

Un modèle général de marche aléatoire directionnelle: application au mouvement animalier.

Aurélien Nicosia

Avec Thierry Duchesne et Louis-Paul Rivest

Université Laval

Colloque ISM 2015

Plan de la présentation

Introduction

Développement des données GPS :



Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Bien comprendre l'organisation spatio-temporelle des animaux permet

Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Bien comprendre l'organisation spatio-temporelle des animaux permet

→ d'organiser la réintroduction d'espèces animales ;

Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Bien comprendre l'organisation spatio-temporelle des animaux permet

- d'organiser la réintroduction d'espèces animales ;
- de préserver l'intégrité écologique (transport de particules par les animaux) ;

Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Bien comprendre l'organisation spatio-temporelle des animaux permet

- d'organiser la réintroduction d'espèces animales ;
- de préserver l'intégrité écologique (transport de particules par les animaux) ;
- de contrôler le déplacement des animaux dans les parcs naturels.

Pourquoi faire ?

Un modèle de déplacement animalier, pourquoi faire ?

Bien comprendre l'organisation spatio-temporelle des animaux permet

- d'organiser la réintroduction d'espèces animales ;
- de préserver l'intégrité écologique (transport de particules par les animaux) ;
- de contrôler le déplacement des animaux dans les parcs naturels.

⇒ Les modèles de déplacement animalier sont très utiles en pratiques !

Introduction aux statistiques directionnelles.

Données directionnelles

Les données directionnelles peuvent être vues comme des directions ou angles, par exemple

Données directionnelles

Les données directionnelles peuvent être vues comme des directions ou angles, par exemple

- la direction du vent ;

Données directionnelles

Les données directionnelles peuvent être vues comme des directions ou angles, par exemple

- la direction du vent ;
- l'heure de connection aux réseaux sociaux ;

Données directionnelles

Les données directionnelles peuvent être vues comme des directions ou angles, par exemple

- la direction du vent ;
- l'heure de connection aux réseaux sociaux ;
- la direction de déplacement d'un animal ;

Données directionnelles

Les données directionnelles peuvent être vues comme des directions ou angles, par exemple

- la direction du vent ;
- l'heure de connection aux réseaux sociaux ;
- la direction de déplacement d'un animal ;

⇒ Données périodiques = données directionnelles.

Pourquoi faire ?

- Vous vous couchez un soir à $x_1 = 23$ h et le lendemain à $x_2 = 1$ h. A quelle heure vous couchez vous en moyenne ?

Pourquoi faire ?

- Vous vous couchez un soir à $x_1 = 23$ h et le lendemain à $x_2 = 1$ h. A quelle heure vous couchez vous en moyenne ?

$$\bar{x} = \frac{x_1 + x_2}{2} = \frac{23 + 1}{2} = 12 \text{ h} !!$$

Pourquoi faire ?

- Vous vous couchez un soir à $x_1 = 23$ h et le lendemain à $x_2 = 1$ h. A quelle heure vous couchez vous en moyenne ?

$$\bar{x} = \frac{x_1 + x_2}{2} = \frac{23 + 1}{2} = 12 \text{ h !!}$$

Les probabilités et statistiques classiques (dans $\mathbb{R}$) ne fonctionnent plus.

Pourquoi faire ?

- Vous vous couchez un soir à $x_1 = 23$ h et le lendemain à $x_2 = 1$ h. A quelle heure vous couchez vous en moyenne ?

$$\bar{x} = \frac{x_1 + x_2}{2} = \frac{23 + 1}{2} = 12 \text{ h} !!$$

Les probabilités et statistiques classiques (dans $\mathbb{R}$) ne fonctionnent plus.

- La notion d'ordre est beaucoup moins claire avec des angles :

$$1 \text{ degrés} = 361 \text{ degrés.}$$

Pourquoi faire ?

- Vous vous couchez un soir à $x_1 = 23$ h et le lendemain à $x_2 = 1$ h. A quelle heure vous couchez vous en moyenne ?

$$\bar{x} = \frac{x_1 + x_2}{2} = \frac{23 + 1}{2} = 12 \text{ h} !!$$

Les probabilités et statistiques classiques (dans $\mathbb{R}$) ne fonctionnent plus.

