

# La masse d'un mammouth : étude d'un modèle convectif

Marine Djaffardjy, Théophile Moreau de Brevans, Émilie Mathian

## Abstract

L'évolution de la masse des espèces animales au cours des ères géologiques est un processus complexe dépendant de nombreux facteurs biologiques et écologiques. Il a pourtant pu être mathématiquement modélisé. Les modèles sont fondés sur le postulat de Cope selon lequel la taille des espèces tend à augmenter au cours des générations. Cette règle, considérée comme universelle, s'applique à de nombreux clades (mammifères, oiseaux, ...). Elle permet d'expliquer l'allure générale de la distribution des masses. Modéliser cette distribution asymétrique grâce à un terme convectif sera l'un des objectifs de notre étude. Nous tenterons aussi d'observer l'influence de la règle de Cope.

## Keywords

Convection — Schéma de Lax — Règle de Cope — Cladogenèse

## Introduction

### Fondamentaux biologiques

La règle de Cope date de 1887 [1]. Suite à l'observation de fossiles, il fait le postulat de l'augmentation de la taille des espèces au cours des générations. Bien que cette règle générale soit controversée, les jeux de données actuels semblent la confirmer. L'une des principales critiques est à l'absence de prise en compte des liens phylogénétiques [2]. En effet, les modèles présument de l'indépendance des données (absence de liens de parenté), en raison de données phylogénétiques peu fiables.

Le gain de taille prévu par la règle de Cope constitue un avantage sélectif. Biologiquement l'augmentation de la taille améliore la thermorégulation, l'endurance, la compétition inter-spécifique, ou encore l'accès aux ressources (augmentation de la gamme d'aliments disponibles) [1]. Un exemple d'adaptation de la taille des espèces face aux changements climatiques a été expliqué pour la période éocène. Les températures ont fortement augmenté au début de l'éocène et ont induit la disparition des forêts pour laisser place à de vastes plaines. Ces nouvelles niches écologiques ont impliqué l'augmentation de la taille des membres des herbivores contraints par la prédation. De plus le passage à une alimentation essentiellement constituée de cellulose est mieux adaptée aux métabolisme des herbivores de grande taille. Le gain de taille des herbivores implique l'augmentation de celle des prédateurs qui doivent exploiter les proies de toutes tailles, restreignant de ce fait la taille de leurs proies [2]. Si la règle de Cope implique une augmentation générale de la taille des animaux, celle-ci n'interdit pas des phases de diminution, bien que plus rares. Toutefois, il a déjà été démontré que la taille des animaux ne pouvait décroître en-dessous de certaines limites en raison de contraintes physiologiques et

thermiques. Ceci permet d'imposer une limite inférieure à notre modèle. La distribution de la masse des espèces suit une forme canonique où la classe modale se situe autour de 40g et où les espèces de petite taille ou de grande taille sont non seulement peu communes mais ont également une répartition asymétrique. Ainsi le Mammouth Impérial aurait eu une masse 10<sup>7</sup>g alors que le plus petit mammifère, la musaraigne pygmée, pèse 1.8g. Cette asymétrie autorise une limite supérieure peu contraignante. Enfin, le mécanisme crucial soutenant le modèle est que l'augmentation de taille accroît le risque d'extinction [3].

## Modèle et paramètres

### Modèle de Aaron Clauset

Aaron Clauset [3][4] est l'auteur de plusieurs études s'intéressant aux modèles de diffusion de la masse modélisée par l'équation suivante :

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} = D \frac{\partial^2 f}{\partial x^2} + (1 - A - Bx)f \quad (1)$$

$f(x, t)$  représente le nombre d'espèces ayant une masse équivalent à  $x = \ln(M)$  au temps  $t$ . Le paramètre  $v = \langle \ln \lambda \rangle$  est le biais de masse entre l'ancêtre et son descendant,  $D$  le coefficient de diffusion équivalent à  $D = \langle (\ln \lambda)^2 \rangle$ . Le paramètre  $\lambda$  est une variable aléatoire tirée d'une loi log-normale de paramètre  $v$ , ou une constante selon les simulations.  $D$  traduit pour chaque événement de spéciation un la masse peut être biaisée vers la gauche ou vers la droite.

