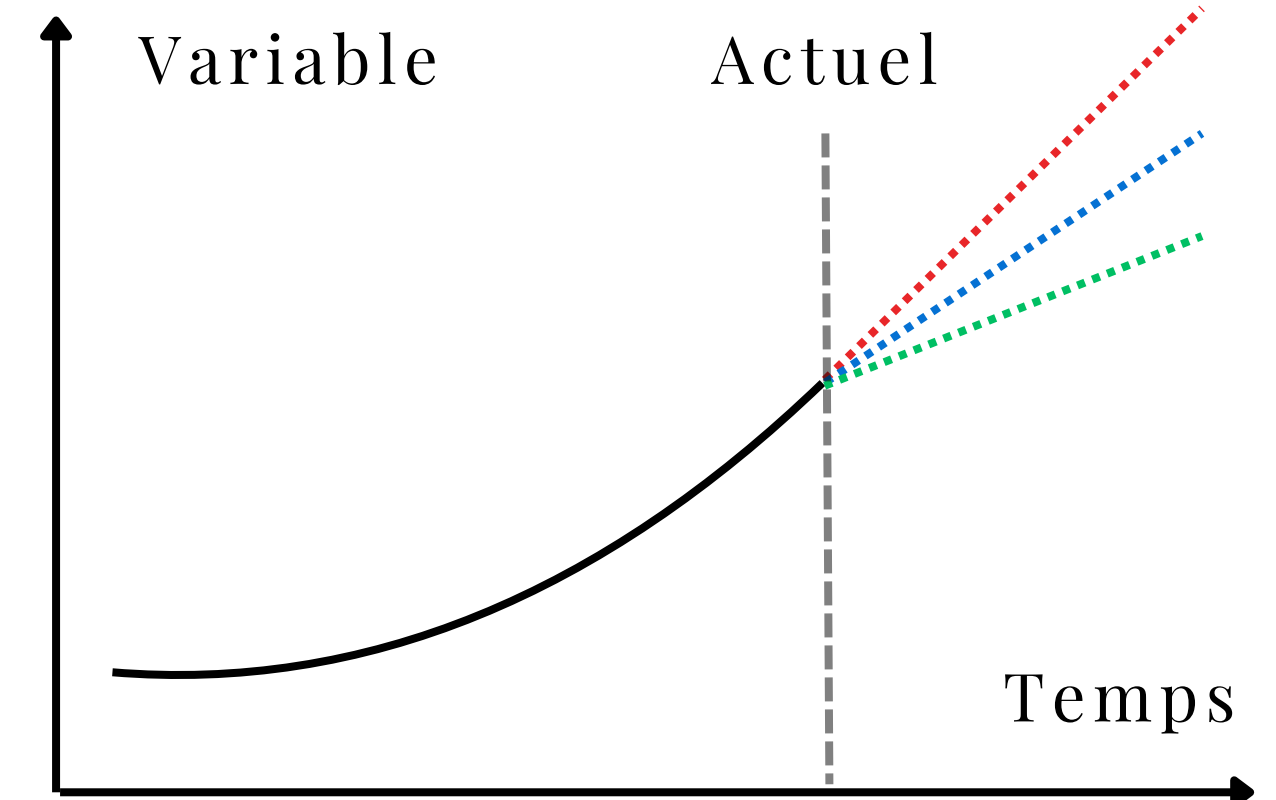
Pourquoi simuler des trajectoires d'oiseaux marins ?