- La notion d'ordre est beaucoup moins claire avec des angles :

$$1 \text{ degrés} = 361 \text{ degrés.}$$

→ **Solution** : Angle = vecteur du cercle unité.

Statistique descriptive

Soient $x_1, \dots, x_n$ des points du cercle $\mathcal{S}^1$, alors il existe $\theta_1, \dots, \theta_n \in [0, 2\pi[$ tels que : $x_i = (\cos \theta_i, \sin \theta_i)^T$

Statistique descriptive

Soient $x_1, \dots, x_n$ des points du cercle S^1 , alors il existe $\theta_1, \dots, \theta_n \in [0, 2\pi[$ tels que : $x_i = (\cos \theta_i, \sin \theta_i)^T$

Centre de masse :

$$\bar{x} = \left(\frac{1}{n} \sum_{i=1}^n \cos \theta_i, \frac{1}{n} \sum_{i=1}^n \sin \theta_i \right) := (\bar{C}, \bar{S})$$

Statistique descriptive

Soient $x_1, \dots, x_n$ des points du cercle $\mathcal{S}^1$, alors il existe $\theta_1, \dots, \theta_n \in [0, 2\pi[$ tels que : $x_i = (\cos \theta_i, \sin \theta_i)^T$

Centre de masse :

$$\bar{x} = \left(\frac{1}{n} \sum_{i=1}^n \cos \theta_i, \frac{1}{n} \sum_{i=1}^n \sin \theta_i \right) := (\bar{C}, \bar{S})$$

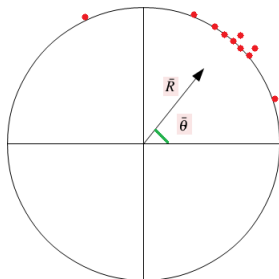
→ deux indices statistiques :

$\bar{\theta}$ et $\bar{R}$

- Direction moyenne : $\bar{\theta} = \arctan^* \left(\frac{\bar{S}}{\bar{C}} \right)$
- Longueur résultante : $\bar{R} = \sqrt{\bar{C}^2 + \bar{S}^2}$

Statistique classique Vs angulaire

	Statistique classique ($\mathbb{R}$)	Statistique angulaire (S^1)
Echantillon	$x_1, \dots, x_n$	$\theta_1, \dots, \theta_n$
Moyenne	$\bar{x}_n = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{\theta} = \arctan^* \left(\frac{\bar{S}}{\bar{C}} \right)$
Dispersion	$\sum_{i=1}^n (x_i - \bar{x}_n) = 0$	$\sum_{i=1}^n \sin(\theta_i - \bar{\theta}) = 0$
Variance	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_n)^2}{n-1}$	$\bar{V} = 1 - \bar{R}$



Densité sur le cercle

Soit Θ une variable aléatoire sur le cercle, alors on peut définir :

Densité de probabilité sur le cercle

On dit que f est une densité sur le cercle si :

- $f(\theta) \geq 0$
- $\int_0^{2\pi} f(\theta) d\theta = 1$
- f est 2π -périodique

Attention : La théorie classique ne fonctionne plus ! Par exemple :

$$\lim_{\theta \rightarrow +\infty} F(\theta) = \infty!$$

Loi de von Mises

La plus importante de toutes (équivalent de la loi normale dans $\mathbb{R}$) !

Loi de von Mises

La plus importante de toutes (équivalent de la loi normale dans $\mathbb{R}$) !

Loi de von Mises $vM(\mu, \kappa)$

La loi de von Mises de paramètres μ et κ a pour densité :

$$f(\theta|\mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \theta \in (-\pi, \pi)$$

où

Loi de von Mises

La plus importante de toutes (équivalent de la loi normale dans $\mathbb{R}$) !