D'après la règle de Cope  $\ln \lambda > 0$ . Enfin  $A - Bx$  traduit la probabilité d'extinction d'une espèce, corrélée au logarithme de la masse par le coefficient  $B > 0$ .

### Conditions initiales et conditions aux bords

Nous allons étudier cette équation avec deux types de conditions aux bords. Premièrement avec les conditions de Dirichlet-Neumann. Nous faisons ainsi les hypothèses sont que le poids minimal 2g, compatible avec la vie, ne peut être franchi et qu'il peut exister un flux au bord droit. Secondement nous étudierons le modèle avec les conditions de Dirichlet-Dirichlet, ce qui nous permet à nouveau de limiter la diffusion en deçà de 2g mais également de représenter l'absence de flux au niveau du bord droit. Sous cette hypothèse nous modélisons l'extinction des espèces ayant une masse en dessous d'une certaine limite fixée à  $10^7$  g. Celle-ci peut être atteignable d'après les conditions biologiques déterminées mais reste néanmoins peu probable.

### Simplification du modèle

Retirons le terme de convection pour l'analyse mathématique l'équation 1 s'écrit telle que :

$$\begin{aligned}\frac{\partial f}{\partial t} &= D \frac{\partial^2 f}{\partial x^2} + (1 - A - Bx)f \\ \frac{\partial f}{\partial t} &= D \Delta_x f + (1 - A - Bx)f\end{aligned}\quad (2)$$

### Fonctions propres :

Nous cherchons à présent les fonctions propres, notées  $\omega(x)$ , associées au deux types de conditions. Rappelons que les fonctions propres  $\omega(x)$ , avec  $x \in [0, M]$  doivent vérifier :

1.  $\omega(x)$  est au moins deux fois dérivable,
2. Il existe  $\lambda \in \mathbb{R}$  tel que  $\omega(x)'' = \lambda \omega(x)$
3.  $\omega(x) \equiv 0$

On cherche donc les solutions telles que  $\omega(x)'' = \lambda \omega(x)$  les solutions sont donc de la forme  $\omega(x) = e^{rx}$  ainsi  $\omega(x)' = re^{rx}$ , et  $\omega(x)'' = r^2 e^{rx}$  d'où  $r^2 e^{rx} = \lambda e^{rx}$ . On obtient  $r^2 = \lambda$ .

Pour les conditions de Dirichlet-Neumann :

D'après ces conditions aux bords on a par définition  $\omega(0) = 0$  et  $\omega(M)' = 0$

Cas 1  $\lambda > 0$  :

On obtient deux racines réelles  $r_1 = \sqrt{\lambda}$  et  $r_2 = -\sqrt{\lambda}$ , et deux solutions  $\omega_1(x) = e^{r_1 x}$  et  $\omega_2(x) = e^{r_2 x}$ . Par combinaison de ces deux solutions on a :

$$\begin{aligned}\omega(x) &= C_1 + e^{r_1 x} + C_2 e^{-r_1 x} \\ \omega'(x) &= r_1 (C_1 + e^{r_1 x} - C_2 e^{-r_1 x})\end{aligned}$$

Or d'après la condition de Dirichlet on a :  $\omega(x) = 0 \Rightarrow C_1 + C_2 = 0 \Rightarrow C_1 = -C_2$ . On peut réécrire la fonction  $\omega(x)$  telle que  $\omega(x) = C_1 (e^{r_1 x} - e^{-r_1 x})$ .

D'après la condition de Neumann on a :

$$\omega(M)' = 0 \Rightarrow r_1 C_1 (e^{r_1 M} + e^{-r_1 M}) = 0$$

On a alors deux solutions possibles :

$$\begin{cases} C_1 = 0 \\ e^{r_1 M} + e^{-r_1 M} = 0 \end{cases}$$

Or comme  $M > 0$  ceci implique que  $e^{r_1 M} + e^{-r_1 M} > 0$ . Donc on aurait  $C_1 = 0$  et  $C_2 = 0$ , ce qui est impossible dans le cadre de la troisième condition.

Cas 2  $\lambda = 0$  :

Si  $\lambda = 0$  alors  $r = 0$ , on a alors  $\omega(x)'' = 0 \Rightarrow \omega(x)' = 0$  et  $\omega(x) = C_1 x + C_2$ .