Exploration de la **plasticité** du mouvement en fonction des **conditions**

Sous échantillonnage spatial



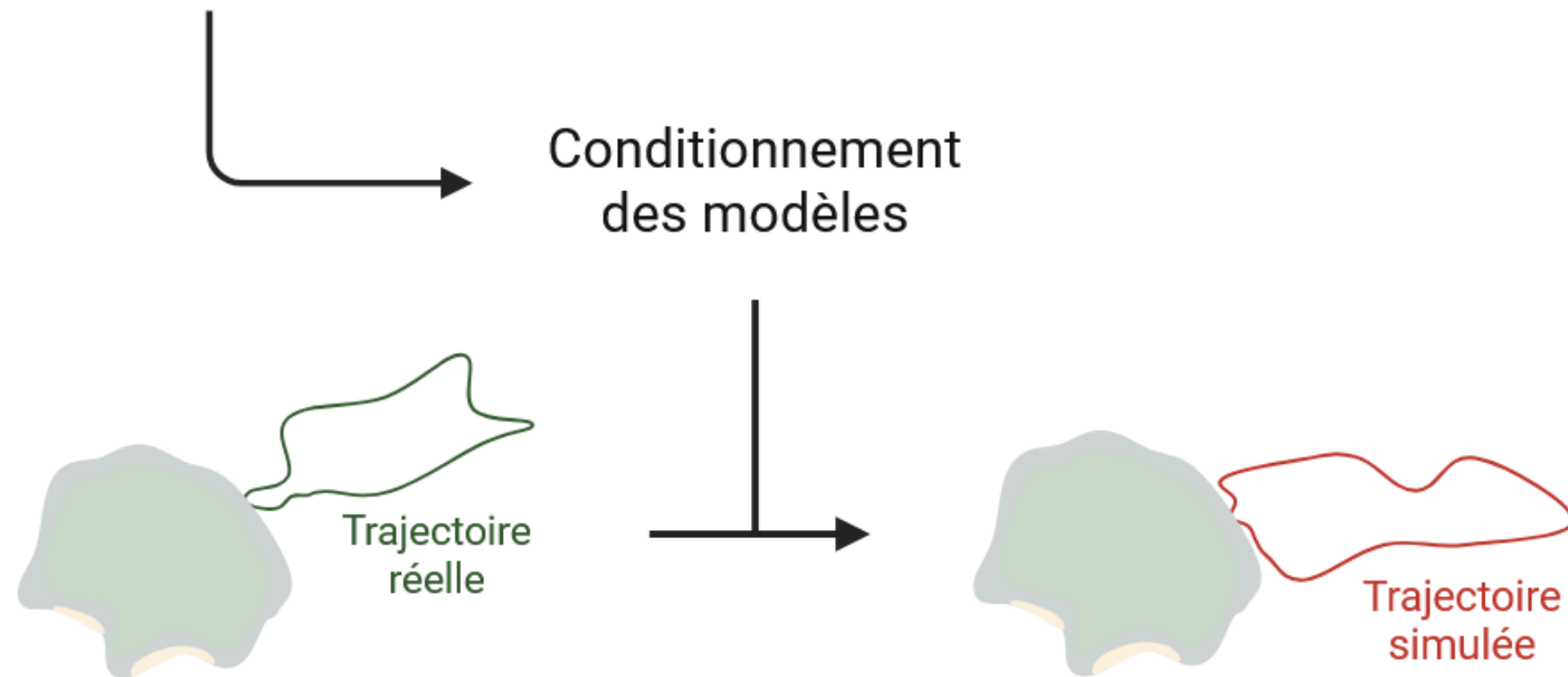
Sous échantillonnage temporel



Inférence mouvement sous différents scénarios

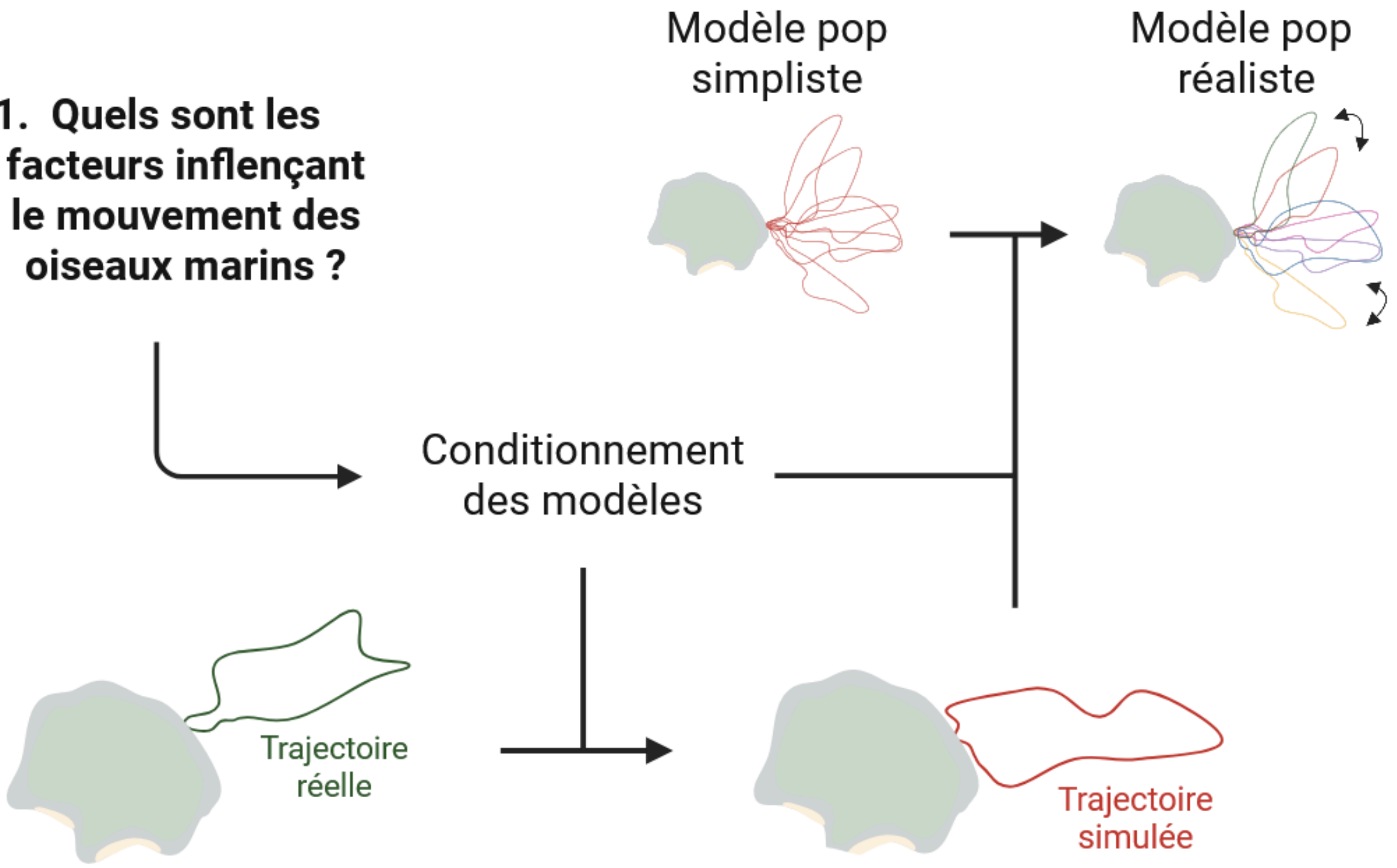
Conditionnement des modèles

1. Quels sont les facteurs influençant le mouvement des oiseaux marins ?

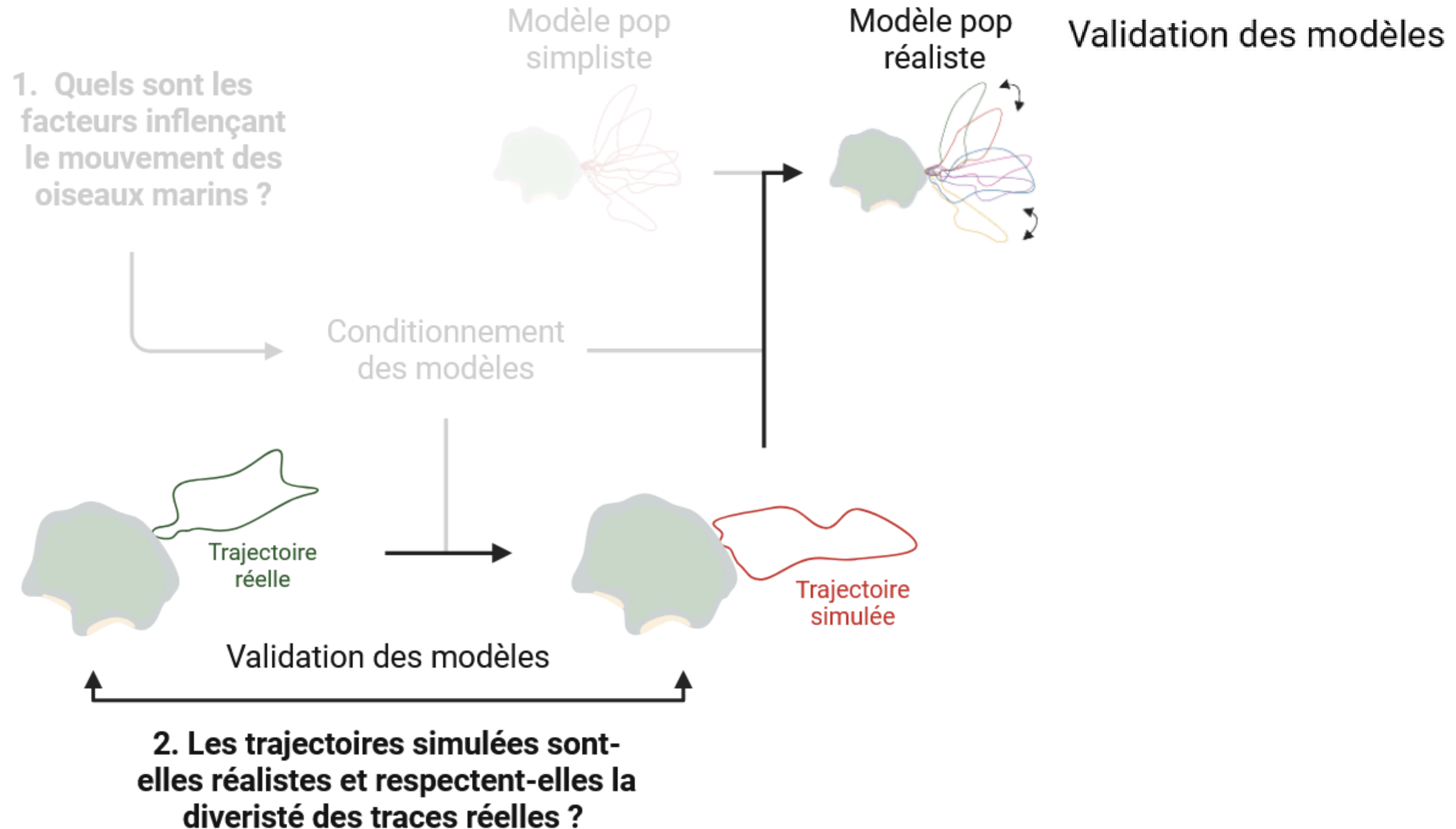


Conditionnement des modèles

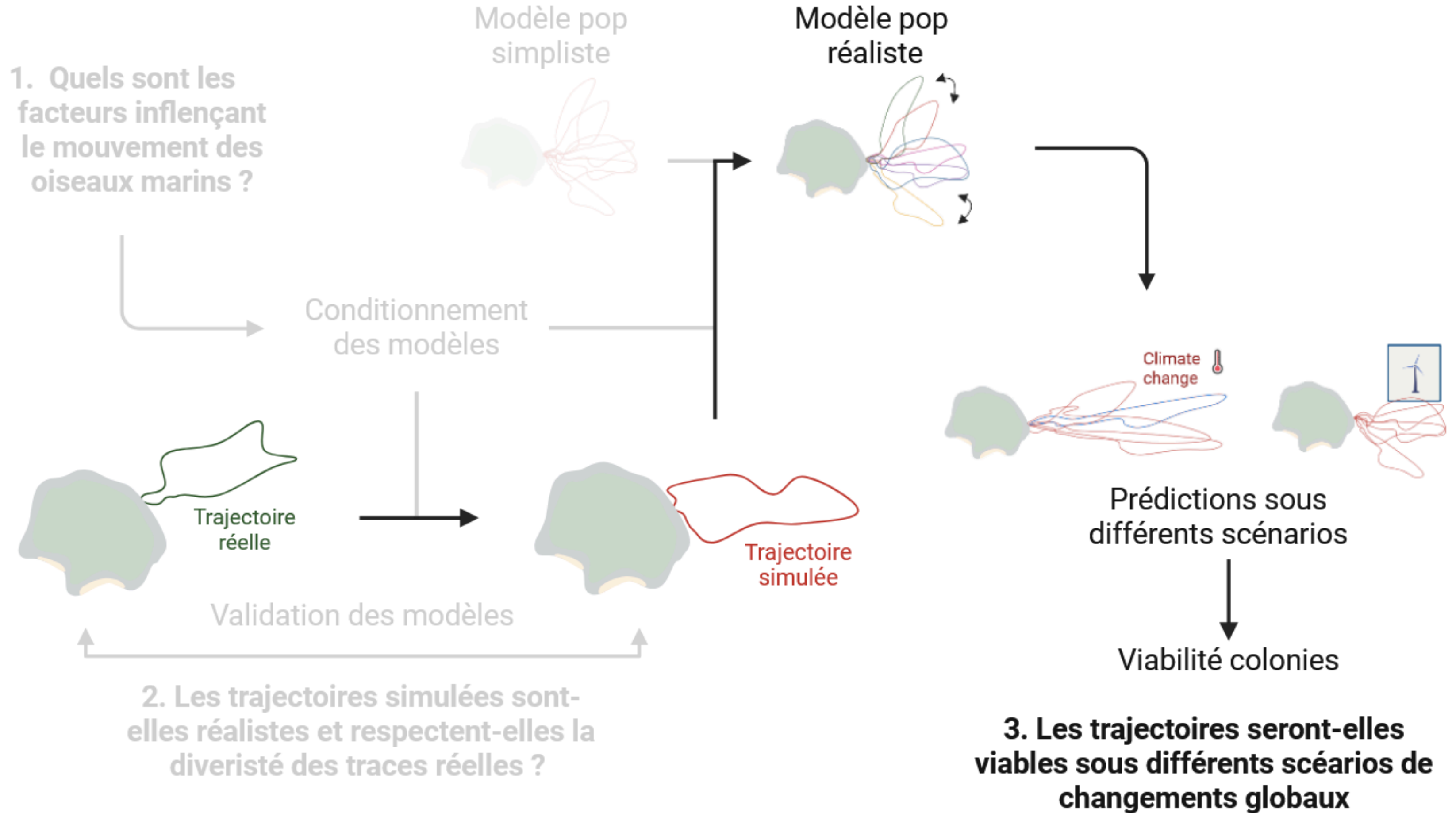
1. Quels sont les facteurs influençant le mouvement des oiseaux marins ?



