



La biodiversité mise en équations... ou presque

Sylvie Méléard, *professeur à l'École polytechnique*

Prédire l'évolution d'une population animale sur une longue période, connaître le fonctionnement d'un écosystème, comprendre l'avantage de la reproduction sexuée pour la survie des espèces... les problèmes issus de la biodiversité sont complexes et leur résolution fait appel à des outils mathématiques sophistiqués.

“After years, I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics: for men thus endowed seem to have an extra-sense”.

(Les années passant, j'ai profondément regretté de ne pas être allé assez loin pour comprendre quelques aspects des principes essentiels des mathématiques: car les hommes dotés d'un tel entendement semblent posséder un sixième sens.)

*Charles Darwin,
Autobiography*

Il peut sembler étonnant d'associer les mathématiques, vues souvent comme une succession de formules figées, et la biodiversité, associée à la vie fourmillante des écosystèmes. Pourtant, la complexité extrême qui apparaît dès que l'on se penche sur la biodiversité justifie l'intérêt des mathématiques et des mathématiciens pour ce domaine. L'enjeu est de simplifier intelligemment les phénomènes pour pouvoir les mettre en équations et construire ainsi des modèles mathématiques novateurs. Le but est de pouvoir en déduire des quantités calculables telles que, par exemple, le temps d'extinction d'une espèce, l'estimation d'une abondance d'espèce, la vitesse d'invasion d'un parasite... ou encore de développer des algorithmes de simulation qui sont des outils d'expérimentation théo-

rique indispensables pour prédire certains comportements. C'est vrai en particulier pour les très grandes échelles de temps de l'évolution, pour lesquelles les observations n'existent pas. Les simulations sont aussi un support visuel qui facilite les échanges scientifiques avec les biologistes. La modélisation peut enfin servir à construire des outils statistiques permettant de prédire et quantifier différents scénarios de la biodiversité au vu des données observées.

Au xix^e et début du xx^e siècle, de nombreux modèles, principalement déterministes, c'est-à-dire pour lesquels une même cause provoque toujours le même effet, ont été développés pour étudier la dynamique des populations (Malthus, Verhulst, Lotka, Volterra...). D'autres modèles plus probabilistes, c'est-à-dire faisant intervenir le hasard, ont également été proposés, pour étudier la génétique des populations (Fisher, Haldane, Wright...). Dans les dernières décennies, les biologistes se sont largement concentrés sur la biologie moléculaire et sur la formidable masse de données que les nouvelles technologies leur ont permis d'obtenir. Les méthodes de séquençage ont apporté beaucoup d'information qu'il a fallu trier, classifier, ordonner et analyser. Forts de ces données, les biologistes ressentent de nouveau la nécessité de créer des modèles pour leur donner une cohérence et en déduire des prédictions sur les comportements biologiques et écologiques.

La complexité des problèmes issus de la biodiversité nécessite donc des outils mathématiques sophistiqués. Réciproquement, les mathématiciens vont puiser dans

ces problèmes une motivation et une inspiration pour créer de nouveaux modèles et théorèmes d'une grande richesse, qui permettront de comprendre certains phénomènes écologiques.

Quelques problématiques d'écologie

Voici quelques exemples de questions posées par les écologues et auxquelles les mathématiciens s'efforcent de donner des éléments de réponses :

- Quel est le comportement d'une population en temps long ? Va-t-elle se stabiliser ou s'éteindre ? Si elle s'éteint, quel sera le temps d'extinction ?
- Les espèces cohabitent et interagissent dans une compétition pour les ressources, dans un réseau hôte-parasite ou une chaîne alimentaire, ou au contraire coopèrent pour leur survie. Comment ces réseaux évoluent-ils ?
- Quel est l'impact de la migration et de la fragmentation de l'habitat sur la biodiversité ?
- Quel est l'avantage de la reproduction sexuée pour la biodiversité et la survie des espèces ? En effet, même si la reproduction sexuée apporte de la variation génétique par recombinaison, nourrir les mâles, rechercher des partenaires, s'exposer à la prédation pendant l'acte sexuel sont des arguments qui rendent plus favorable la reproduction asexuée.



- Quel est l'impact des variations de l'environnement sur la biodiversité ?
- Quel est le lien entre l'ADN des populations, leurs interactions et leur évolution ?

La notion d'espèce et les modèles probabilistes

Pour comprendre la biodiversité, il faut commencer par comprendre ce qu'est une espèce.