D'après Dirichlet on a  $\omega(0) = 0$  donc  $C_2 = 0$ . Notre fonction propre peut alors être écrite telle que  $\omega(x) = C_1 x \Rightarrow \omega(x)' = C_1$ , d'après les conditions de Neumann on a :  $\omega(M)' = 0$  et donc  $C_1 = 0$  ce qui est impossible en vertu de la condition 3.

Cas 3  $\lambda < 0$  :

On obtient deux racines complexes telles que :

$r_1 = i\sqrt{|\lambda|}$  et  $r_2 = -i\sqrt{|\lambda|}$  de la forme  $r = \alpha + i\beta$  et deux solutions possibles  $\omega_1(x) = e^{i\sqrt{|\lambda|x}}$  et  $\omega_2(x) = e^{-i\sqrt{|\lambda|x}}$ .

Les solutions sont alors données par :

$$\omega(x) = e^{\alpha x} (C_1 \cos(\beta x) + C_2 \sin(\beta x))$$

où  $\alpha = 0$  et  $\beta = \sqrt{|\lambda|}$  donc

$$\omega(x) = C_1 \cos(\sqrt{|\lambda|x}) + C_2 \sin(\sqrt{|\lambda|x})$$

et  $\omega'(x) = \sqrt{|\lambda|} (-C_1 \sin(\sqrt{|\lambda|x}) + C_2 \cos(\sqrt{|\lambda|x}))$ .

D'après les conditions de Dirichlet on a :

$$\omega(0) = 0 \Rightarrow \sqrt{|\lambda|} (C_1 \sin(0) + C_2 \cos(0)) = 0$$

donc  $C_1 = 0$ , on peut réécrire  $\omega(x)$  telle que :

$$\omega(x) = C_2 \sin(\sqrt{|\lambda|x}) \Rightarrow \omega(x)' = \sqrt{|\lambda|} C_2 \cos(\sqrt{|\lambda|x})$$

Or d'après les conditions de Neumann  $\omega(M)' = 0$  donc

$$\sqrt{|\lambda|} C_2 \cos(\sqrt{|\lambda|M}) = 0$$

On a deux solutions sont possibles :

$$\begin{cases} C_2 = 0 \\ \cos(\sqrt{|\lambda|M}) = 0 \end{cases}$$

Pour respecter la condition 3 la seule solution possible est  $\cos(\sqrt{|\lambda|M}) = 0 \Rightarrow \sqrt{|\lambda|M} = \frac{k\pi}{2}$ . On déduit que les valeurs propres sont de la forme  $\lambda_k = -(\frac{k\pi}{2M})^2$  et les fonctions propres sont telles que  $\omega_k(x) = C_2 \sin(\frac{k\pi}{M}x)$ , ce qui peut être généralisé par  $\omega_k(x) = \sin(\frac{k\pi}{M}x)$ .

Pour les conditions de Dirichlet-Dirichlet :

D'après ces conditions aux bords on a par définition  $\omega(0) = 0$  et  $\omega(M) = 0$

Cas 1  $\lambda > 0$  :

On obtient deux racines réelles  $r_1 = \sqrt{\lambda}$  et  $r_2 = -\sqrt{\lambda}$ , et deux solutions  $\omega_1(x) = e^{r_1 x}$  et  $\omega_2(x) = e^{-r_1 x}$ . Par combinaison de ces deux solutions on a :

$$\begin{aligned}\omega(x) &= C_1 + e^{r_1 x} + C_2 e^{-r_1 x} \\ \omega'(x) &= r_1 (C_1 + e^{r_1 x} - C_2 e^{-r_1 x})\end{aligned}$$

Or d'après la condition de Dirichlet on a :

$$\omega(x) = 0 \Rightarrow C_1 + C_2 = 0 \Rightarrow C_1 = -C_2$$

. On peut réécrire la fonction  $\omega(x)$  telle que  $\omega(x) = C_1(e^{r_1x} - e^{r_2x})$ .