Validation des modèles

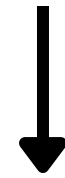


Simulations sous différents scénarios

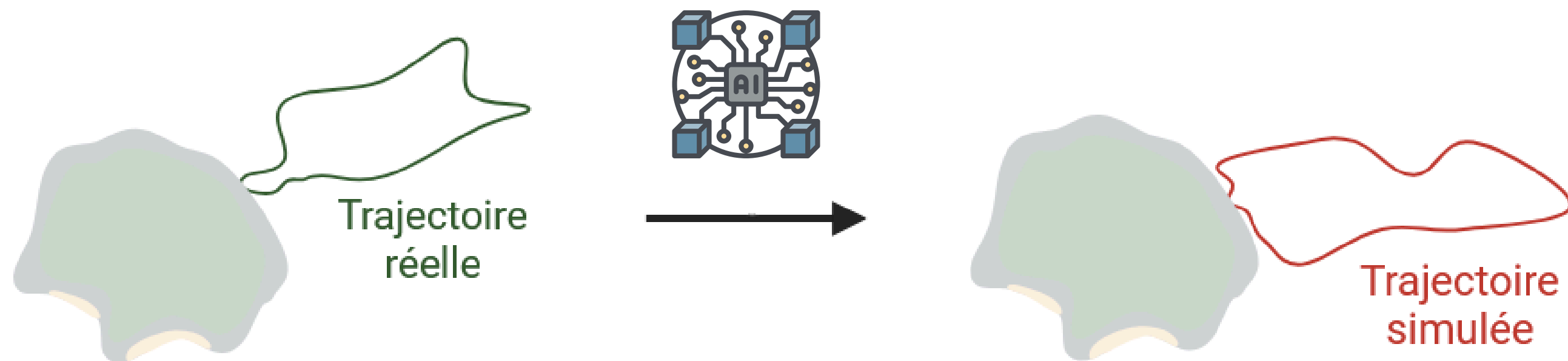


Simuler des trajectoires d'oiseaux marins

Les méthodes de l'état de l'art pour l'analyse du mouvement animal (HMM, SSF,...) simulent des trajectoires peu réalistes



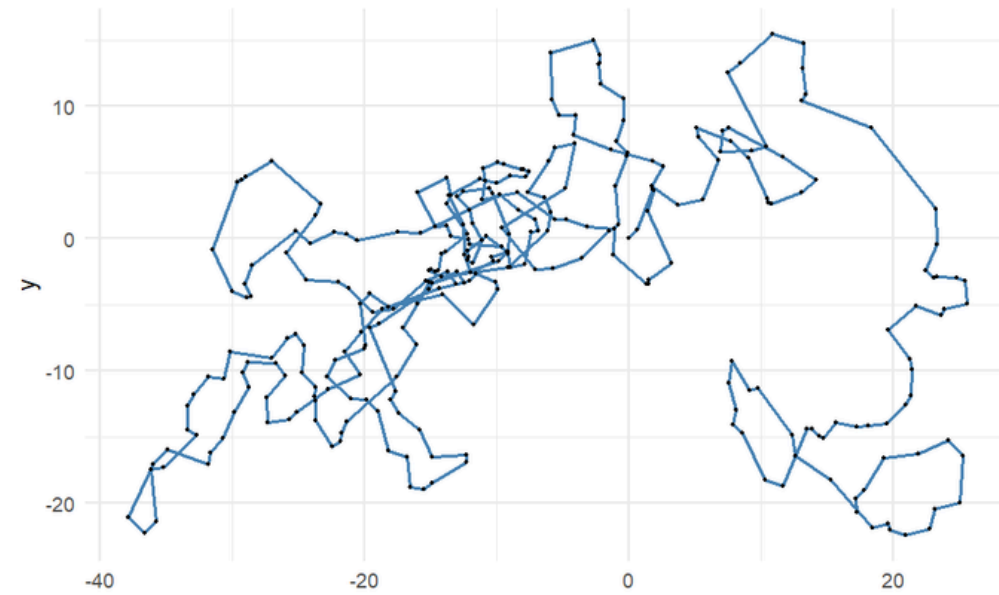
Utilisation de méthodes d'IA générative pour simuler des trajectoires d'oiseaux marins



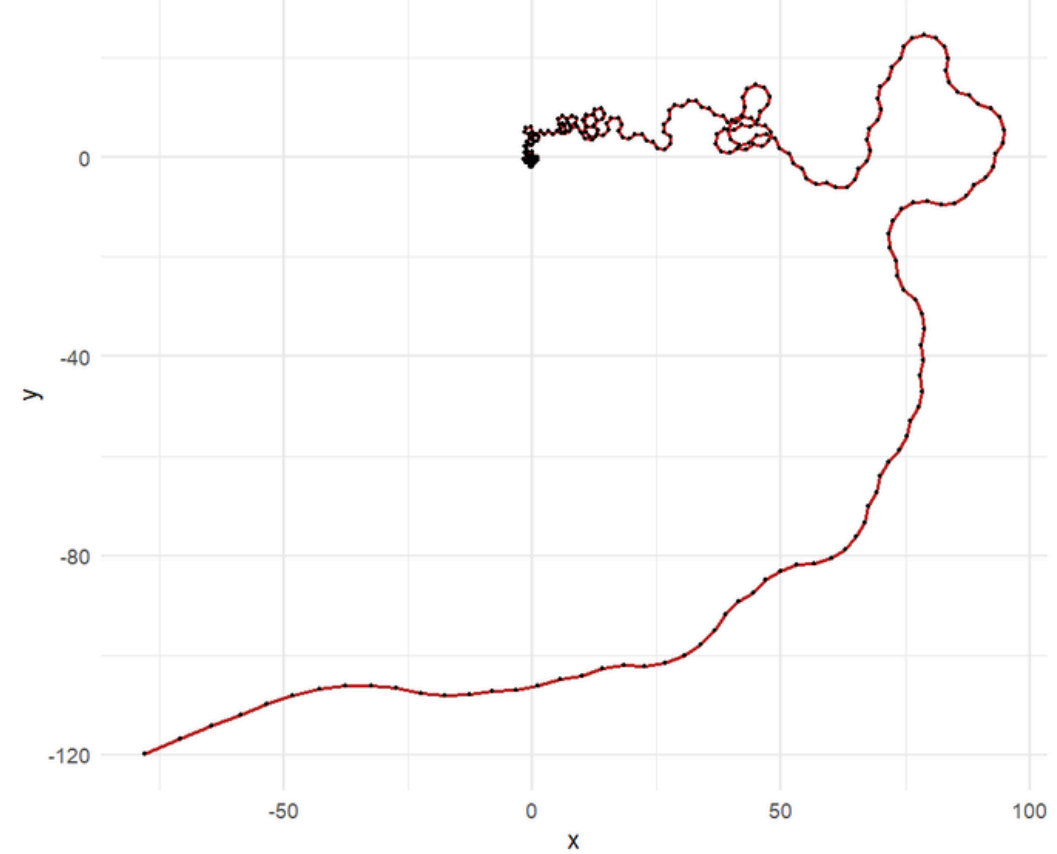
Comparaison simple de métriques

Mêmes distributions de pas et d'angle mais association entre variables différentes

