

La biogéographie est la science qui s'interroge sur les causes de la répartition de la biodiversité dans les différentes parties du globe. Le modèle déterministe de MacArthur et Wilson décrit l'évolution de la biodiversité sur les îles vers un équilibre, mais la migration des espèces vers des îles et leur extinction potentielle sont des phénomènes aléatoires. Comment un modèle déterministe peut-il refléter ce hasard ?

La Biodiversité en territoires isolés

Kévin Cazelles
UQAR

Un des modèles les plus puissants en biogéographie est celui proposé par MacArthur et Wilson dans leur passionnante théorie de la biogéographie des îles. Ces deux illustres pionniers nous ont livré un paradigme simple et puissant pour envisager la construction de la biodiversité sur une île. Notons dès maintenant que l'île n'est pas nécessairement un monticule de sable fin au milieu de l'océan mais plutôt – et plus généralement – un territoire isolé. Le modèle permet de comprendre l'impact des capacités de dispersion et de survie des espèces sur la biodiversité de l'île étudiée. Considérons que cette île est accessible depuis un continent qui présente un nombre constant d'espèces P . Les P espèces sont celles que nous pourrions retrouver sur l'île, posons S le nombre d'espèces sur l'île, nous avons alors $S \leq P$. À un temps t donné, les S espèces de l'île

peuvent s'éteindre avec un taux e . De même, les $P - S$ espèces du continent (absentes de l'île) peuvent coloniser l'île avec un taux c . En langage mathématique, nous obtenons l'équation différentielle déterministe suivante pour décrire l'évolution temporelle de S :

$$\begin{aligned} \frac{dS}{dt} &= c(P - S) - eS \\ &= cP - (c + e)S. \end{aligned}$$

Cette équation est linéaire et non homogène, nous pouvons la résoudre facilement par la méthode de la variation de la constante (voir l'encadré *Solution déterministe*). Pour un temps infini, la biodiversité tend vers la valeur :

$$S_{eq} = P \left(\frac{c}{c + e} \right),$$

interprétée comme le nombre d'espèces provenant du continent et présentes sur l'île à

La biogéographie

La biogéographie est la science qui s'interroge sur les causes de l'agencement spatial des espèces à la surface de notre planète et cherche à expliquer, par exemple, la répartition et les différences entre les grands biomes terrestres. Le premier élément de réponse est l'implication des facteurs climatiques. La température, l'humidité, les précipitations sont autant de variables qui limitent les conditions d'existence des espèces qui composent et modèlent les paysages terrestres. C'est avec ces contraintes que les écologues élaborent des modèles dits « bioclimatiques » afin de décrire l'évolution de la biodiversité avec

les contraintes climatiques de demain. Cependant, les facteurs qui expliquent la distribution des espèces ne sont pas toujours climatiques. Les mouvements des espèces, leurs interactions et leurs histoires évolutives sont ainsi des moteurs fondamentaux en biogéographie. Les capacités de dispersion des espèces leur donnent accès à de nouveaux territoires qui, avec le temps peut-être, les verront disparaître. C'est ainsi qu'un territoire voit sa composition en espèces évoluer, avec l'arrivée de nouvelles espèces et l'extinction d'autres. C'est autour de cette idée majeure que s'articule le modèle présenté.



l'équilibre. Pour les valeurs arbitraires $c = 0,2$, $e = 0,1$ et la condition initiale suivante : à $t = 0$, l'île est inhabitée ($S(0) = 0$), nous représentons graphiquement la solution de l'équation dans l'encadré jaune ci-dessous.

Comment un modèle déterministe émerge-t-il de phénomènes aléatoires?

Modèle stochastique

Il est très intéressant de réaliser que l'équation différentielle à la page précédente, d'apparence déterministe, renferme un modèle stochastique. « Stochastique » est un terme signifiant aléatoire, avec une part de hasard. Deux réalisations d'un modèle stochastique ne donneront donc pas nécessairement le même résultat. Pour comprendre où se cache le hasard dans le modèle que nous étudions, nous avons besoin d'objets mathématiques particuliers appartenant au domaine des probabilités : variables aléatoires et processus stochastiques (voir l'encadré Variables aléatoires).