D'après la seconde condition de Dirichlet on a :

$$\omega(M) = 0 \Rightarrow C_1(e^{r_1M} + e^{r_2M}) = 0$$

On a alors deux solutions possibles :

$$\begin{cases} C_1 = 0 \\ e^{r_1M} + e^{r_2M} = 0 \end{cases}$$

Or  $r_1M = \ln(M) = 0$  impliquerait que  $\omega(x) \equiv 0$  ce qui est impossible en raison de la troisième condition.

Cas 2  $\lambda = 0$  :

Si  $\lambda = 0$  alors  $r = 0$ , on a alors  $\omega(x)'' = 0 \Rightarrow \omega(x)' = 0$  et  $\omega(x) = C_1x + C_2$ .

D'après Dirichlet on a  $\omega(0) = 0$  donc  $C_2 = 0$ . La fonction  $\omega(x)$  peut alors être écrite telle que  $\omega(x) = C_1x$ . Or d'après la seconde condition de Dirichlet  $\omega(M) = C_1M = 0 \Rightarrow C_1 = 0$ . En raison de la condition 3 cette solution est impossible.

Cas 3 :  $\lambda < 0$  :

On obtient deux racines complexes telles que :

$r_1 = i\sqrt{|\lambda|}$  et  $r_2 = -i\sqrt{|\lambda|}$  de la forme  $r = \alpha + i\beta$ .

Les solutions sont alors données par :

$$\omega(x) = e^{\alpha x}(C_1 \cos(\beta x) + C_2 \sin(\beta x))$$

où  $\alpha = 0$  et  $\beta = \sqrt{|\lambda|}$  donc

$$\omega(x) = C_1 \cos(\sqrt{|\lambda|x}) + C_2 \sin(\sqrt{|\lambda|x})$$

et  $\omega'(x) = \sqrt{|\lambda|}(C_1 \sin(\sqrt{|\lambda|x}) + C_2 \cos(\sqrt{|\lambda|x}))$ .

D'après les conditions de Dirichlet on a :

$$\omega(0) = 0 \Rightarrow \sqrt{|\lambda|}(C_1 \sin(0) + C_2 \cos(0)) = 0$$

donc  $C_1 = 0$ , on peut réécrire  $\omega(x)$  telle que :

$$\omega(x) = C_2 \sin(\sqrt{|\lambda|x})$$

Or d'après la seconde condition de Dirichlet  $\omega(M) = 0$  donc :

$$\sqrt{|\lambda|}C_2 \sin(\sqrt{|\lambda|M}) = 0$$

On a deux solutions possibles :

$$\begin{cases} C_2 = 0 \\ \sin(\sqrt{|\lambda|M}) = 0 \end{cases}$$

Pour respecter la condition 3 la seule solution possible est  $\sin(\sqrt{|\lambda|M}) = 0 \Rightarrow \sqrt{|\lambda|M} = \frac{k\pi}{2}$ . On déduit que les valeurs propres sont de la forme  $\lambda_k = -(\frac{k\pi}{M})^2$  avec  $k \in \mathbb{N}^*$  et les fonctions propres sont telles que  $\omega_k(x) = C_2 \sin(\frac{k\pi x}{M})$ , ce qui peut être généralisé par  $\omega_k(x) = \sin(\frac{k\pi x}{M})$ .

### Estimation des paramètres :

Reprenons l'estimation statistique des paramètres du modèle établi sur un jeu de données répertoriant 4022 espèces de mammifères fossiles [5] :

- $x_{min} = 2g$
- $x_{max} = 10^7g$
- $A = 0.221$
- $B = 0.4064$
- $D = 0.508$
- $v = 0.109$

Rappelons par ailleurs que le modèle estime la distribution du logarithme de la masse.

### Caractérisation des équilibres :

Notons  $f^*$  les équilibres de l'équation 2 rappelons que ceux-ci vérifient :

$$\frac{\partial f^*}{\partial t} = 0 \quad \text{et} \quad \Delta x f^* = 0$$

Alors on a  $(1 - A - Bx)f^* = 0$  notons  $g(f, x)$  la fonction telle que :

$$g(f, x) = (1 - A - Bx)f$$

Nous avons donc deux équilibres notés  $f_0^* = 0$  et  $x_0^* = \frac{1-A}{B}$ .