La notion d'espèce a évolué au cours du temps. Déjà, Maupertuis (1698-1759), mathématicien et naturaliste, soulignait l'apparition au hasard de mutants qui se

développent et forment une nouvelle population: « Nous voyons paraître des races de chiens, de pigeons, de serins qui n'étaient point auparavant dans la nature. Ce n'ont été d'abord que des individus fortuits; l'art et les générations répétées en ont fait des espèces ». Dans l'œuvre de Lamarck (1744-1829), l'image linéaire de la grande chaîne des êtres est remplacée par un « arbre buissonnant ». Finalement, dans son fameux ouvrage *On the origin of species by natural selection*, Darwin (1809-1882) introduit l'idée de sélection naturelle et d'arbre des espèces, avec disparition de certaines espèces par sélection des plus adaptées, et apparition de nouvelles, telles les différentes espèces de pinsons qu'il met en évidence sur les îles Galapagos.

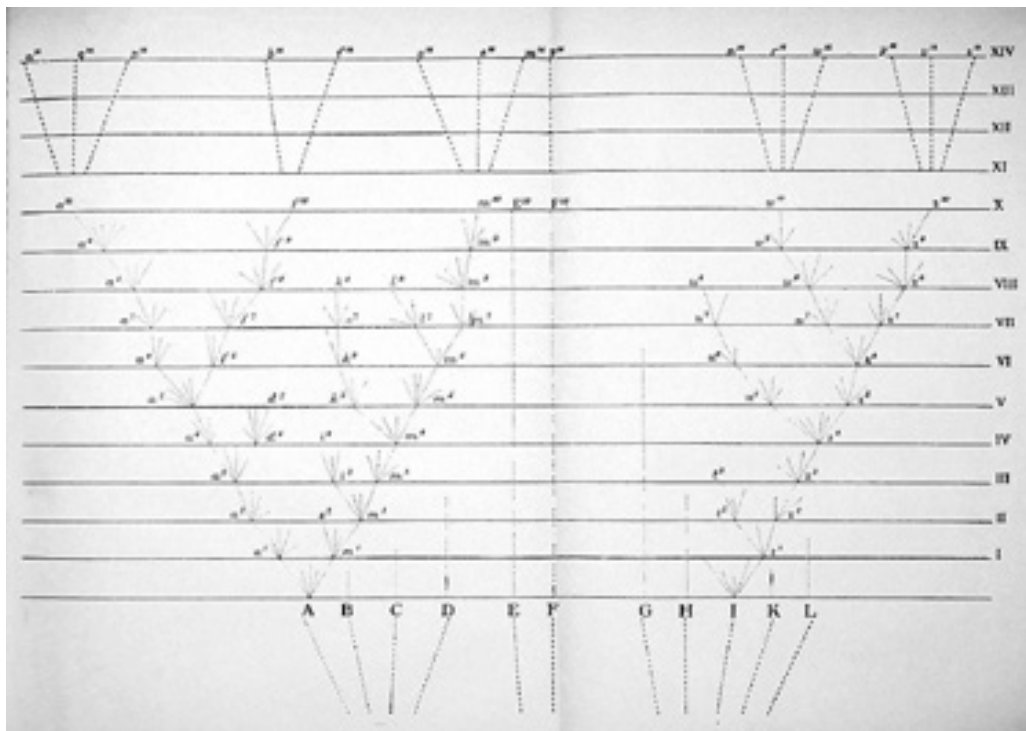


Figure 1. Arbre des espèces, avec disparition de certaines espèces et apparition de nouvelles, d'après C. Darwin, *On the origin of species by means of natural selection*.



C'est cette dynamique de la biodiversité que l'on souhaite mieux appréhender, grâce à des modèles mathématiques et en particulier à des modèles probabilistes qui prennent en compte les naissances et les morts aléatoires des individus et que l'on appelle de manière générique *processus de branchement* ou *processus de naissance et mort*. Le prototype en est le processus de Galton-Watson (voir encadré). Il peut permettre de calculer la probabilité d'extinction d'une espèce.

Une démarche inverse, que le séquençage de l'ADN a beaucoup stimulée, consiste à reconstruire l'histoire de vie d'une population. Étant donné un groupe d'individus, à quelle époque se situait leur plus récent ancêtre commun? Comment retrouver les lignées ancestrales des individus? Cette démarche est celle des généticiens des populations et a conduit au modèle de Wright-Fisher (voir encadré).