Loi de von Mises $vM(\mu, \kappa)$

La loi de von Mises de paramètres μ et κ a pour densité :

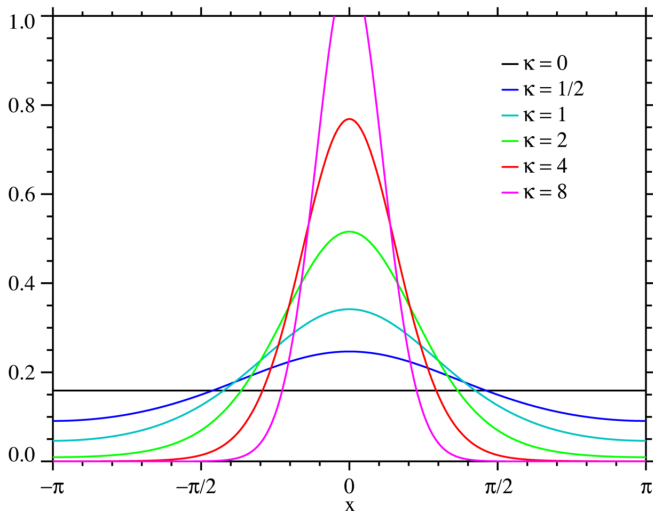
$$f(\theta | \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta - \mu)}, \theta \in (-\pi, \pi)$$

où

μ : direction moyenne.

κ : paramètre de concentration.

où $I_p(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos(p\theta) e^{\kappa \cos(\theta)} d\theta$ est la fonction de Bessel modifiée de première espèce et d'ordre p

Loi de von Mises ($\mu = 0$)

Régression angle-angle

Question : Comment prédire un angle par plusieurs angles explicatifs ?

Régression angle-angle

Question : Comment prédire un angle par plusieurs angles explicatifs ? Soit Y_i un angle observé et $x_i, i = 1, \dots, p$ angles prédicteurs. On définit alors :

Régression angle-angle

Question : Comment prédire un angle par plusieurs angles explicatifs ? Soit Y_i un angle observé et $\mathbf{x}_i$, $i = 1, \dots, p$ angles prédicteurs. On définit alors :

$$Y_i = \underbrace{\mu(y_i | \mathbf{x}_i; \kappa)}_{\text{direction moyenne}} + \epsilon_i$$

où

Régression angle-angle

Question : Comment prédire un angle par plusieurs angles explicatifs ? Soit Y_i un angle observé et $\mathbf{x}_i$, $i = 1, \dots, p$ angles prédicteurs. On définit alors :

$$Y_i = \underbrace{\mu(y_i | \mathbf{x}_i; \kappa)}_{\text{direction moyenne}} + \epsilon_i$$

où

- $\mu(y_i | \mathbf{x}_i, \kappa)$ est la direction (par rapport à l'origine) du vecteur :

$$\sum_{j=1}^p \kappa_j \begin{pmatrix} \cos(x_{ij}) \\ \sin(x_{ij}) \end{pmatrix}$$

Régression angle-angle

Question : Comment prédire un angle par plusieurs angles explicatifs ? Soit Y_i un angle observé et $\mathbf{x}_i$, $i = 1, \dots, p$ angles prédicteurs. On définit alors :

$$Y_i = \underbrace{\mu(y_i | \mathbf{x}_i; \kappa)}_{\text{direction moyenne}} + \epsilon_i$$

où

- $\mu(y_i | \mathbf{x}_i, \kappa)$ est la direction (par rapport à l'origine) du vecteur :

$$\sum_{j=1}^p \kappa_j \begin{pmatrix} \cos(x_{ij}) \\ \sin(x_{ij}) \end{pmatrix}$$

- $\epsilon_i \sim vM(0, \ell_i)$, où ℓ_i peut dépendre des prédicteurs (non-homogène).

Modèle de mouvement animalier.



Trajectoire

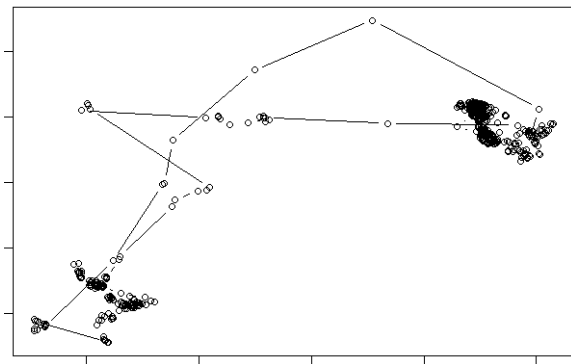


FIGURE : Caribou, forêt Boréale (Canada).

→ Avec les distances de déplacement, on soupçonne un comportement de type multi-états !