### Étude de la stabilité :

Au point d'équilibre  $f_0^*$  :

$$\frac{\partial g}{\partial f} = (1 - A - Bx) \rightarrow g(f_0^*, x)' = (1 - A - Bx)$$

D'après nos conditions  $\lambda_k = -(\frac{k\pi}{2M})^2$  et l'estimation des paramètres nous obtenons :

$$\begin{aligned} g(f_0^*, x)' + D\lambda_k &= (1 - A - Bx) - D(\frac{\pi}{M})^2 k^2 \\ &= 1 - 0.221 - 0.4064x - 0.508(\frac{\pi}{\ln(10^7)})^2 k^2 \end{aligned}$$

Cet équilibre est localement asymptotiquement stable (LAS)

si  $g(f_0^*, x)' + D\lambda_k < 0$  donc si  $k \in [-\sqrt{\frac{A-1+Bx}{-D(\pi/M)^2}}, \sqrt{\frac{A-1+Bx}{-D(\pi/M)^2}}]$ .

Les solutions de cette inéquation existent si  $x > 6.8g$ , soit lorsque  $x$  tend vers l'infini.

D'après cette caractérisation nous pouvons affirmer qu'il existe un moment où la répartition de la masse sera à l'équilibre, et donc qu'il existe une condition optimale de stabilité des espèces en terme de masse.

Au point d'équilibre  $x_0^*$  :

$$\frac{\partial g}{\partial x} = Bf \rightarrow g(f, x_0^*)' = Bf$$

De même cet équilibre est LAS si  $g(f, x_0^*)' + D\lambda_k < 0$ . Or par hypothèse  $f(x, t) > 0$  et comme  $D\lambda_k < 0$ , cette équilibre sera instable quelle que soit  $k$ .

L'interprétation de cette équilibre reste relativement, en effet il implique qu'un ensemble d'espèce de mammifères d'une masse proche de 2g risquent de s'éteindre ou d'augmenter en masse.

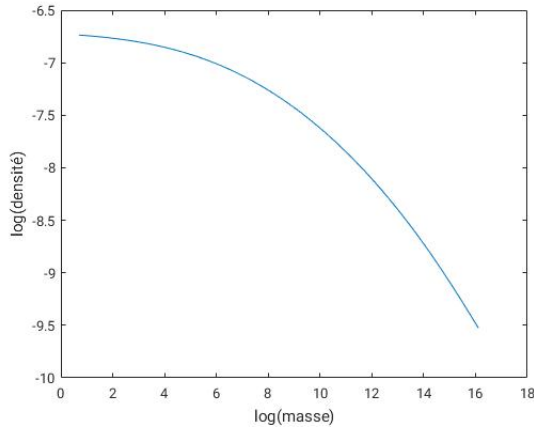
## Implémentation numérique

Afin d'observer l'influence du terme de convection, nous proposons d'implémenter deux modèles un premier sans ce terme de convection et un second prenant en compte ce dernier. Nous étudierons le problème par deux approches numériques, dans un premier temps avec un modèle continu dans le temps, puis par une modélisation individu centré.

### Schéma numérique continu

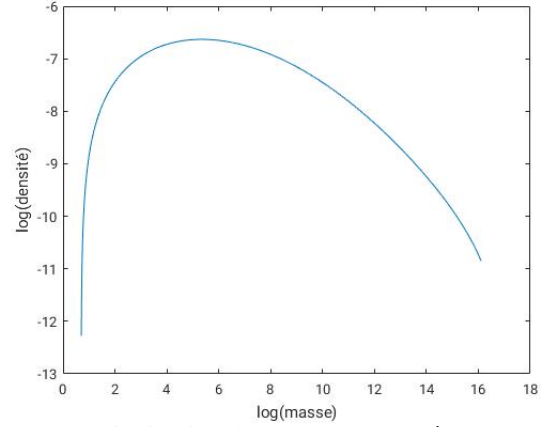
#### Schéma sans convection

La simulation effectuée avec ce modèle donne à voir une distribution gaussienne dans les premiers temps. Puis, au bout d'un temps suffisamment grand la distribution perd ce caractère gaussien en s'étalant sur la gamme de masses autorisée par le modèle (fig. 1). On perd donc la valeur modale de 40g observée dans les données au profit des espèces de masse minimale. L'intervalle des densités est également très réduit.



