



Que peuvent avoir à se dire Probabilistes et Biologistes ?

Trois retours d'expériences

Sylvain Billiard

Université de Lille - France - Lab. Evo-Eco-Paléo - UMR CNRS 8198

Programme

1. Introduction

Historique et questionnements principaux

2. avec Chi Tran (Univ. G. Eiffel) :

Modélisation stochastique d'un système de reproduction chez les plantes à fleurs, et classes de problèmes mathématiques

3. avec Charline Smadi (INRAE, Grenoble) :

Dynamique d'une communauté à diverses échelles, et utilité des théorèmes

4. avec Vincent Bansaye (E. Polytechnique) :

Comment les individus interagissent ? Convergence, approximations et inférence



Great Scientist \neq Good at Math

E. O. Wilson Shares a Secret: Discoveries Emerge from Ideas, Not Number-Crunching

Fortunately, exceptional mathematical fluency is required in only a few disciplines, such as particle physics, astrophysics and information theory. Far more important throughout the rest of science is the ability to form concepts, during which the researcher conjures images and processes by intuition.

Over the years, I have co-written many papers with mathematicians and statisticians, so I can offer the following principle with confidence. Call it Wilson's Principle No. 1: It is far easier for scientists to acquire needed collaboration from mathematicians and statisticians than it is for mathematicians and statisticians to find scientists able to make use of their equations.



Bernd Sturmfels

Can Biology Lead to New Theorems?

Yet, as these ubiquitous initiatives and programs unfold, many mathematicians remain unconvinced, and some secretly hope that this “biology fad” will simply go away soon. They have not seen any substantive impact of quantitative biology in their area of expertise, and they rightfully ask: **where are the new theorems?**



Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better

Joel E. Cohen

Challenges

- *"Replace the tree of life with a network or tapestry to represent lateral transfers of heritable features"*
- *"Find better ways to model multi-level systems, for example, cells within organs within people in human communities in physical, chemical, and biotic ecologies"*
- *"Understand probability, risk, and uncertainty. Despite three centuries of great progress, we are still at the very beginning of a true understanding."*

Probabilistes et biologiste des populations : Une tradition

(Cohen 2004)

Table 1. Mathematics Arising from Biological Problems

Subject	Illustrative Reference
Age structure of stable populations	Euler 1760
Logistic equation for limited population growth	Verhulst 1838
Branching processes, extinction of family names	Galton 1889
Correlation	Pearson 1903
Markov chains, statistics of language	Markov 1906
Hardy-Weinberg equilibrium in population genetics	Hardy 1908; Weinberg 1908
Analysis of variance, design of agricultural experiments	Fisher 1950
Dynamics of interacting species	Lotka 1925; Volterra 1931
Birth process, birth and death process	Yule 1925; Kendall 1948, 1949
Traveling waves in genetics	Fisher 1937; Kolmogorov et al. 1937
Game theory	von Neumann and Morgenstern 1953
Distribution for estimating bacterial mutation rates	Luria and Delbrück 1943
Morphogenesis	Turing 1952
Diffusion equation for gene frequencies	Kimura 1994
Circular interval graphs, genetic fine structure	Benzer 1959
Threshold functions of random graphs	Erdős and Rényi 1960
Sampling formula for haplotype frequencies	Ewens 1972
Coalescent genealogy of populations	Kingman 1982a, 1982b

382

GENETICS: S. WRIGHT

PROC. N. A. S.