Le modèle

Hypothèse :

Le modèle

Hypothèse :

- La trajectoire de l'animal est régie par différents "états" (comportement) **non observés**.

Le modèle

Hypothèse :

- La trajectoire de l'animal est régie par différents "états" (comportement) **non observés**.
- La direction de déplacement est potentiellement influencée par des cibles environnementales.

Les données GPS

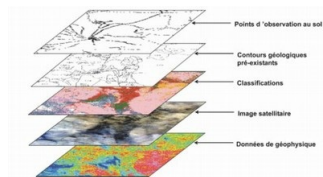
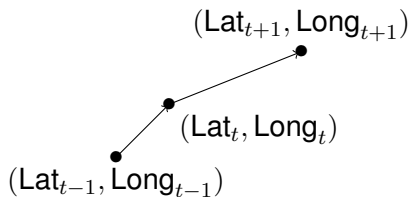


FIGURE : Données environnementales (GIS).

Les données

Les données sont :

$$\mathcal{F}_T := \{(y_t, d_t, \mathbf{x}_t), t = 0, \dots, T\}$$

où au temps t :

- $y_t \in [0, 2\pi)$ angle de direction.
- $d_t \geq 0$ distance parcourue.
- $\mathbf{x}_t = (x_{1t}, \dots, x_{pt})$ p variables explicatives qui sont potentiellement utiles à la prédiction de y_t .

Processus caché

- Supposons que l'animal manifeste K "états" différents non observés.

Processus caché

- Supposons que l'animal manifeste K "états" différents non observés.

$(S_t, t = 1, \dots, T)$ est un processus caché homogène à valeurs dans $\{1, \dots, K\}$, par exemple :

- S_t chaîne de Markov ;

Processus caché

- Supposons que l'animal manifeste K "états" différents non observés.

$(S_t, t = 1, \dots, T)$ est un processus caché homogène à valeurs dans $\{1, \dots, K\}$, par exemple :

- S_t chaîne de Markov ;

$\rightarrow p_{i,j} = \mathbb{P}(S_t = i | S_{t-1} = j)$ à estimer.

Distance de déplacement

- N'importe quelle loi dans $\mathbb{R}^+$: Weibull, Gamma...etc.

Distance de déplacement

- N'importe quelle loi dans $\mathbb{R}^+$: Weibull, Gamma...etc.
- Les paramètres de la distribution des distances dépendent du processus caché sous-jacent.

Direction de déplacement

Si l'animal est dans l'état k ($S_t = k$), on suppose que y_t est prédit par la direction précédente y_{t-1} et les angles aux cibles environnementales :

$$y_t = \mu_k(y_t | y_{t-1}, \mathbf{x}_t; \kappa) + \epsilon_t^k$$

Direction de déplacement

Si l'animal est dans l'état k ($S_t = k$), on suppose que y_t est prédit par la direction précédente y_{t-1} et les angles aux cibles environnementales :

$$y_t = \mu_k(y_t | y_{t-1}, \mathbf{x}_t; \kappa) + \epsilon_t^k$$

où ϵ_t^k est une von Mises de paramètre de concentration $\ell_t^{(k)}$.

Avantages :

Direction de déplacement

Si l'animal est dans l'état k ($S_t = k$), on suppose que y_t est prédit par la direction précédente y_{t-1} et les angles aux cibles environnementales :

$$y_t = \mu_k(y_t | y_{t-1}, \mathbf{x}_t; \kappa) + \epsilon_t^k$$

où ϵ_t^k est une von Mises de paramètre de concentration $\ell_t^{(k)}$.

Avantages :

→ Compromis directionnel.

Direction de déplacement

Si l'animal est dans l'état k ($S_t = k$), on suppose que y_t est prédit par la direction précédente y_{t-1} et les angles aux cibles environnementales :

$$y_t = \mu_k(y_t | y_{t-1}, \mathbf{x}_t; \kappa) + \epsilon_t^k$$

où ϵ_t^k est une von Mises de paramètre de concentration $\ell_t^{(k)}$.

Avantages :

- Compromis directionnel.
- Concentration consensus.