Les objets mathématiques requis

Pour modéliser l'évolution de la biodiversité de l'île, nous gagnons à décrire l'évolution de la présence individuelle de chacune des espèces considérées. Afin d'éviter d'entrer dans le détail des interactions entre les espèces et de l'effet de la taille des populations, nous comptons simplement 1 lorsque l'espèce est sur l'île, 0 sinon. Nous introduisons donc X_i , la variable aléatoire de présence sur l'île de l'espèce i choisie parmi les P

Solution déterministe

Il est facile de vérifier que la solution générale de l'équation homogène :

$$\frac{dS}{dt} = -(c+e)S$$

est $S(t) = C \exp(-(c+e)t), \forall t \in \mathbb{R}^+$, où C est une constante. La solution de l'équation non homogène est alors de la forme :

$$S(t) = f(t)\exp(-(c+e)t).$$

En remplaçant dans l'équation différentielle,

$$f'(t)\exp(-(c+e)t) = cP,$$

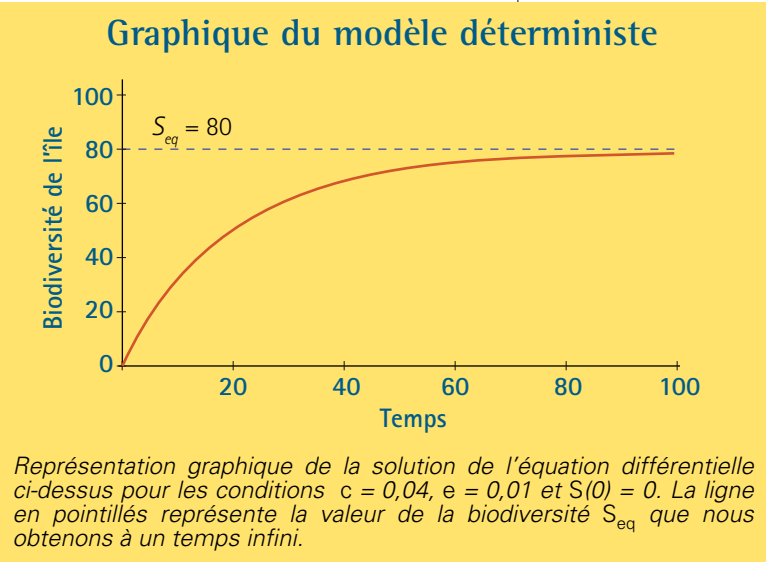
ou encore,

$$f'(t) = cP\exp((c+e)t).$$

En intégrant,

$$S(t) = P \frac{c}{c+e} + K \exp(-(c+e)t),$$

où K est une constante arbitraire. Avec $S(0) = 0$, $K = -P \frac{c}{c+e}$, nous obtenons

$$S(t) = P \frac{c}{c+e} [1 - \exp(-(c+e)t)].$$


espèces du continent. X_i est égale à 0 si l'espèce n'est pas sur l'île et égale à 1 si elle est sur l'île. C'est une variable aléatoire de Bernoulli (voir encadré Loi binomiale). Ainsi, le nombre S d'espèces présentes sur l'île est égale à la somme des X_i . Nous enregistrons ces valeurs au cours du temps pour définir le processus stochastique $X_{i,t} > 0$, qui n'est autre qu'une succession de 1 et de 0 indiquant à chaque instant si oui ou non l'espèce i est sur l'île. A priori, les suites $X_{i,t}$ ne sont pas

prévisibles, ce qui n'exclut pas que la fonction S présente, en moyenne, un comportement tout à fait régulier.