*THE DIFFERENTIAL EQUATION OF THE DISTRIBUTION OF
GENE FREQUENCIES*

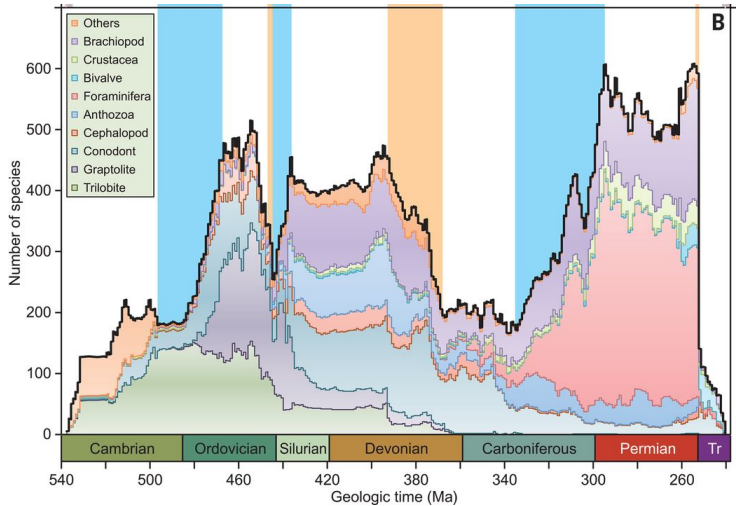
BY SEWALL WRIGHT

DEPARTMENT OF ZOOLOGY, THE UNIVERSITY OF CHICAGO

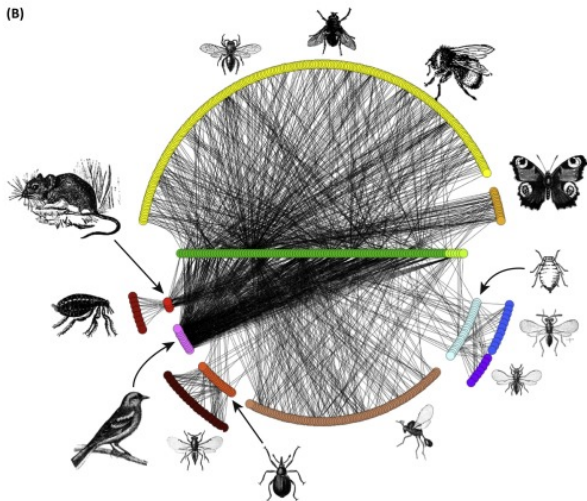
Communicated October 15, 1945

Dr. A. Kolmogorov⁸ has recently been kind enough to send me a reprint of an important paper on this subject which was published in 1935 but which had not previously come to my attention.

0. La grande question (1)



0. La grande question (2)



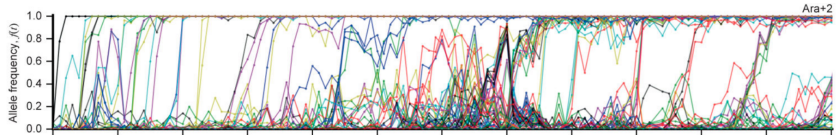
1. Tout est stochastique

1. Tout est stochastique

Proposition 1.

- La dérive génétique n'existe pas
- La stochasticité démographique non plus

Good, Benjamin H., Michael J. McDonald, Jeffrey E. Barrick, Richard E. Lenski, and Michael M. Desai. 2017. "The Dynamics of Molecular Evolution Over 60,000 Generations." *Nature* 551 (7678): 45-50.



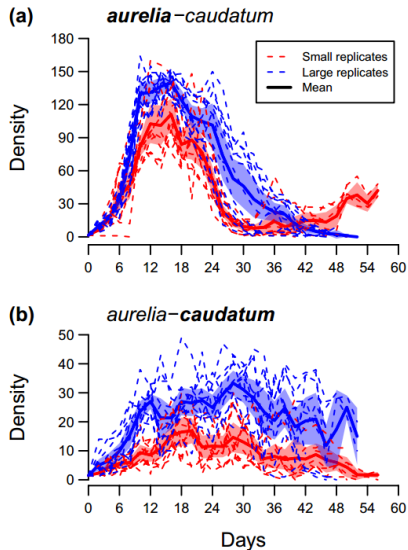
1. Tout est stochastique

Proposition 2.

- La sélection naturelle est une interprétation, pas un mécanisme
- L'exclusion compétitive, les interactions mutualistes, etc., idem

Demographic stochasticity alters expected outcomes in experimental and simulated non-neutral communities

Geoffrey Legault, Jeremy W. Fox and Brett A. Melbourne



1. Tout est stochastique

Proposition 3.

- Tout **motif** (ou *pattern*) que l'on observe est le résultat statistique d'un certain nombre de particules en interaction, à une certaine échelle
- Tout ce que l'on interprète comme **processus** (ou *process*) : idem.

1. Tout est stochastique

Conséquences.

- Et les interactions elles-mêmes alors ?
→ Cours avec Vincent
- Spéciation et extinctions d'espèce ?
→ Cours avec Chi
- Modèle déterministe = approximation à une certaine échelle d'un modèle stochastique à une autre échelle
→ Cours avec Charline
- Changer notre façon d'expliquer les choses (ex. pertinence de "neutre" vs. "sélection" ?)
- Qu'appelle-t-on **motif** et **processus** ?

1. Tout est stochastique

Conséquences.