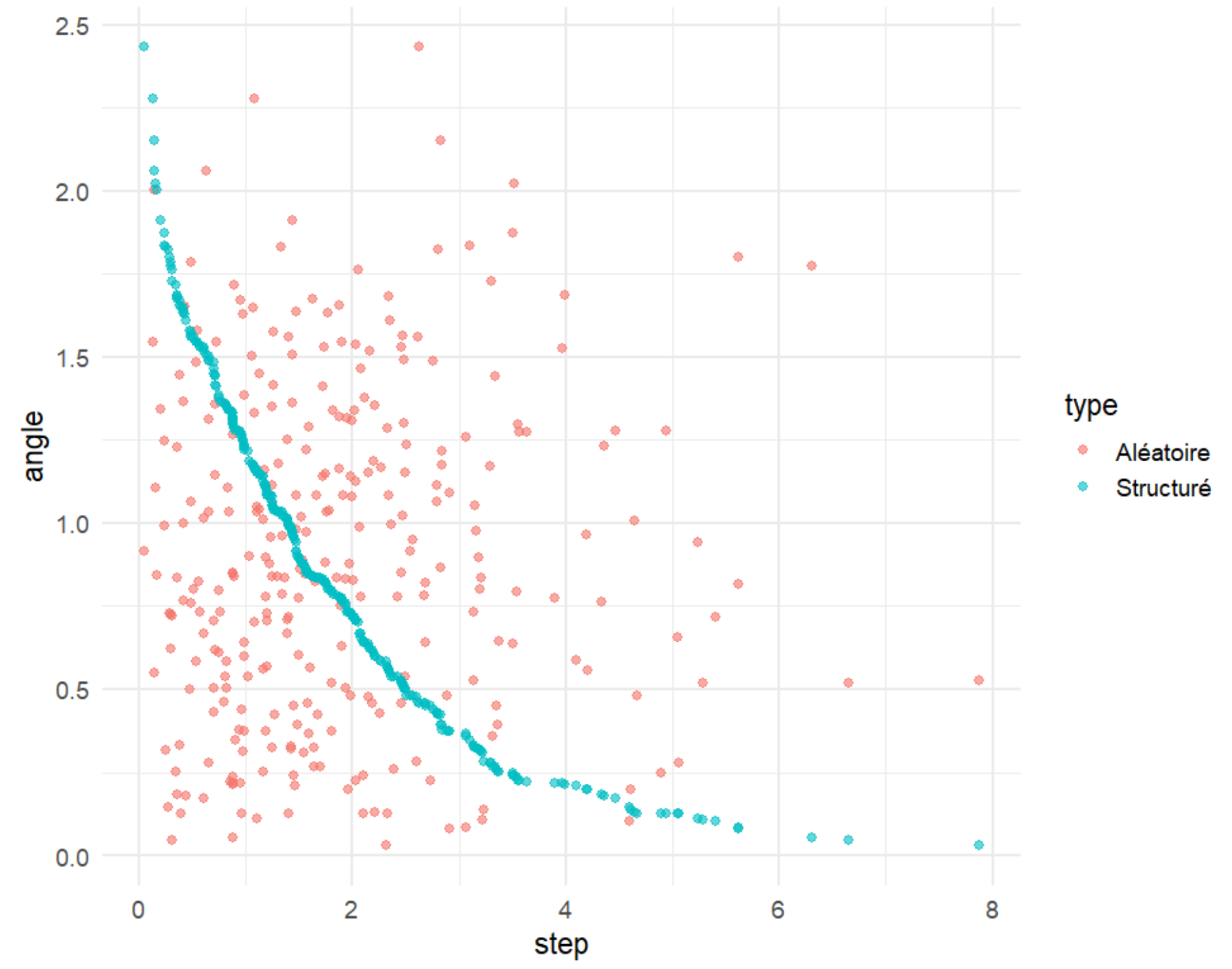
Association aléatoire



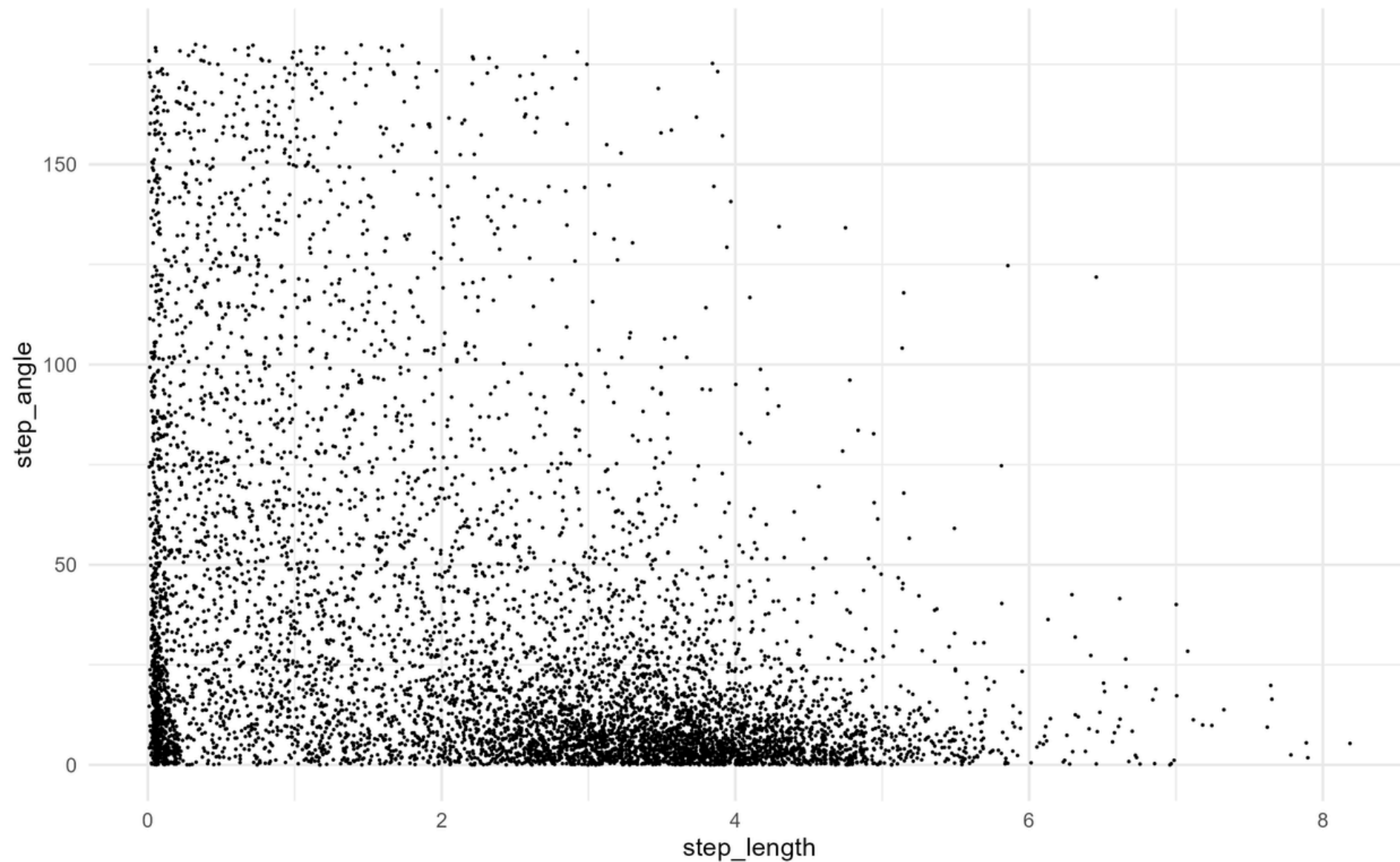
Association structuré



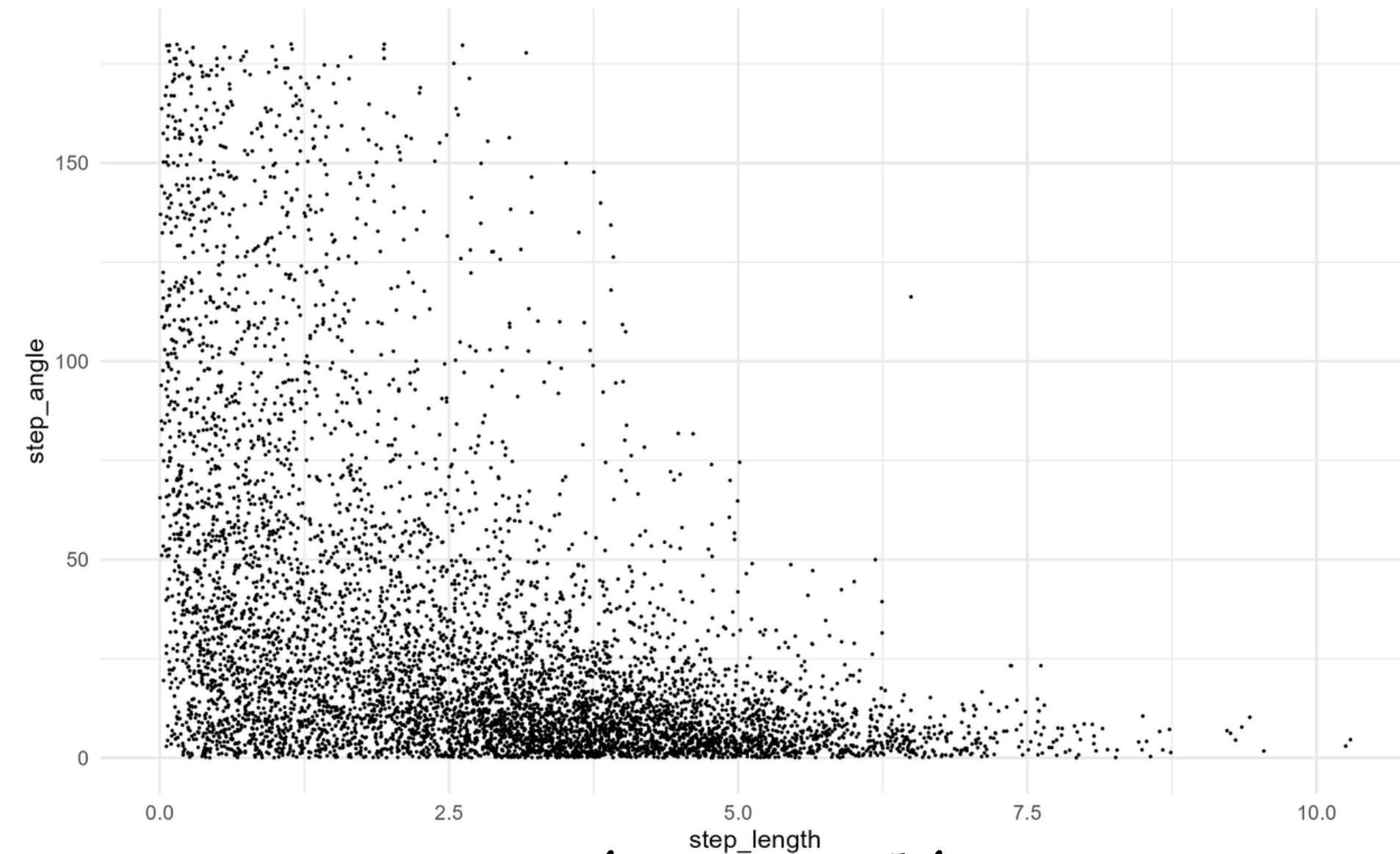
Structure d'association différente



Comparaison simple de métriques



Données réelles

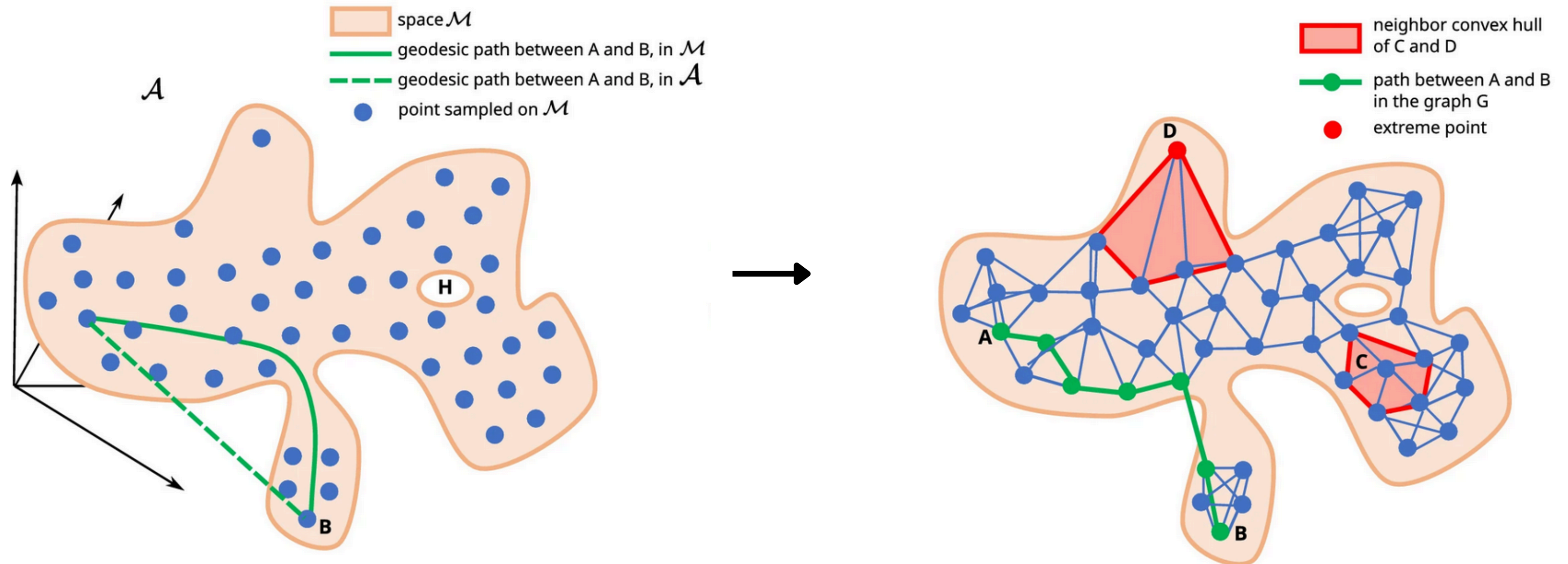


Données simulées
(modèle de diffusion)