Les briques élémentaires du modèle stochastique

Au temps $t = 0$, l'espèce i n'est pas sur l'île. Pour construire la suite de son histoire, il nous faut une règle simple pour décrire l'évolution de sa présence entre deux pas de temps très proches. Le modèle de MacArthur et Wilson postule que si l'espèce considérée est sur l'île au temps t , elle s'éteint avec un taux e ; si elle est sur le continent, elle le colonise avec un taux c . Pour simplifier, nous supposons que les taux d'extinction et de colonisation sont les mêmes pour toutes les espèces. Nous donnons maintenant une signification en terme de probabilité à ces taux : c est la probabilité de colonisation par unité de temps (on parle de densité du processus), e est la probabilité d'extinction par unité de temps. Ainsi, edt désigne la probabilité d'extinction pendant l'intervalle de temps dt . De même, cdt est la probabilité de colonisation

pendant l'intervalle de temps dt . En faisant appel aux probabilités conditionnelles, en supposant dt assez petit, nous distinguons quatre cas, voir ci-contre.

Quatre cas distincts

$\forall t \in \mathbb{R}^+$ et $dt > 0$,

$$P(X_{i,t+dt} = 1 \mid X_{i,t} = 0) = cdt$$

$$P(X_{i,t+dt} = 0 \mid X_{i,t} = 1) = edt$$

$$P(X_{i,t+dt} = 0 \mid X_{i,t} = 0) = 1 - cdt$$

$$P(X_{i,t+dt} = 1 \mid X_{i,t} = 1) = 1 - edt$$

Il faut interpréter la première équation ainsi : sachant que l'espèce i était absente au temps t , la probabilité qu'elle soit sur l'île au temps $t + dt$ est égale à la probabilité qu'elle colonise l'île durant l'intervalle dt , c'est-à-dire cdt . Nous n'avons pas mentionné que durant dt , une espèce peut coloniser, s'éteindre et re-coloniser. Seulement, tous ces événements sont en fait beaucoup moins probables et nous pouvons les ignorer complètement à la limite quand dt tend vers 0. Les trois autres équations s'interprètent avec un raisonnement similaire. Ces quatre équations sont les briques du modèle stochastique que nous construisons. Nous devons maintenant les assembler correctement.

Assemblons les briques !

L'assemblage demande de considérer les différents cas disjoints possibles avec leurs probabilités associées (voir l'encadré *Probabilité conditionnelle et formule des probabilités totales*). Nous pouvons en déduire la probabilité de présence de l'espèce i à l'instant $t + dt$:

$$P(X_{i,t+dt} = 1) = P(X_{i,t+dt} = 1 \mid X_{i,t} = 0)P(X_{i,t} = 0) + P(X_{i,t+dt} = 1 \mid X_{i,t} = 1)P(X_{i,t} = 1)$$

Variable aléatoire

Les variables aléatoires sont définies sur l'ensemble des résultats possibles d'un événement aléatoire. On associe à chaque valeur de la variable sa probabilité, de sorte que la somme des probabilités sur l'ensemble des valeurs possibles soit égale à 1. Dans notre exemple, X_i est une variable aléatoire, 0 et 1 sont ses valeurs, et nous nous intéressons à leur probabilité $P(X_i = 1)$ et $P(X_i = 0)$ avec $P(X_i = 1) + P(X_i = 0) = 1$. Les processus aléatoires ou stochastiques sont des collections de variables aléatoires indexées par le temps t , c'est-à-dire ordonnées. Nous suivons ainsi l'évolution de X_i au cours du temps d'où la notation $X_{i,t>0}$. Attention ! $X_{i,t>0}$ est un processus aléatoire mais $X_{i,t}$ est l'une des variables aléatoires du processus.