- Et les interactions elles-mêmes alors ?
→ Cours avec Vincent
- Spéciation et extinctions d'espèce ?
→ Cours avec Chi
- Modèle déterministe = approximation à une certaine échelle d'un modèle stochastique à une autre échelle
→ Cours avec Charline
- Changer notre façon d'expliquer les choses (ex. pertinence de "neutre" vs. "sélection" ?)
- Qu'appelle-t-on **motif** et **processus** ?

1. Tout est stochastique

Conséquences.

- Et les interactions elles-mêmes alors ?
→ Cours avec Vincent
- Spéciation et extinctions d'espèce ?
→ Cours avec Chi
- Modèle déterministe = approximation à une certaine échelle d'un modèle stochastique à une autre échelle
→ Cours avec Charline
- Changer notre façon d'expliquer les choses (ex. pertinence de "neutre" vs. "sélection" ?)
- Qu'appelle-t-on **motif** et **processus** ?

1. Tout est stochastique

Conséquences.

- Et les interactions elles-mêmes alors ?
→ Cours avec Vincent
- Spéciation et extinctions d'espèce ?
→ Cours avec Chi
- Modèle déterministe = approximation à une certaine échelle d'un modèle stochastique à une autre échelle
→ Cours avec Charline
- Changer notre façon d'expliquer les choses (ex. pertinence de "neutre" vs. "sélection" ?)
- Qu'appelle-t-on motif et processus ?

1. Tout est stochastique

Conséquences.

- Et les interactions elles-mêmes alors ?
→ Cours avec Vincent
- Spéciation et extinctions d'espèce ?
→ Cours avec Chi
- Modèle déterministe = approximation à une certaine échelle d'un modèle stochastique à une autre échelle
→ Cours avec Charline
- Changer notre façon d'expliquer les choses (ex. pertinence de "neutre" vs. "sélection" ?)
- Qu'appelle-t-on **motif** et **processus** ?

"[...] Buckle demonstrated **statistically** that if only **a sufficient number of people** is taken into account, then not only is the **number of natural events** like death, illness, etc. perfectly **constant**, but also the number of so-called voluntary actions - marriages at a given age, crimes, and suicides. **It occurs no differently among molecules.**"

— Ludwig Boltzmann 1886 (in Gigerenzer, « *The empire of chance* »)



The Origin of Species was published toward the end of the year 1859. The preceding years since 1846 had been one of the most productive seasons – or if extended so as to cover the great book we are considering, *the* most productive period of equal length in the entire history of science from its beginnings until now. [For] the idea that chance begets order, which is one of the cornerstones of modern physics . . . was at that time put into its clearest light (Peirce, 1893, p. 183).

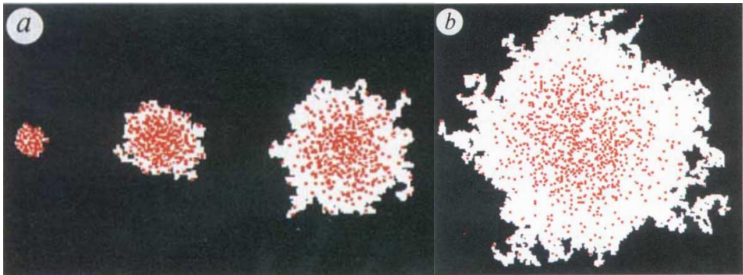
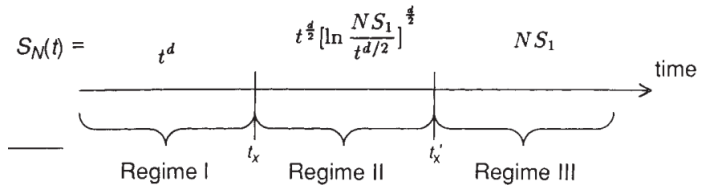
Causalité statistique (Wise 2015)

2. Le concept d'échelle : central, pas simple et polysémique

- Echelles de temps, d'espace, de paramètres

Territory covered by N diffusing particles

Hernan Larralde*, Paul Trunfio*, Shlomo Havlin*†, H. Eugene Stanley* & George H. Weiss†