Comment valider les modèles en considérant plus de métriques ?

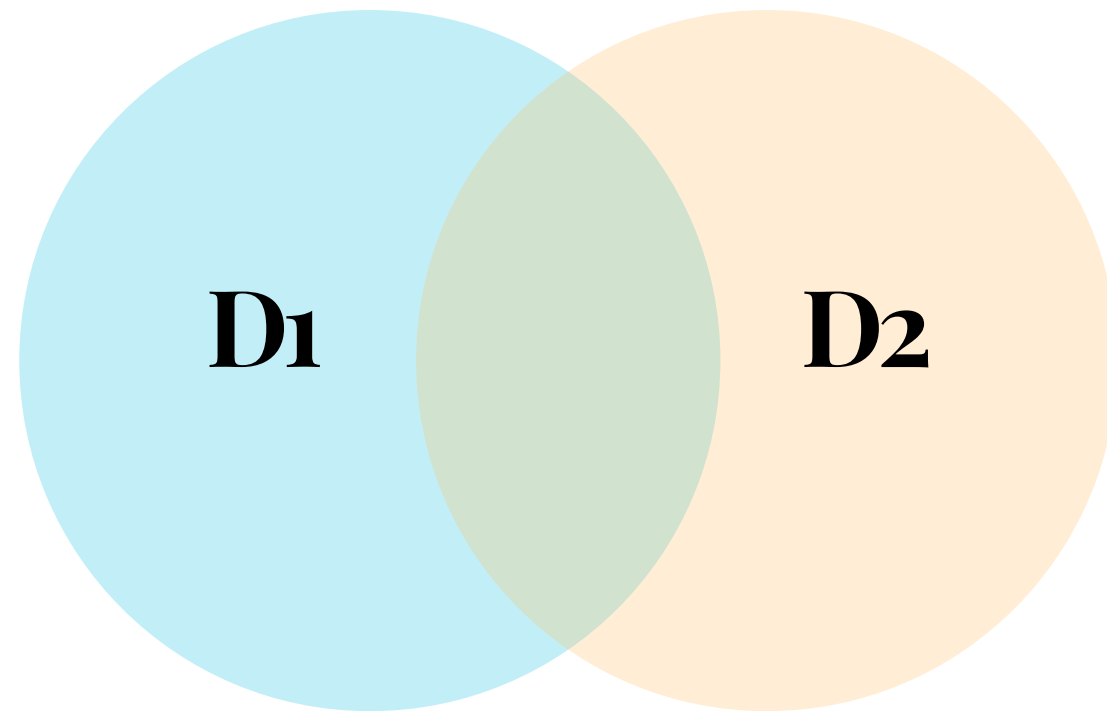
Comment comparer jeux de métriques ?

Datascape: combinaison d'un graphe des plus proches voisins, d'un ensemble d'enveloppes convexes et d'une métrique de distance respectant la topologie des données



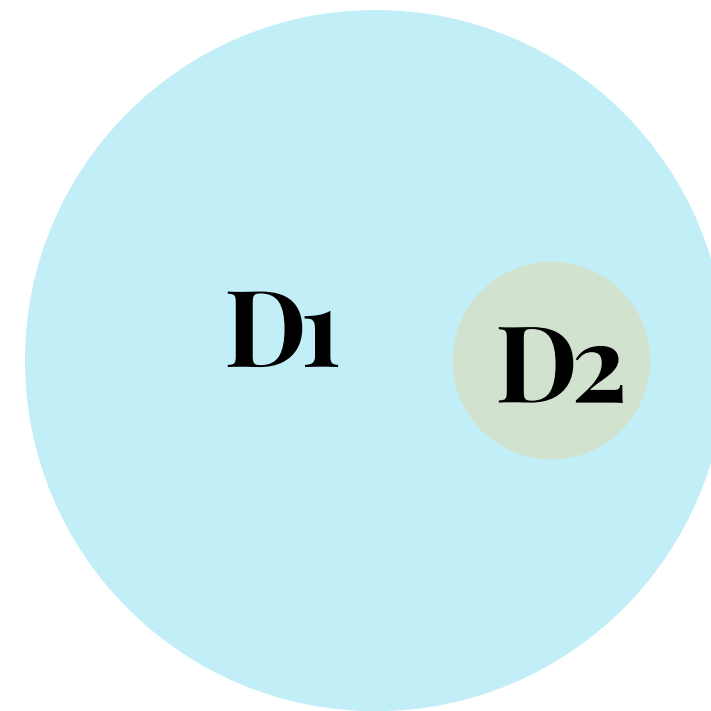
Comparer des datascares pour comparer des jeux de données