Loi binomiale

Une variable aléatoire X suit un schéma de Bernoulli lorsqu'elle prend uniquement deux valeurs 0 ou 1. Si nous notons $p = P(X = 1)$, nous avons

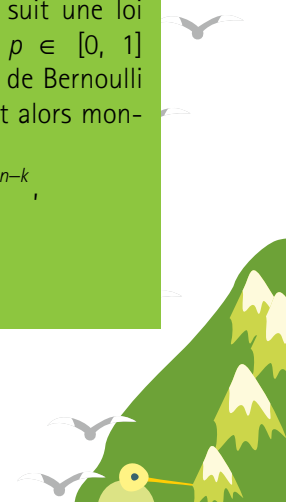
$$P(X = 0) = 1 - p, E(X) = p \text{ et } V(X) = p(1 - p).$$

X est alors une variable aléatoire de Bernoulli de paramètre p . Une variable aléatoire X suit une loi Binomiale de paramètres $n \in \mathbb{N}$ et $p \in [0, 1]$ lorsqu'elle est la somme de n variables de Bernoulli indépendantes de paramètre p . On peut alors montrer que

$$\forall k \leq n, P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k},$$

$E(x) = np$ est l'espérance de X

et $V(X) = np(1 - p)$, sa variance.



Il s'agit d'une somme de deux termes couvrant toutes les possibilités puisque, à l'instant t , l'espèce i était, soit absente de l'île, soit présente : c'est un système complet d'événements. Si l'espèce i était absente (premier terme) à t , elle sera présente à $t + dt$ si elle colonise pendant l'intervalle de temps dt (premier des quatre cas distincts). Si elle était déjà sur l'île au temps t , elle s'y maintient à condition de ne pas s'éteindre (quatrième des quatre cas distincts).

En remarquant que $P(X_{i,t+dt} = 0) = 1 - P(X_{i,t+dt} = 1)$, nous pouvons écrire :

$$P(X_{i,t+dt} = 1) = cdt[1 - P(X_{i,t} = 1)] + (1 - edt)P(X_{i,t} = 1).$$

Pour simplifier l'écriture, on représente $P(X_{i,t} = 1)$ par $s_{i,t}$, on a alors :

$$\begin{aligned} s_{i,t+dt} &= cdt(1 - s_{i,t}) + (1 - edt)s_{i,t} \\ &= cdt(1 - s_{i,t}) + s_{i,t} - edts_{i,t}. \end{aligned}$$

Réarrangeant et divisant les deux membres par dt ,

$$\frac{s_{i,t+dt} - s_{i,t}}{dt} = c(1 - s_{i,t}) - es_{i,t}.$$

En faisant alors tendre dt vers 0, $\frac{ds_i}{dt} = c(1 - s_i) - es_i$.

Nous retrouvons alors notre équation de départ pour l'évolution de S , avec $P = 1$. Sa solution est donc

$$s_i(t) = \frac{c}{c+e} [1 - \exp(-(c+e)t)] = s(t).$$

Retrouvons le modèle classique

Nous avons presque retrouvé la solution de l'équation déterministe ; suffirait-il de multiplier par le nombre d'espèces P ? L'idée est bonne, mais demande justification ! Nous allons effectivement considérer non pas une, mais bien les P espèces du continent pour lesquelles l'équation déterministe est valable. Nous allons alors définir un nouveau processus stochastique qui est simplement la somme des P processus de présence que nous supposons indépendants $X_{i,t>0}$,

$$Y_{t>0} = X_{1,t>0} + X_{2,t>0} + \dots + X_{P,t}.$$

Le problème est alors de connaître à tout instant t la probabilité d'avoir un nombre donné k d'espèces présentes sur l'île. À un instant donné t nous faisons une somme de variable aléatoires de Bernoulli. La somme de P variables de Bernoulli indépendantes est une variable aléatoire suivant une loi binomiale. Nous avons donc :

$$\forall t \in \mathbb{R}^+, P(Y_t = k) = \binom{P}{k} s(t)^k (1 - s(t))^{P-k}.$$