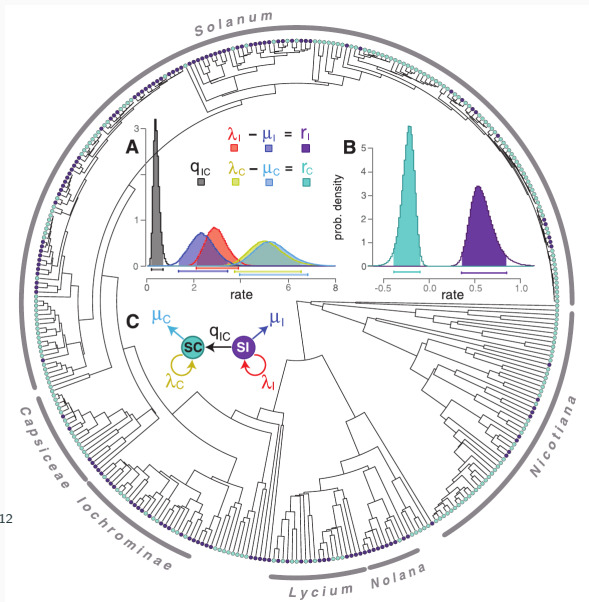
2. Le concept d'échelle: central, pas simple et polysémique

- Echelles d'organisation, d'observation, d'abstraction, d'individuation
→ Clarifier notre emploi de "*hasard*"

2. Le concept d'échelle: central, pas simple et polysémique

- Même modèles probabilistes à toutes échelles ?
- Si oui lesquels ?
- Si non, pourquoi ?
- Comment répondre à ces questions ?
- Conflits entre échelles ?
- Imbrication des échelles ?

Peut-on parler de sélection à l'échelle des espèces ?



Goldberg et al. 2012

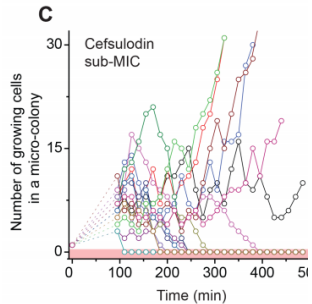
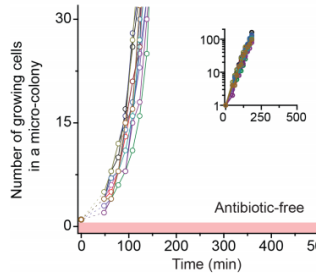
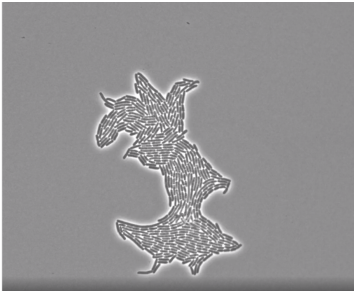
3. Implications pour l'inférence

Ne pas faire de différences entre

- Modèles basés sur les processus
- Modèles statistiques et d'inférences
- Expérimentations, observations, mesures

Antibiotic-induced population fluctuations and stochastic clearance of bacteria

Jessica Coates^{1†}, Bo Ryoung Park^{2†}, Dai Le³, Emrah Şimşek², Waqas Chaudhry³, Minsu Kim^{1,2,3*}



Réflexe de biologistes... et vous ?

3. Implications pour l'inférence

Quatre sources de variabilité

- Extrinsèque (ou environnementale)
- Entre individus
- Erreurs d'observations, d'expérimentations
- Intrinsèque (d'un processus sous-jacent)

3. Implications pour l'inférence

Enjeu : quantifier les quatre sources dans un jeu de données

- Quel modèle inférentiel ?
- Quel schéma expérimental ?
- Quelle information tirer ?

3. Implications pour l'inférence

Les fluctuations nous renseignent sur le processus d'intérêt ...

... abusons-en !

$$Y = F(\Theta, E, Individu) + \varepsilon(\Theta, E, Individu, Erreurs)$$

1. Introduction

Historique et questionnements principaux

2. avec Chi Tran (Univ. G. Eiffel) :

Modélisation stochastique d'un système de reproduction chez les plantes à fleurs, et classes de problèmes mathématiques

3. avec Charline Smadi (INRAE, Grenoble) :

Dynamique d'une communauté à diverses échelles, et utilité des théorèmes

4. avec Vincent Bansaye (E. Polytechnique) :

Comment les individus interagissent ? Convergence, approximations et inférence

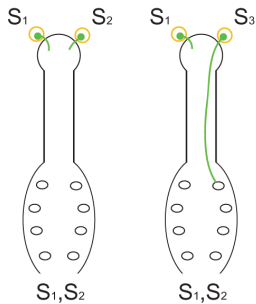
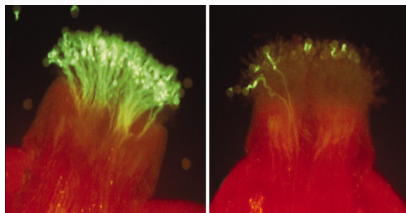
Système d'auto-incompatibilité chez les plantes

- Un système d'appariement
- Chez des dizaines de familles (évolutions indépendantes)
- Chez $\sim 50\%$ espèces
- Un phénotype aux classes discrètes
- Un déterminisme génétique simple
- Intérêt économique, etc.