**Figure 1.** Distribution de la masse des espèces d'après le modèle sans convection en échelle  $\log(\text{densité})=f(\log(\text{masse}))$

Précisons que la figure 1 est obtenue avec des conditions de Neumann imposées aux bords. Le modèle serait sûrement plus proche des données si on imposait une condition de Dirichlet à gauche pour tenir compte de l'impossibilité physiologique pour un mammifère de peser moins de 2g. Si on fixe la valeur 2g à 0, on obtient la courbe suivante (fig. 2) après un temps de simulation grand.



**Figure 2.** Distribution de la masse des espèces d'après le modèle sans convection en échelle  $\log(\text{densité})=f(\log(\text{masse}))$

L'ajout du terme de convection a présenté certaines difficultés algorithmiques qui nous ont conduit à utiliser la méthode de Lax-Friedrichs.

#### Schéma avec convection

Afin de simplifier l'implémentation numérique de l'équation (1), nous proposons de scinder cette équation en une somme de plusieurs termes. Nous avons choisi le schéma de Lax-Friedrichs pour calculer le terme de convection  $v \frac{\partial f}{\partial x}$ , la méthode de Crank Nicolson pour la diffusion et un schéma explicite pour le terme de réaction. Écrivons à présent la méthode numérique, le pas de temps est noté  $n$  et le pas d'espace  $j$ , enfin la fonction  $f(x,t)$  discrétisée est notée  $U$ .

Schéma de Lax-Friedrichs - convection :

$$\frac{\partial U}{\partial t} + v \frac{\partial U}{\partial x} = v \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} + \frac{U_j^{n+1} - \frac{1}{2}U_{j+1}^n + U_{j-1}^n}{\Delta t}$$

Schéma de Crank Nicholson - Diffusion:

$$\begin{aligned} \frac{\partial U}{\partial t} &= D \frac{\partial^2 U}{\partial x^2} \\ \frac{U_j^{n+1} - U_j^n}{\Delta t} &= D \frac{(U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}) + (U_{j+1}^n - 2U_j^n + U_{j-1}^n)}{2(\Delta x)^2} \end{aligned}$$

Schéma explicite - Réaction

$$\begin{aligned} \frac{\partial U}{\partial t} &= f(t, U) \\ U_{j+1}^{n+1} &= U_j^n + \Delta t(1 - A - Bx)U_j^n \end{aligned}$$

Finalement en isolant le terme  $U_j^{n+1}$  nous obtenons un schéma de forme générale :

$$\begin{aligned} \frac{U_j^{n+1}}{\Delta t} &= -\frac{1/2}{\Delta t}(-U_{j+1}^n + U_{j-1}^n) + \\ \frac{D}{2} \frac{L_0^{n+1} + L_0^n}{\Delta x^2} - v \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} + (1 - A - Bx)U_j^n \end{aligned}$$

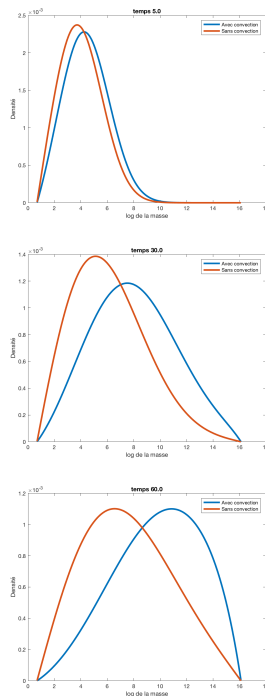
avec  $L0$  une matrice trigonale telle que :

$$L0 = \begin{pmatrix} -2 & 1 & \cdots & 0 \\ \vdots & 1 & -2 & 1 \\ 0 & \cdots & 1 & -2 \end{pmatrix}$$

Afin de modéliser les deux types de conditions aux bords nous n'omettrons pas de modifier les coefficient  $L0(1,1)$  et  $L0(J,J)$