Compromis directionnel

$\mu_{\mathbf{k}}(y_t | y_{t-1}, \mathbf{x}_t; \kappa)$ est la direction du vecteur :

$$\underbrace{\kappa_0^{(\mathbf{k})} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix}}_{\text{Persistence directionnelle}} + \underbrace{\sum_{j=1}^p \kappa_j^{(\mathbf{k})} \begin{pmatrix} \cos(x_{jt}) \\ \sin(x_{jt}) \end{pmatrix}}_{\text{Influence des cibles}}$$

Compromis directionnel

$\mu_{\mathbf{k}}(y_t | y_{t-1}, \mathbf{x}_t; \kappa)$ est la direction du vecteur :

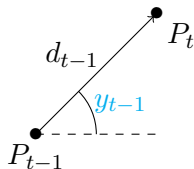
$$\underbrace{\kappa_0^{(\mathbf{k})} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix}}_{\text{Persistence directionnelle}} + \underbrace{\sum_{j=1}^p \kappa_j^{(\mathbf{k})} \begin{pmatrix} \cos(x_{jt}) \\ \sin(x_{jt}) \end{pmatrix}}_{\text{Influence des cibles}}$$

Interpretation des paramètres :

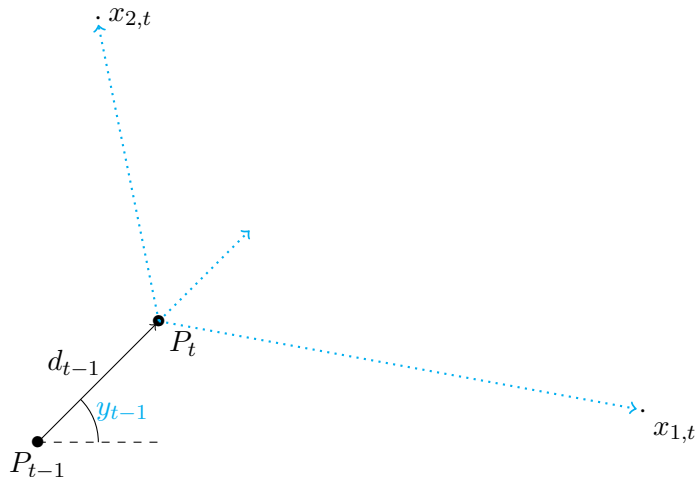
→ $\kappa_j^{(\mathbf{k})}$ mesure l'influence de la cible j sur la direction de l'animal.

- $\kappa_j^{(\mathbf{k})} < 0$: cible répulsive.
- $\kappa_j^{(\mathbf{k})} \approx 0$: cible non influente.
- $\kappa_j^{(\mathbf{k})} > 0$: cible attractive.