Comment ça marche ?



(photo: P. Saumitou)



Self-Incompatibility in Brassicaceae: Identification and Characterization of *SRK*-Like Sequences Linked to the *S*-Locus in the Tribe Biscutelleae

Jean-Baptiste Leducq, Célia C Gosset, Rita Gries, Kevin Calin, Éric Schmitt, Vincent Castric and Xavier Vekemans

G3: GENES, GENOMES, GENETICS June 1, 2014 vol. 4 no. 6 983-992;



	Sequences		S-haplotypes		Pollen donor																							
	Ind.	1	2	1	2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
Pollen receptor	1			=	S01	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
	2			=	S01	+	x	+	+	+	+	+	+	+	+	+	+	+	+	+	+	x	+	+	+	+	+	
	3			S01	S01		+	+	-	+	+	+	+	-	+	-	+	-	+	+	+	+	+	+	+	+	+	+
	4			S01	S01	S01		S02	S02	S02	S02	+	+	+	+	+	+	+	+	+	+	x	+	+	+	+	+	+
	5			+	x	+	S02		S02	S02	S02	+	x	+	+	+	+	+	+	+	+	x	+	-	+	+	+	+
	6			+	+	+	S02	S02		S02	+	+	+	+	+	+	+	+	+	+	+	x	S07	+	+	+	+	+
	7			+	+	+	S02	S02	S02		S03	S03	S03	+	+	+	+	+	+	+	+	x	+	S03	+	+	+	+
	8			+	+	+	+	+	+	+	S03		S03	S03	+	+	-	+	+	+	+	+	+	+	S03	+	+	+
	9			+	+	+	+	+	+	+	S03	S03		S03	+	+	+	+	+	+	+	x	+	S03	+	+	+	+
	10			+	+	+	+	+	+	+	S03	S03	S03		S04	S04	-	+	+	+	+	+	+	S03	+	+	+	+
	11			+	+	+	+	+	+	+	+	+	+	+	S04		S04	S06	+	+	+	+	S06	+	+	+	+	S04
	12			+	-	+	+	+	+	+	+	+	+	+	S04	S04		S05	S05	S05	S05	+	+	+	+	+	+	S04
	13			+	+	+	+	+	+	+	+	+	+	+	+	+	S05		=	S05	S05	S06	+	+	+	+	+	+
	14			+	+	+	+	+	+	+	+	+	+	+	+	+	S05	=		S05	S05	S06	+	+	+	+	+	+
	15			+	+	+	+	+	+	+	+	+	+	+	+	+	S05	S05	S05		=	+	+	-	+	+	+	+
	16			+	+	+	+	x	+	+	+	+	+	+	-	+	S05	S05	S05	=		x	+	+	+	S10	+	+
	17			-	+	-	x	+	x	x	-	x	-		S06	+	+	+	+	+	+			-	-	+	+	+
	18			+	+	+	+	+	S07	+	+	+	+	+	+	+	+	+	+	+	+	+	S06	+	+	+	+	+
	19			+	+	+	+	+	+	+	+	+	+	+	S03	+	+	-	+	+	+	+	+	-	+	S08	+	+
	20			+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	S08		+	+
	21			+	+	-	+	+	+	+	+	+	-	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+

Une vieille obsession... empirique

TABLE 1. Number of incompatibility alleles known or estimated to be present in populations, n_p and N_p , of 19 flowering plants