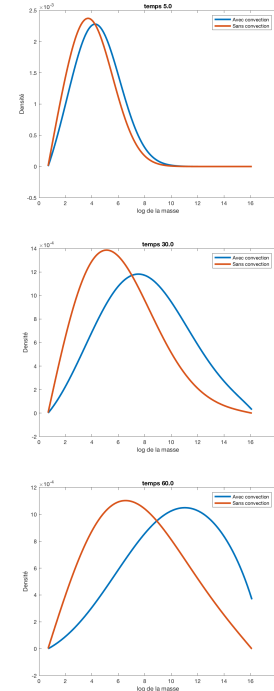
### Résultats et interprétation

Pour modéliser l'évolution des espèces de mammifères durant le crétacé nous avons avec convection nous avons conservé les paramètres cités ci-dessus pour A, B et v. Nous avons choisi un poids minimal de 2g étant la condition au bord Dirichlet et un poids maximal de  $10^7$ g représentant une condition de Dirichlet ou de Neumann selon la simulation. L'espace a été discrétisé avec un pas de 0.01. Enfin nous avons effectué une simulation sur 60 millions d'années période recouvrant les données [5].



**Figure 3.** Densité des poids en fonction de log de la masse sans convection (rouge) avec convection (bleu) avec les conditions de Dirichlet-Dirichlet : Graphique 1) t=5M années , Graphique 2) t=30M années, Graphique 3) t=60M années

Au regard des résultats, le terme de convection permet de générer un biais sur la diffusion. Au lieu d'observer une simple dispersion non orientée de la masse au cours de la cladogenèse, l'ajout du terme de convection permet de modéliser la règle de Cope. En effet celle-ci permet d'introduire le fait que l'augmentation de la masse augmente la fitness des espèces. La masse permettrait ainsi une meilleure thermorégulation,



**Figure 4.** Densité des poids en fonction de log de la masse sans convection (rouge) avec convection (bleu) avec les conditions de Dirichlet-Neumann : Graphique 1) t=5M années , Graphique 2) t=30M années, Graphique 3) t=60M années

ou encore de faciliter l'accès aux ressources. Or si la convection permet de représenter cet effet, pourrait-on imaginer une augmentation de masse jusqu'à l'infini? Étant donné que les hypothèses physiologiques ne contraignent le modèle dans ce sens, nous avons choisi de fixer une condition au bord supérieure très supérieure à la masse espèces mammifères fossiles et actuelles connue avec un poids maximal de 100 tonnes.

Néanmoins les hypothèses sous-jacentes d'une modélisation avec une condition de Dirichlet ou de Neumann différent. En effet la condition de Dirichlet implique qu'une espèce de masse supérieure à la plus volumineuse ne peut exister, tandis qu'une condition de Neumann implique qu'elle ne peut engendrer une espèce plus grande. Toutefois l'une des caractéristique du modèle est que la masse maximale fixée par la condition au bord droit n'est pas atteinte à la fin de la simulation, ce qui suggère une autre contrainte à la borne supérieure. Celle-ci peut être expliquée par la compromis entre le gain de fitness permis par une augmentation de la masse et l'augmentation du risque d'extinction corrélé avec la masse, qui est contrôlé par la constante B de l'équation.

### Schéma individu-centré

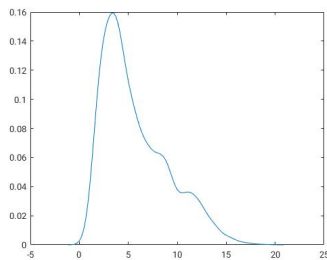
Ce schéma prend en sorte la convection.

La convection est prise en compte sous forme d'un paramètre



de "biais", que nous allons utiliser pour appliquer la règle de Cope : le biais est plus grand (et fonction de la masse) pour les petites masses, et uniforme et inférieur pour les plus grandes masses. "Petites masses" correspond aux masses inférieures à un paramètre qui est le logarithme de l'intercept de la taille. Ce paramètre, en particulier, sert à modéliser le premier "virage" qu'on observe sur la courbe de la répartition des masses (il y en a un deuxième vers les grandes masses). Ces valeurs semblent justifier la règle de Cope, qui fait l'hypothèse que plus des animaux sont petits, plus ils ont tendance à grandir lors de l'évolution.