Cette dualité entre la dynamique des populations au cours du temps et la reconstruction des généalogies permet ainsi d'estimer la biodiversité passée en fonction des relations de parenté des espèces actuelles.

Des modèles construits sur les comportements individuels et leurs interactions écologiques

Les modèles cités précédemment sont mathématiquement très simples. Même si les processus de Galton-Watson et les modèles de Wright-Fisher jouent un rôle

fondamental – leur simplicité permettant de faire des calculs et de donner certaines réponses quantitatives – ils ne permettent pas de prendre en compte la diversité génétique et les interactions écologiques. Aussi, les recherches actuelles des biologistes théoriciens et des mathématiciens ont pour but d'introduire de la diversité génétique (ADN) et de l'interaction (compétition, coopération, prédation) dans les paramètres démographiques des modèles.

Les recherches actuelles ont pour but d'introduire de la diversité génétique et de l'interaction.

Les modèles les plus anciens introduisant de la compétition sont des modèles déterministes décrivant le comportement de grandes populations. Le plus célèbre est l'*équation logistique*, qui décrit la dynamique de la taille d'une population :

$$n'(t) = n(t)(b - cn(t))$$

Dans cette équation, le paramètre b décrit le taux de croissance de la population et c mesure la compétition entre deux individus pour le partage des ressources. Pour des temps t très grands, cette taille se stabilise sur la valeur limite b/c , appelée *capacité de charge*. Si la population comporte des individus de plusieurs types, elle est alors décrite par un système d'équations. Dans le modèle de proies-prédateurs par exemple, si $n_1(t)$ et $n_2(t)$ sont le nombre de proies et de prédateurs au temps t , alors on peut écrire les équations :



$$n_1'(t) = a n_1(t) - b n_1(t) n_2(t)$$

$$n_2'(t) = -c n_2(t) + d n_1(t) n_2(t)$$

On observe alors un comportement cyclique de ces dynamiques.

Dans le cas d'un continuum de types (la taille à la naissance, l'âge à la maturité...), la dynamique d'une grande population sera décrite par des équations plus compliquées (équations aux dérivées partielles ou intégro-différentielles) reliant les types des individus et le temps.

Ces modèles déterministes sont des modèles macroscopiques (c'est-à-dire qu'ils considèrent le comportement global d'un ensemble d'individus) et ne prennent pas en compte les comportements individuels, ni les naissances aléatoires d'individus mutants. Ils ne sont pas représentatifs des comportements des petites populations pour lesquels les fluctuations aléatoires sont essentielles. Ils ne peuvent pas non plus rendre compte de toutes les échelles de temps auxquelles l'on peut être confronté : par exemple le lien entre l'échelle écologique (les temps de naissance et de mort des individus) et celle de l'évolution (apparition et fixation des mutations). Pour pouvoir intégrer ces événements au modèle, il est important de revenir à une description du comportement aléatoire de chaque individu en prenant en compte ses propres caractéristiques. Les modèles probabilistes sont alors des généralisations du processus de Galton-Watson.

Exemple d'une population caractérisée par la taille des individus à la naissance

Dans cet exemple, les individus sont caractérisés par leur taille x à la naissance, qui varie par exemple entre 0 et 4. Le temps qu'un individu de taille x met à se reproduire est une variable aléatoire de moyenne $b(x)$. On peut supposer par exemple que $b(x) = 1/(4 - x)$, ce qui traduit le fait qu'un grand individu dépense beaucoup d'énergie à se nourrir et en a moins pour se reproduire.

A chaque reproduction, l'individu transmet héréditairement son type x , mais il arrive que son descendant mute en un individu de type y . Ce dernier va se reproduire et créer ensuite sa propre sous-population de type y . Les mutations introduisent ainsi de la variabilité génétique dans le modèle. Par ailleurs, la mort de l'individu dépend de ses propres paramètres génétiques, mais aussi de la compétition pour les ressources qu'il subit de la part de ses congénères. Par exemple, un individu de type $y > x$ consommera plus de ressources, au dépend de celui de trait x qui va ainsi voir ses possibilités de survie s'amoinrir.

La sélection naturelle, qui va permettre à la population de s'adapter et entraîner son évolution, est donc le résultat d'un compromis entre le fait de favoriser les petits individus pour accroître leurs capacités de reproduction et celui de favoriser les gros individus pour leur permettre d'être plus forts dans la compétition pour les ressources. C'est ce type d'optimisation évolutive que les mathématiciens cherchent à comprendre



ainsi que les mécanismes mathématiques qui peuvent expliquer l'apparition de nouvelles espèces.