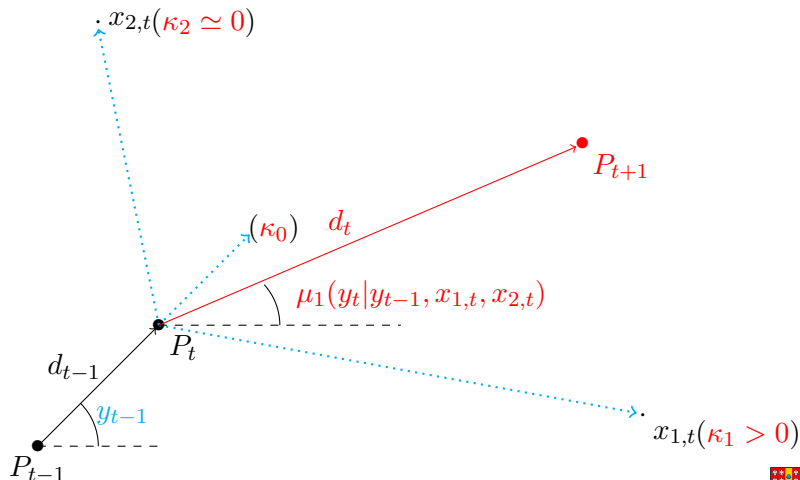
Compromis directionnel - état 1



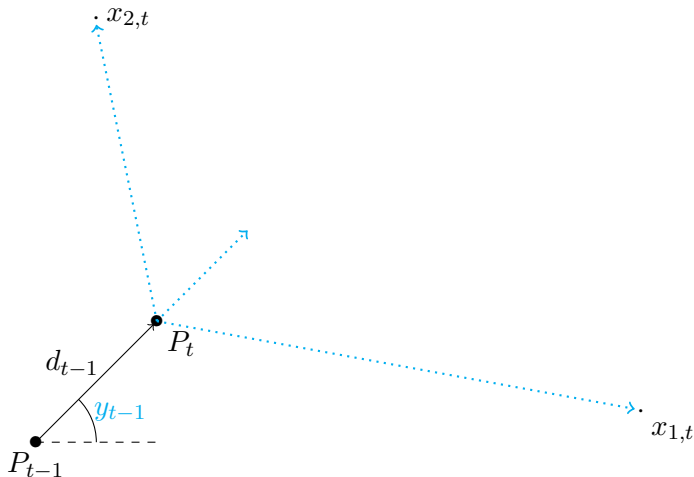
Compromis directionnel - état 1



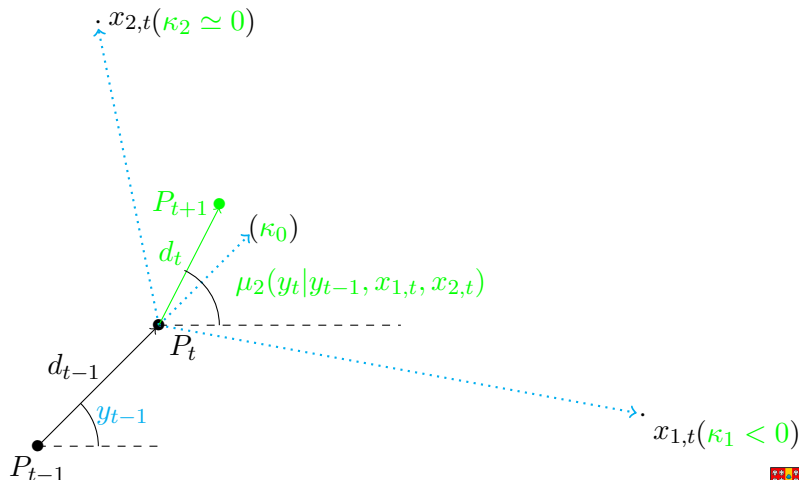
Compromis directionnel - état 1



Compromis directionnel - état 2



Compromis directionnel - état 2



Concentration consensus

La concentration des erreurs $\ell_t^{(k)}$ est la longueur du vecteur :

$$\kappa_0^{(k)} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix} + \sum_{j=1}^p \kappa_j^{(k)} \begin{pmatrix} \cos(x_{jt}) \\ \sin(x_{jt}) \end{pmatrix}$$

Interpretation :

Concentration consensus

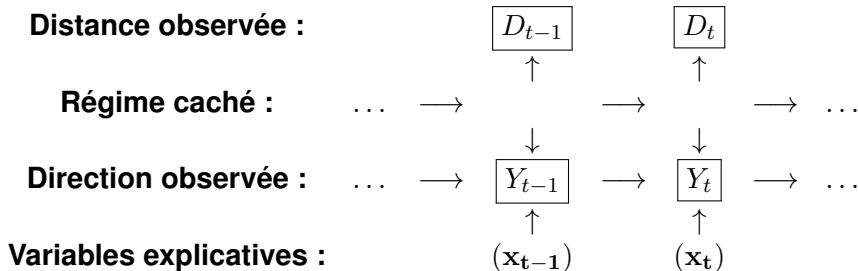
La concentration des erreurs $\ell_t^{(k)}$ est la longueur du vecteur :

$$\kappa_0^{(k)} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix} + \sum_{j=1}^p \kappa_j^{(k)} \begin{pmatrix} \cos(x_{jt}) \\ \sin(x_{jt}) \end{pmatrix}$$

Interpretation :

- Si y_{t-1} et $x_{1t}, \dots, x_{pt}$ pointent dans la même direction $\Rightarrow$ la concentration $\ell_t^{(k)}$ est grande (erreur ϵ_t faible).
- La magnitude des erreurs est inversement proportionnelle au degré de consensus entre la persistance directionnelle et les cibles environnementales.