On peut observer ce virage sur cette courbe :

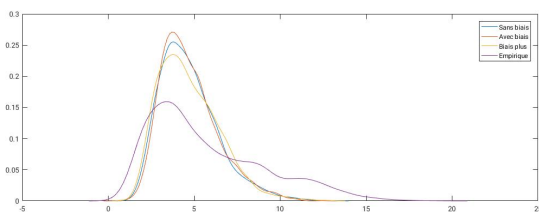


**Figure 5.** Répartition empirique de la masse d'espèces mammifères du crétacé

Ce biais est pris en compte lors de l'étape de cladogénèse : en effet, à chaque tour, l'espèce (qui fait office d'individu dans le modèle), stocké sous forme de la masse moyenne des individus dans un vecteur, a une probabilité d'évoluer. Son évolution sera modélisée comme la disparition de l'espèce et l'apparition de deux autres espèces, différentes. Leurs masses dépendront de la masse de l'individu(espèce) mère. En effet, les masses des nouvelles espèces ne peuvent pas être trop éloignées de l'espèce mère.

Les espèces ont aussi, à chaque tour, une probabilité d'extinction, mais elle n'est pas fonction de la masse.

On a fait la simulation avec plusieurs paramètres : sans biais (donc sans paramètre de convection), avec un biais pour modéliser le premier virage, et avec un deuxième paramètre de biais pour tenter de modéliser le deuxième.



On peut voir que les différences entre les distributions des logarithmes des masses sont plus importantes entre les modélisations qu'avec les données empiriques. Le paramètre de convection, notamment, ajoute une différence dans la répartition.

Le deuxième virage a été modélisé sur le même modèle que

le premier : En utilisant un palier de la même forme, en choisissant un biais sur le même modèle... Mais il aurait pu être amélioré pour mieux coller à l'allure des données empiriques, même si c'est celui qui est le plus proche. On peut donc se pencher sur cette question et suggérer que la raison de ce virage n'a rien à voir avec la règle de Cope, mais à d'autres paramètres que le modèle ne prend pas en compte. De plus, le paramètre de seuil  $c(3)$  n'a pas été choisi de manière très rigoureuse, mais en approximant à partir des données empiriques. Mathématiquement, quand on compare les distributions pour voir laquelle est la plus proche des données empiriques, on obtient, en comparant les moyennes :

- sans biais : 4.8493 ;
- avec biais : 4.6590 ;
- biais plus : 4.6230 ;
- empirique : 5.7215.

Ce qui ne semble pas cohérent avec les premières approximations que l'on peut faire sur le sujet. On voit bien que l'asymétrie de la répartition des données joue beaucoup. De plus, le deuxième virage a un rôle important au niveau de la répartition des masses : c'est un phénomène qui ne semble pas négligeable dans la modélisation.

## Conclusion

Si le modèle l'implémentation numérique fonctionne, nous ne sommes pas parvenu à ajuster l'échelle de la modélisation aux données. Nous devrions retravailler sur l'estimation des paramètres biologiques déduit par un test de Komolgorov Smirnov à partir des données. Nous devrions également vérifier les équations numériques. Enfin si nous sommes parvenu à reconstituer la forme de la distribution des masses des espèces de mammifères au crétacé grâce à l'implémentation de deux biais sur l'évolution des masses, nous pouvons toujours questionner la pertinence de la règle de Cope. En effet d'une part une approximation aussi fine de la masse d'espèces fossiles est complexe, et d'autre part dans le jeu de données rien ne nous permet d'établir des liens de parenté. Ainsi cette modélisation représente une théorie biologique mais ne la démontre pas.

## References

- [1] Steven M Stanley. An explanation for cope's rule. *Evolution*, 27(1):1–26, 1973.
- [2] BG Lovegrove and MO Mowoe. The evolution of mammal body sizes: responses to cenozoic climate change in north american mammals. *Journal of evolutionary biology*, 26(6):1317–1329, 2013.
- [3] Aaron Clauset, David J Schwab, and Sidney Redner. How many species have mass m? *The American Naturalist*, 173(2):256–263, 2008.

- [4] Aaron Clauset. How large should whales be? *PLoS One*, 8(1):e53967, 2013.
- [5] Aaron Clauset and Douglas H Erwin. The evolution and distribution of species body size. *Science*, 321(5887):399–401, 2008.

-