Critères d'inclusion entre deux datascares



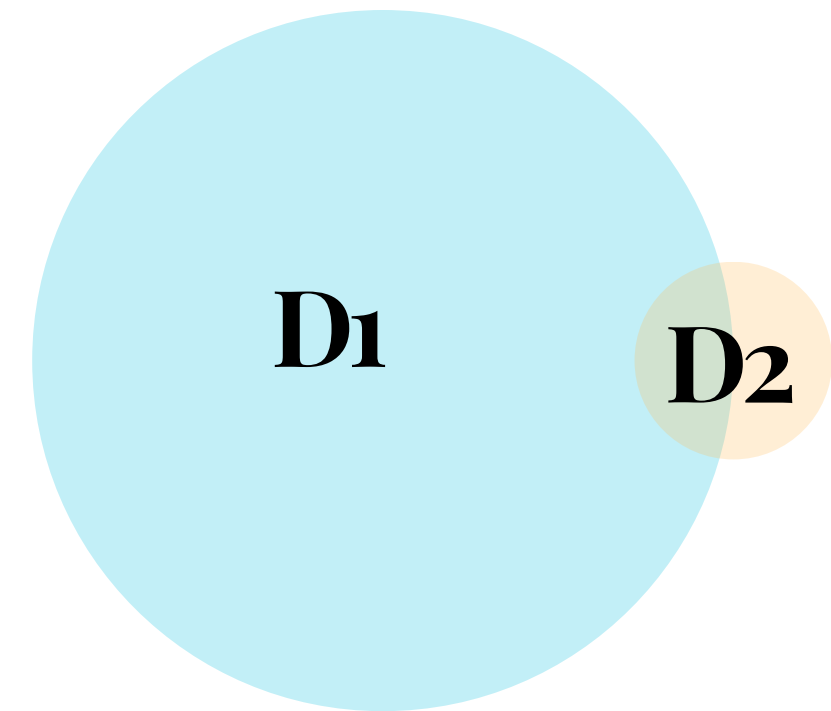
D1 in D2: 25%

D2 in D1: 25%



D1 in D2: 10%

D2 in D1: 100%



D1 in D2: 5%

D2 in D1: 50%

Critères de comparaison de deux datascares D_1 et D_2

Pourquoi le critère d'inclusion des points dans le datascape n'est pas suffisant ?



Tous les points de D_1 sont inclus dans D_2 et inversement. Pourtant distribution différente

Il faut regarder les densités (~ volumes enveloppes convexes)

Avantages et limites du Datascape

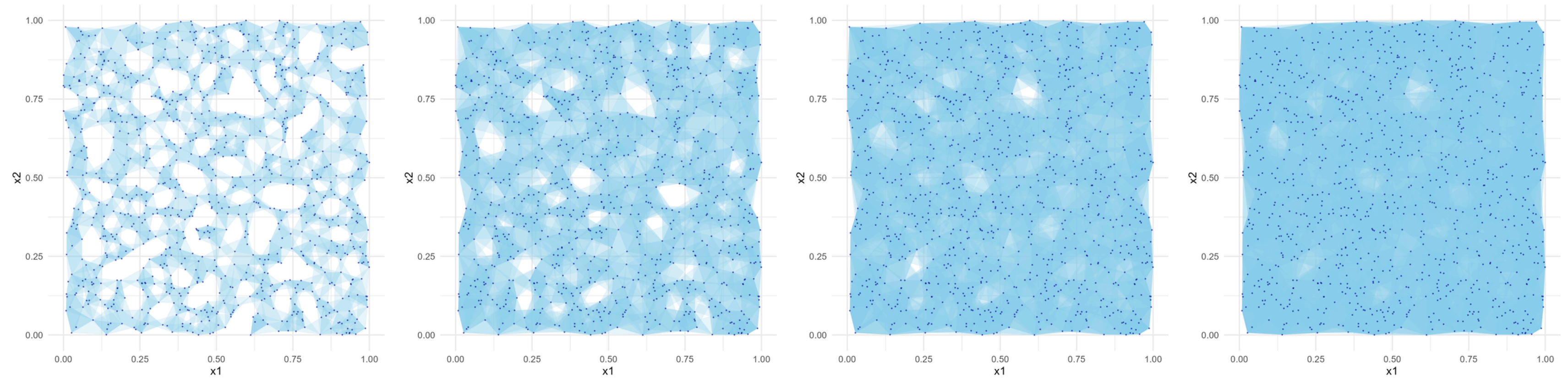
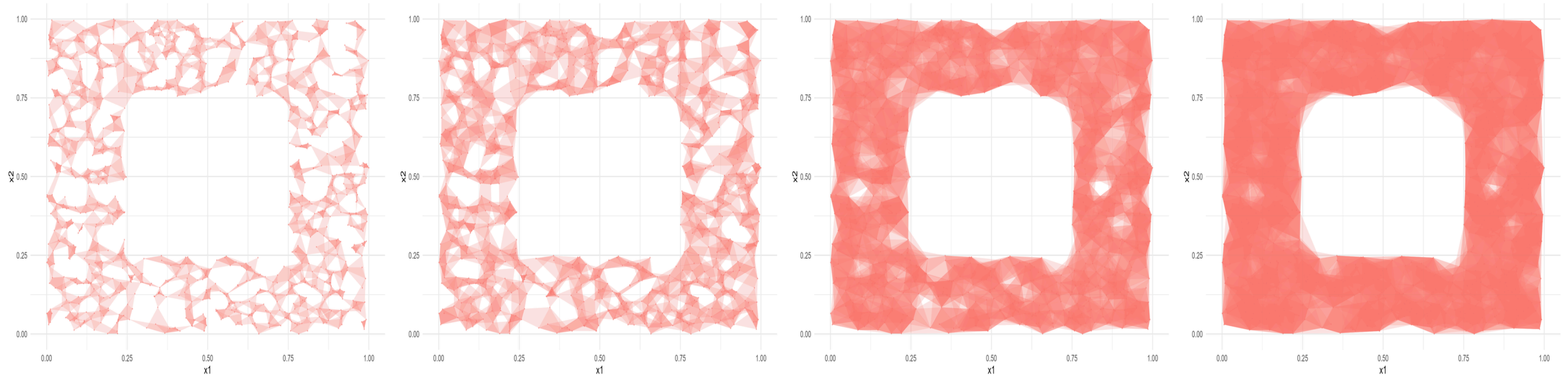
Avantages

- étude multidimensionnelle sans réduction de dimension
- calcul de métriques objectives possibles pour comparer des jeux de données

Limites et développements futurs

- Manque de stabilité des résultats: sensibilité au **nombre de voisins du graphe**, au **nombre de points** et à la **dimension** (paramétrage)
- risque de sur apprentissage
- trouver critères de comparaison pertinents

Comparaison de données test en fonction de k (lois uniformes)