Le modèle décrit ci-dessus peut expliquer la diversité des pinsons de Darwin à travers la taille de leur bec. La figure ci-dessous montre une simulation du modèle (voir figure 2) qui explique l'apparition de quatre

espèces à partir d'une espèce unique, dans une très longue échelle de temps.

Il est bien sûr totalement illusoire de vouloir réduire la biodiversité à de simples équations, mais les modèles mathématiques peuvent fournir un nouvel angle d'étude et un point de vue dépassionné et objectif des écosystèmes.

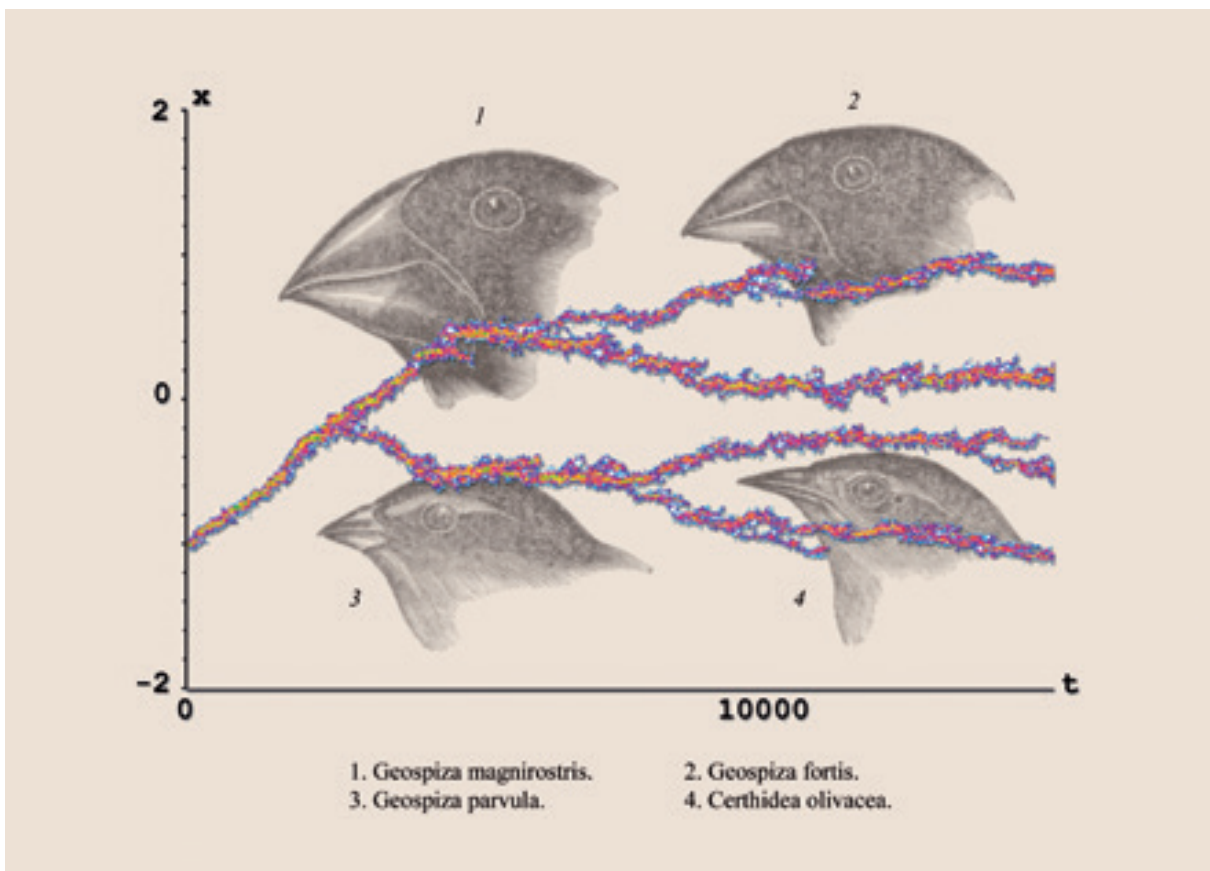


Figure 2. Simulation du modèle des pinsons de Darwin.