Structure de dépendance



Estimation

- Modèle complètement paramétrique $\Rightarrow$ Estimateur du Maximum de Vraisemblance.

Estimation

- Modèle complètement paramétrique $\Rightarrow$ Estimateur du Maximum de Vraisemblance.
- Maximisation par l'algorithme EM.

Données réelles



FIGURE : Caribou, forêt Boréale (Canada).

Trajectoire

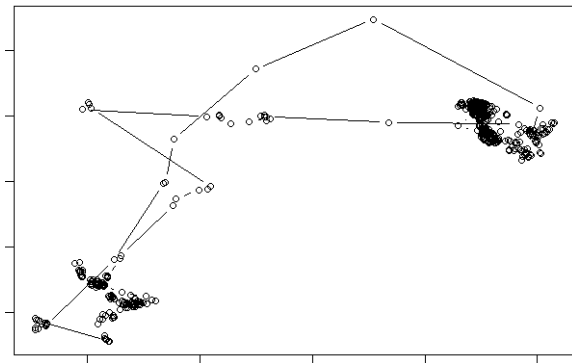


FIGURE : Caribou, forêt Boréale (Canada).

Exploration des différents états

Cas :	Distance $< d_{\text{critical}}$	Distance $> d_{\text{critical}}$
chang. de direction		
n	500	117
$\bar{\theta}$	1.933	0.40
$\bar{R}$	0.0253	0.3199
P_{value} Uniforme	> 0.15	< 0.01

Cibles environnementales

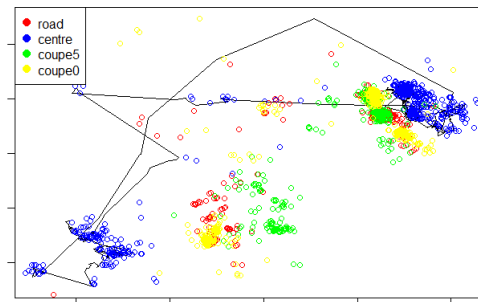


FIGURE : Caribou trajectory with explanatory variables.

Modèle à un état → centre et coupe5 ont de l'effet sur la direction du caribou.

Tableaux des résultats

état 1	Estim.	S.e.	état 2	Estim.	S.e.
$\kappa_{\text{persis.}}$	1.3569(*)	0.3821	$\kappa_{\text{persis.}}$	0.0465	0.0599
κ_{centre}	0.4455	0.3357	κ_{centre}	0.2587(*)	0.0677
κ_{coupe5}	0.1724	0.3285	κ_{coupe5}	0.1480(*)	0.0678
λ	0.3943	0.0804	λ	5.5963	0.2728
p_{11}	0.7436	0.1007	p_{22}	0.9851	0.0068

Comportement estimé



Conclusion

On a construit une marche aléatoire générale directionnelle
qui :

Conclusion

On a construit une marche aléatoire générale directionnelle
qui :

- prend en compte les caractéristiques environnementales
et les différents comportements non observés de l'animal ;

Conclusion

On a construit une marche aléatoire générale directionnelle qui :

- prend en compte les caractéristiques environnementales et les différents comportements non observés de l'animal ;
- modélise l'angle de direction et la distance parcouru selon ces différents états.

Conclusion

On a construit une marche aléatoire générale directionnelle qui :

- prend en compte les caractéristiques environnementales et les différents comportements non observés de l'animal ;
- modélise l'angle de direction et la distance parcouru selon ces différents états.

L'application numérique sur le Caribou à montré que

- la distance spécifie bien les comportements cachés mais aussi la direction de déplacement !

Conclusion

On a construit une marche aléatoire générale directionnelle qui :

- prend en compte les caractéristiques environnementales et les différents comportements non observés de l'animal ;
- modélise l'angle de direction et la distance parcouru selon ces différents états.

L'application numérique sur le Caribou à montré que

- la distance spécifie bien les comportements cachés mais aussi la direction de déplacement !
- le processus caché estimé donne une bonne idée spatio-temporelle du comportement de l'animal.

Futurs travaux

- Équivalence avec d'autres modèles de déplacement (régression logistique).

Futurs travaux

- Équivalence avec d'autres modèles de déplacement (régression logistique).
- Temps continu.

Merci pour votre attention !