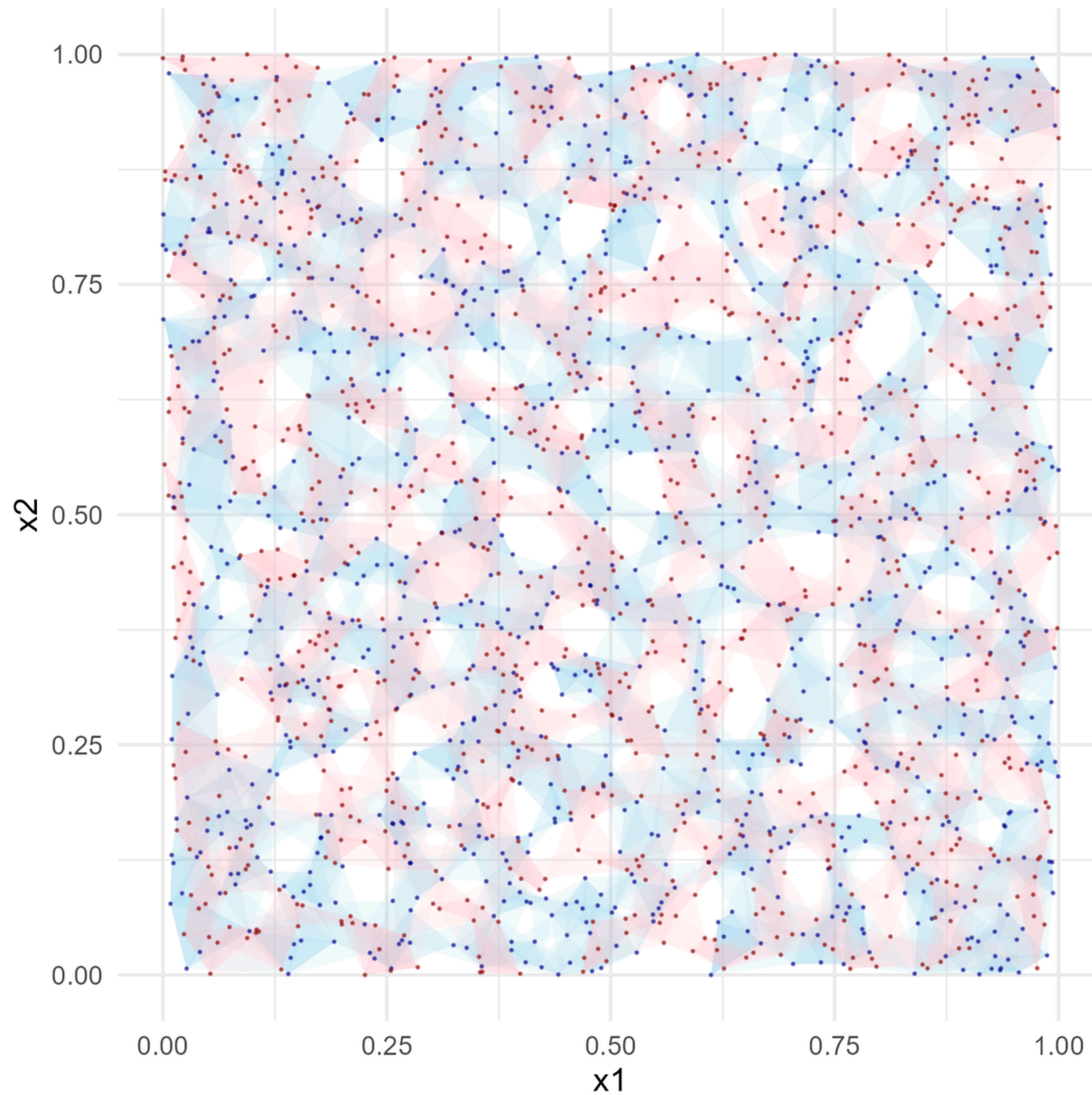
k=5

k=10

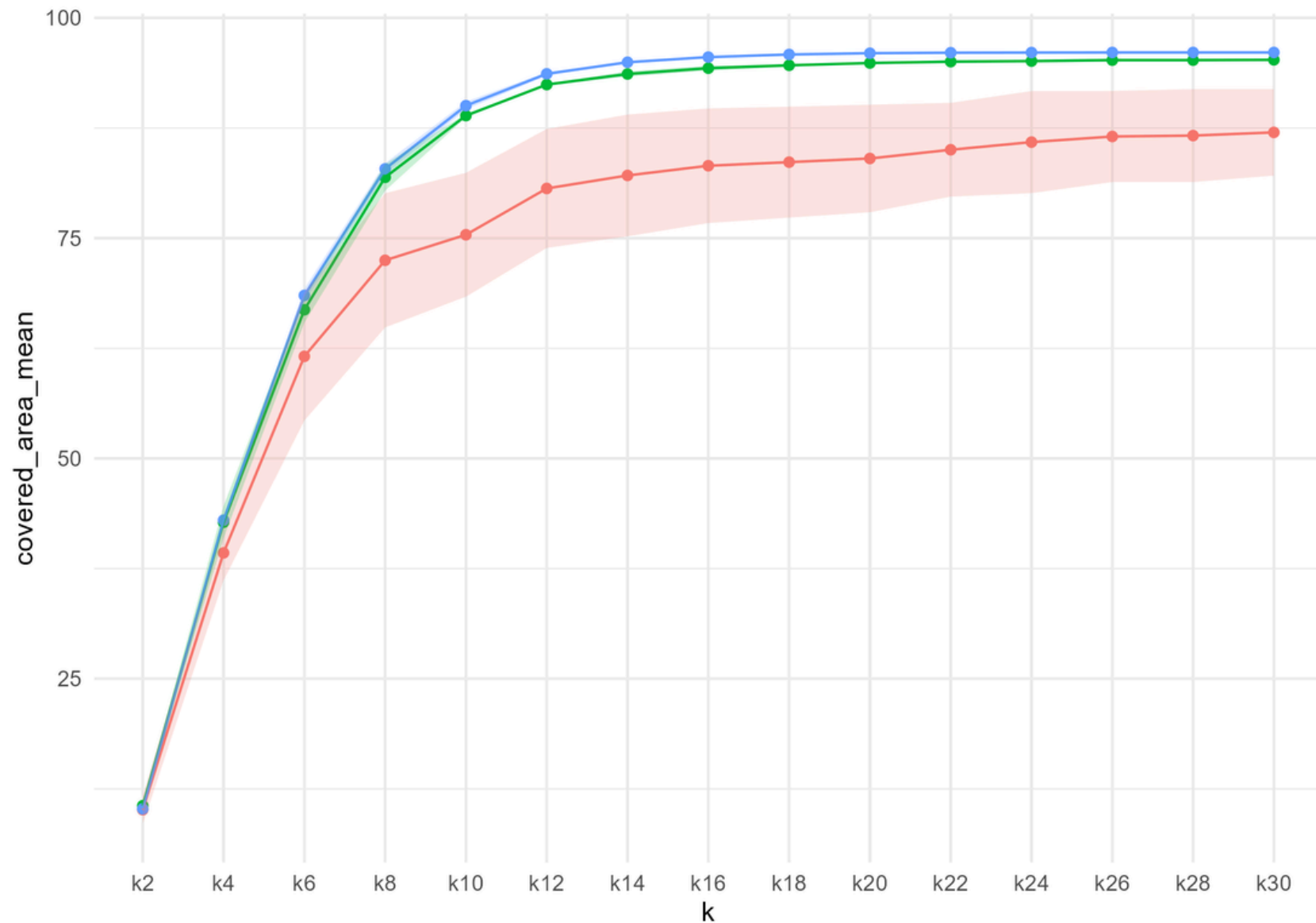
k=15

k=20

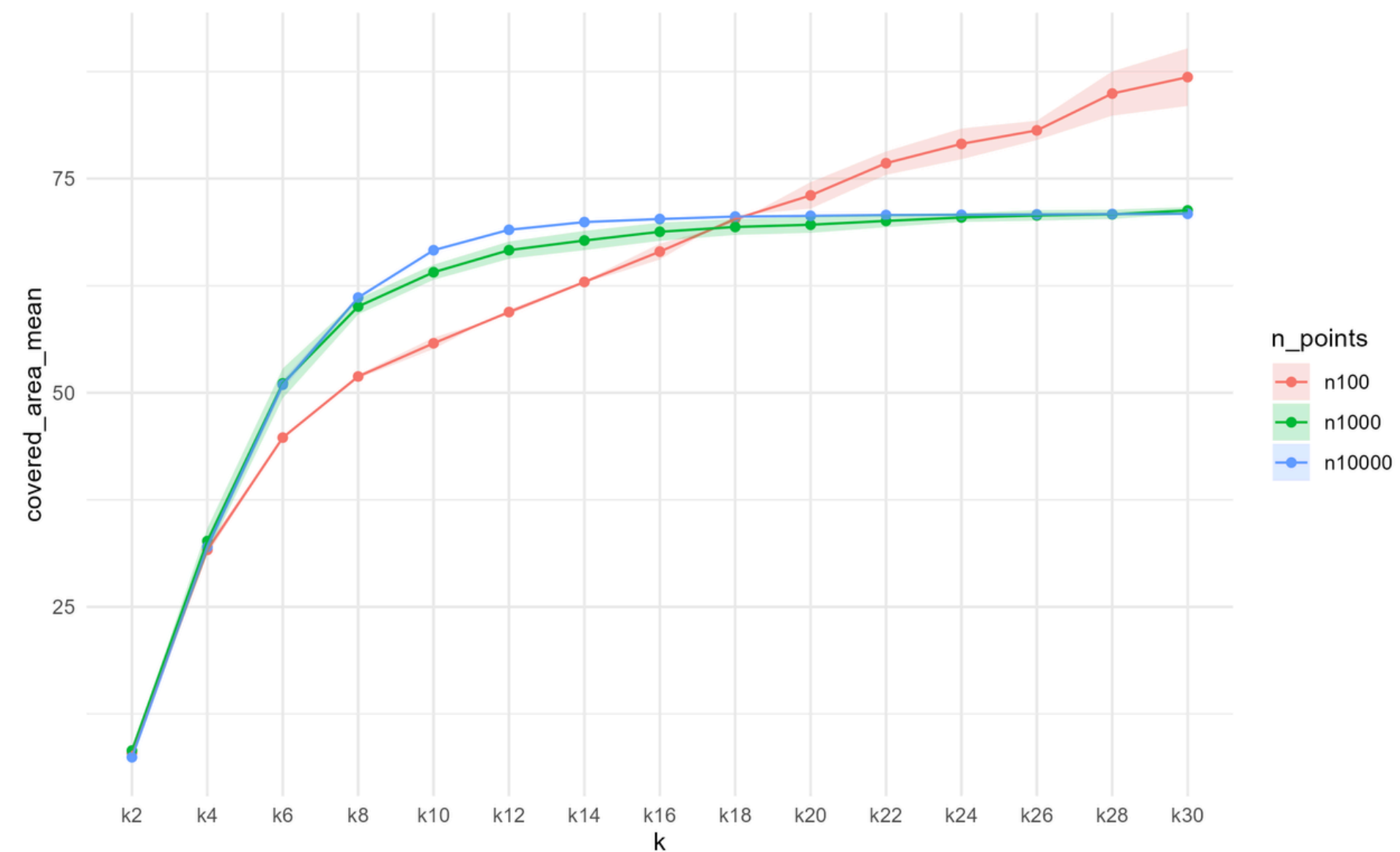
Comparaison de données test en fonction de k (lois uniformes)