Le processus de Galton-Watson

Le processus de Galton-Watson est une suite de variables aléatoires (Z_n) décrivant l'évolution d'une population qui, à chaque génération, se reproduit aléatoirement. Si Z_n est la taille de la population à la $n^{\text{ième}}$ génération, alors $Z_{n+1} = X_1 + \dots + X_{Z_n}$, où X_i est le nombre d'enfants du $i^{\text{ème}}$ individu. Les variables X_i ont toutes la même loi X , et en particulier le nombre moyen m d'enfants par individu est constant. La probabilité d'extinction est alors obtenue comme limite d'une suite récurrente (dont le terme général donne la probabilité d'extinction à la génération n), et est

solution de l'équation $g(p) = p$, où $g(s)$ est la moyenne de s^X . On peut montrer que si m est un nombre supérieur à 1, la population peut exploser exponentiellement, alors que si m est inférieur ou égal à 1, la population va s'éteindre. Par exemple, sachant que le nombre moyen des petits de baleines noires est estimé à 0,976 et que l'abondance des femelles était estimée à 150 en 1994, on peut montrer par ce modèle qu'il y a 99 % de chances pour que ces baleines aient totalement disparu en 2389.

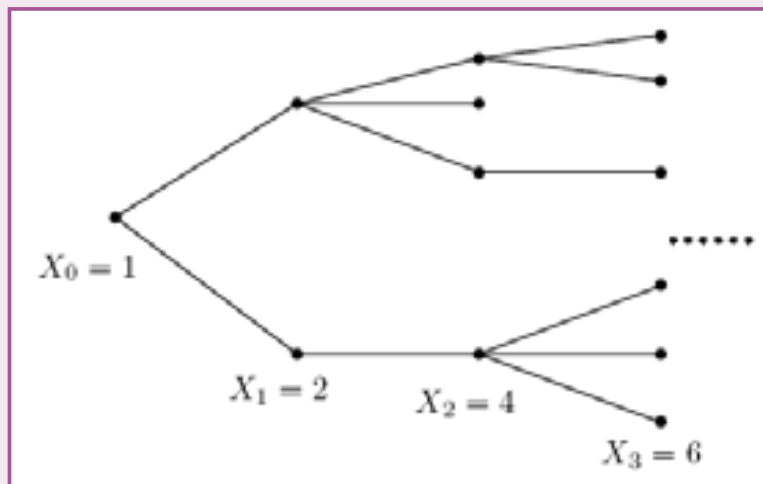


Figure 3. Une réalisation du processus de Galton-Watson.



Le modèle de Wright-Fisher

Dans ce modèle, la taille de la population N est constante et, à chaque génération, chaque individu choisit son ancêtre uniformément au hasard dans la génération précédente. Ainsi la probabilité que deux individus aient le même parent est $1/N$, qui tend vers 0 quand N tend vers l'infini. Il faut donc changer l'échelle de temps pour observer les généalogies pour une grande population et prendre des temps de génération proportionnels à N . Dans ce cas, on construit un objet limite appelé *coalescent de Kingman* qui est un objet géométrique probabiliste décrivant les lignées ancestrales.

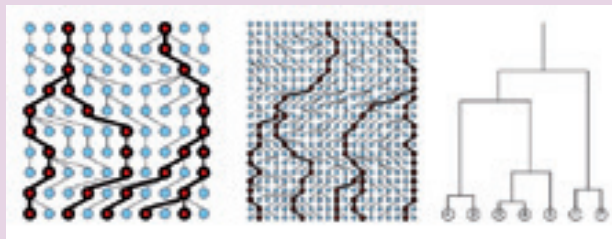


Figure 4. Le coalescent de Kingman comme limite d'échelle d'arbres généalogiques (d'après B. Mallein, Culture Math, ENS Ulm, 2011)

Pour aller plus loin

Bacaer N., (2008). *Histoires de mathématiques et de populations*. Le sel et le fer, Cassini.

<http://www.math.ens.fr/culturemath/articles-ens/mallein11/coalescent-de-kingman.html>

Haccou P., Jagers P., Vatutin V.A., (2005). *Branching Processes: Variation, growth and extinction of populations*, Cambridge University Press.

Tangente Hors série n° 42 (2011). *Mathématiques et biologie – L'organisation du vivant*.

Istas J., (2000). *Introduction aux modélisations mathématiques pour les sciences du vivant*. Springer.

Darwin online

<http://darwin-online.org.uk/content/frameset?pageseq=60&itemID=F1497&viewtype=side>

Mallein B., (2011). *Généalogie de populations: le coalescent de Kingman*, Culture Math, ENS Ulm.