De plus, on peut démontrer que

$$\forall t \in \mathbb{R}^+, E(Y_t) = Ps(t),$$

$$\forall t \in \mathbb{R}^+, V(Y_t) = Ps(t)(1 - s(t)).$$

Grâce à ces résultats de probabilité, on a

$$\forall t \in \mathbb{R}^+, E(Y_t) = P \frac{c}{c+e} [1 - \exp(-(c+e)t)].$$

Nous retrouvons la solution déterministe ! Ainsi, le modèle déterministe peut être considéré comme reflétant l'évolution de l'espérance des variables Y_t pour $t > 0$, c'est donc l'espérance du processus stochastique. Nous pouvons alors simuler le modèle stochastique grâce au modèle déterministe : à chaque instant t , il suffit de simuler une loi binomiale de paramètres P et $E(Y_t)$. Le graphique à la page suivante montre des différences fondamentales entre les deux approches. Le modèle déterministe donnera des valeurs continues et, pour des paramètres donnés, toujours le même résultat. De son côté, le modèle stochastique livrera des valeurs discrètes. De plus, deux simulations du modèle aléatoire ne donneront pas nécessairement les mêmes courbes.



Applications et perspectives

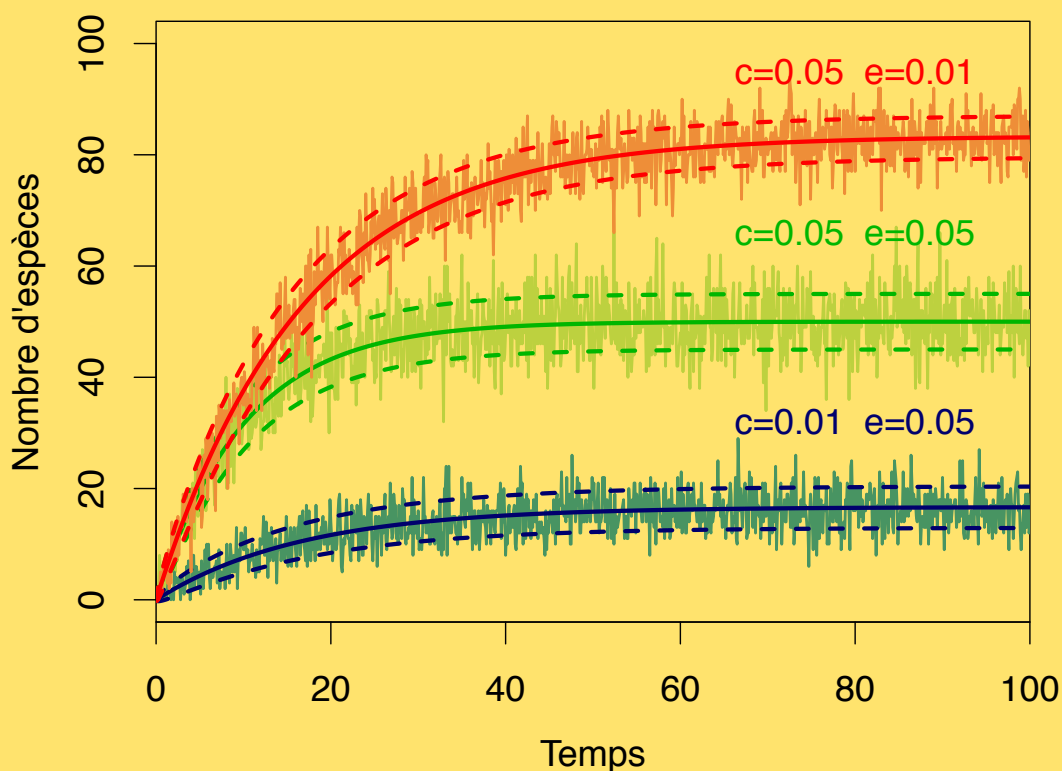
Bien que datant du milieu des années 1960, le modèle de MacArthur et Wilson demeure très utile, notamment pour étudier les habitats fragmentés (des îles !) et prendre des décisions relatives à la conservation des espèces. Prenons un exemple : une nouvelle aire protégée a récemment été mise en place que nous assimilerons à une île. Cette zone a longtemps été exploitée par l'homme de sorte qu'au temps $t = 0$ le nombre d'espèces est $S = 0$. Dans la région, la richesse en espèces est de $P = 100$. À l'instar de MacArthur et Wilson, nous allons supposer que la valeur du taux de colonisation c dépend de la difficulté d'accès. De plus, la survie d'une espèce sur cette zone dépend de la superficie de la zone : ainsi plus l'aire de conservation est grande, moins les

espèces s'éteignent ; c'est donc plus faible pour les grandes aires de protection. Prenons trois situations :

- Grande zone et accès aisé : $e=0,01$ et $c=0,05$
- Petite zone et accès aisé : $e=0,05$ et $c=0,05$
- Petite zone et accès difficile : $e=0,05$ et $c=0,01$.