Couverture en fonction de k et n



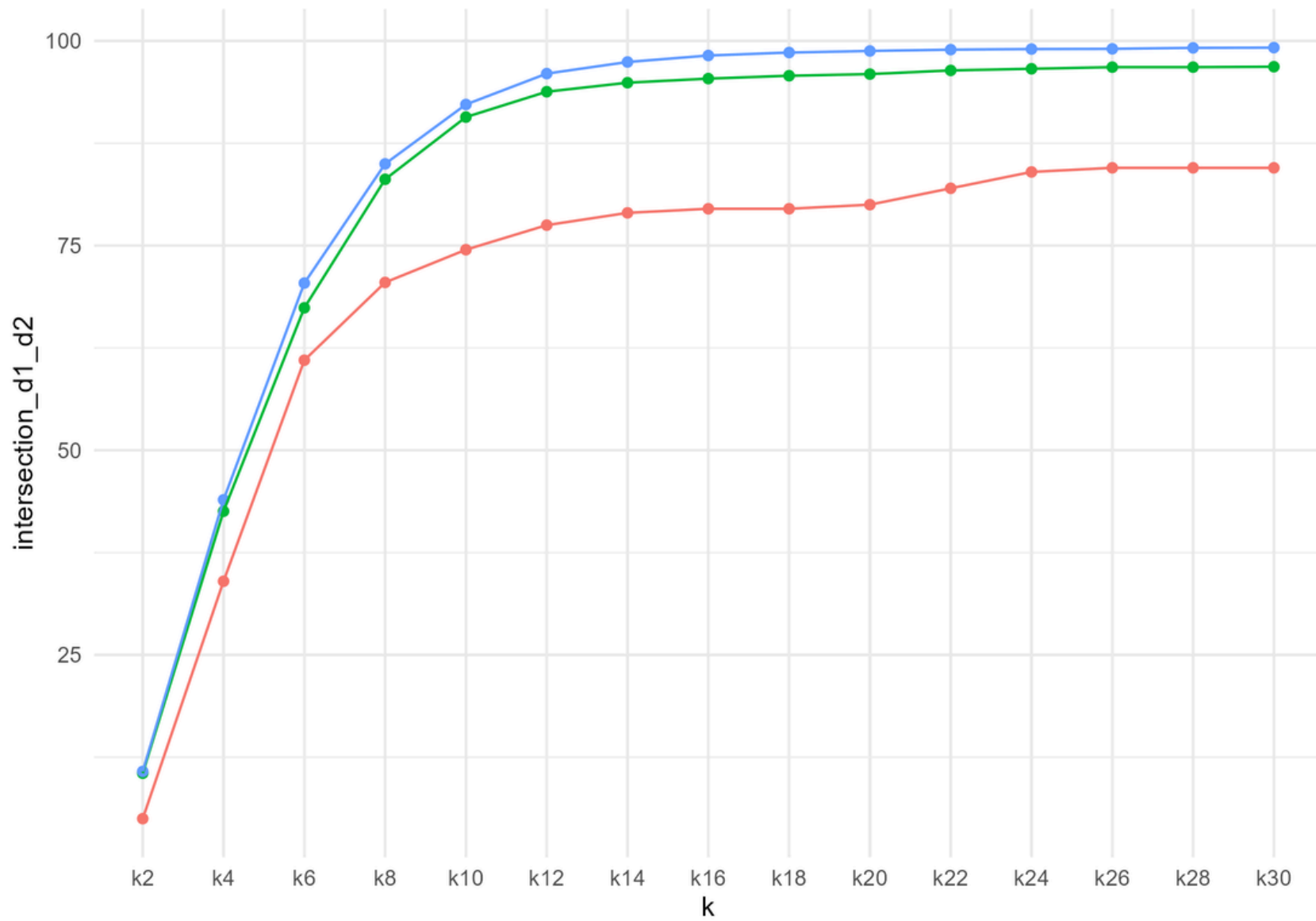
Carré plein



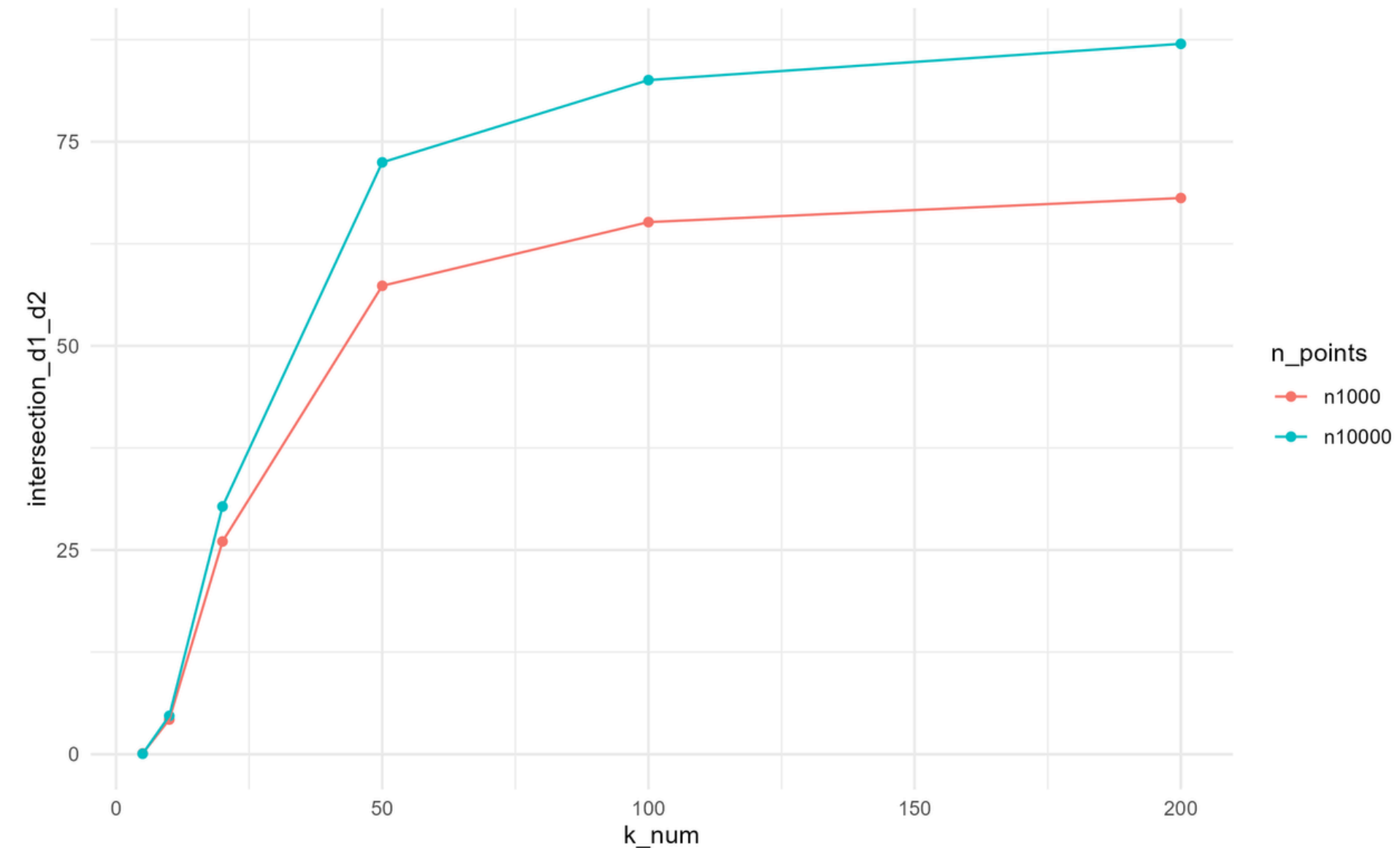
Carré avec trou

Le datascape ne permet pas de capturer la topologie de la donnée avec échantillonnage trop faible

Dimension



2D



5D

**Merci pour votre
attention !**