Pour mettre en place un suivi et savoir comment se comporte la réserve, il faut avoir une idée de la dynamique attendue ! Ce que nous permet le modèle étudié, sous réserve de valider les hypothèses utilisées. Nous simulons pour les trois situations, avec les deux approches du modèle. Les résultats

Dynamique de la biodiversité



Dynamique de la biodiversité de trois zones protégées. Les caractéristiques de la zone protégée sont données par les valeurs de c et e . Les symboles rouges sont relatifs à une zone protégée de grande taille et facilement accessible, en vert, ils font référence à une petite zone facilement accessible ; enfin, en bleu, sont présentés les résultats pour une petite zone difficilement accessible. Les courbes striées en couleurs pastels sont les réalisations du modèle stochastique. Les courbes en traits pleins, font référence au modèle déterministe qui est également la moyenne du modèle stochastique. Enfin, les courbes en pointillés représentent la moyenne du modèle stochastique plus ou moins l'écart type.

sont donnés par la figure dynamique de la biodiversité. Sans chercher à rendre compte de la distribution des espèces et de leurs interactions, ce modèle très simple permet néanmoins d'apprécier l'effet de la taille de la zone protégée (à travers le paramètre e) et celui de son accessibilité (à travers le paramètre c). Le modèle stochastique permet d'étudier la variance associée aux différents scénarios. De plus, l'approche stochastique révèle toute sa force pour enrichir le modèle. Ici, nous avons fait plusieurs hypothèses simplificatrices. Notamment, nous avons considéré des taux de colonisation et

d'extinction constants. Cette supposition peut être pertinente lorsque les espèces sont biologiquement proches, par exemple, en considérant seulement des amphibiens. Mais si nous étudions des espèces aux capacités de survie et de dispersion très différentes, les taux du modèle dépendent alors des caractéristiques du territoire et des groupes d'espèces. De plus, les effets de populations et leurs interactions sont totalement occultés. Pour aller plus loin, nous travaillons sur les probabilités associées aux différents états de l'île et bâtissons des approches plus complexes certes, mais aussi plus réalistes.

Probabilité conditionnelle et formule des probabilités totales

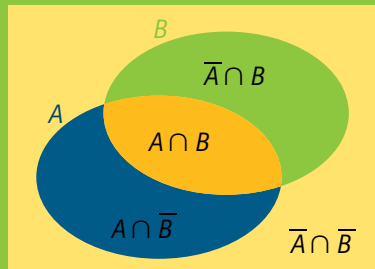
Une probabilité conditionnelle, est la probabilité qu'un événement se réalise à condition qu'un autre soit déjà réalisé. Soit deux événements A et B , la probabilité de A sachant B , notée ici $P(A|B)$, est la probabilité que A se réalise lorsque B est réalisé,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Un système complet d'événements est un ensemble d'événements exclusifs dont la somme des probabilités est 1. Par exemple, l'événement B et son complémentaire \bar{B} forment un tel système.

La formule des probabilités totales est le calcul de la probabilité d'un événement A à partir de la connaissance de sa réalisation conjointe avec un système complet d'événements (B et \bar{B} , par exemple) que nous pouvons également formuler en utilisant les probabilités conditionnelles. Ainsi, pour connaître la probabilité de A , nous pouvons écrire :

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$



Représentation de la formule des probabilités totales. A et B sont deux événements, la probabilité de A peut être décrite comme la somme des probabilités de ses intersections avec B et \bar{B} qui constituent un système complet d'événements.