System	Species	R	Populations	
			n_p	N_p
Gametophytic one locus	<i>Nicotiana glauca</i> ¹	0.59	15	18
	<i>Oenothera organensis</i> ²	0.65	28	30
	<i>Trifolium repens</i> ³	0.07	17	139
		0.19	28	80
		0.28	36	74
	<i>Trifolium pratense</i> ⁴	0.22	39	101
		0.16	41	143
		0.09	33	193
	<i>Trifolium hybridum</i> ⁵	0.47	12	17
	<i>Papaver rhoeas</i> ⁶	0.65	27	35+
		0.68	25	32+
		0.72	31	38+
	<i>Phlox drummondii</i> ⁷	0.40	30	45
	<i>Solanum carolinense</i> ⁸	0.68	11	12
	<i>Physalis crassifolia</i> ⁹	0.57	12	14
		0.39	28	44
		<i>Physalis cinerascens</i> ¹⁰	0.57	12
<i>Lycium andersonii</i> ¹¹		0.34	22	38
<i>Witheringia maculata</i> ¹²		0.67	10	14
Gametophytic two loci	<i>Lolium perenne</i> ¹³	0.58	17S	31+S
		0.59	17Z	31+Z
Sporophytic one locus	<i>Iberis amara</i> ¹⁴			> 22?
	<i>Raphanus raphanistrum</i> ¹⁵			
	<i>Brassica oleracea</i> ¹⁶			
	<i>Brassica campestris</i> ¹⁷	0.58	16	22
		0.52	18	31
	<i>Sinapis arvensis</i> ¹⁸	0.51	35	43
	<i>Ipomoea trifida</i> ¹⁹	0.96	5	
		0.83	15	
		0.77	19	
0.83		16		
0.76		21		
	0.93	6		

(Lawrence 2000)

THE DISTRIBUTION OF SELF-STERILITY ALLELES IN POPULATIONS

SEWALL WRIGHT

The University of Chicago, Chicago, Illinois

Received February 18, 1939

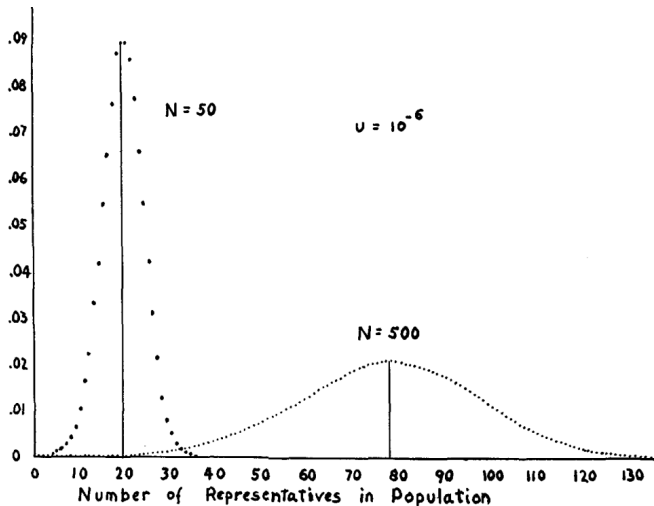
$$\Delta q = \frac{q(1-q)(1+R)}{2[R+q(1-R)]} - [1+u+m(1-q)]q + (v+mq)(1-q)$$

$$\sigma_{\Delta q}^2 = \frac{q(1-q)}{2N} .$$

THE DISTRIBUTION FORMULA

$$\phi(q) = C[R+q(1-R)]^{2N(1+R)/(1-R)} q^{4N(v+mq)-1} (1-q)^{4N[1+u+m(1-q)]-1} .$$

Une vieille obsession... probabiliste



Un nouveau regard probabiliste (avec P. Csuppon)

Birth rate of A_{ij} individuals

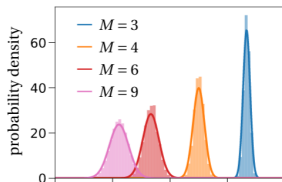
$$T_{ij}^+ = \left(\underbrace{\frac{1}{2} \sum_{k \neq i, j} A_{jk} \frac{p_i}{1 - p_j - p_k}}_{i \text{ pollen fertilization}} + \underbrace{\frac{1}{2} \sum_{k \neq i, j} A_{ik} \frac{p_j}{1 - p_i - p_k}}_{j \text{ pollen fertilization}} \right) \underbrace{\frac{N - A_{ij}}{N}}_{\text{non-} A_{ij} \text{ replacement}}$$

Death rate of A_{ij} individuals

$$T_{ij}^- = \left(N - \frac{1}{2} \sum_{k \neq i, j} A_{ik} \frac{p_j}{1 - p_i - p_k} - \frac{1}{2} \sum_{k \neq i, j} A_{jk} \frac{p_i}{1 - p_j - p_k} \right) \frac{A_{ij}}{N}$$

Stationary distribution

$$p^* \sim \mathcal{N} \left(\frac{1}{M}, \frac{(M-2)^3}{NM^2(2M-3)} \right)$$



(en cours)

A general stochastic model for sporophytic self-incompatibility

Sylvain Billiard · Viet Chi Tran

$$Z_t(dg) = \sum_{i=1}^{N_t} \delta_{\{S^1(i), S^2(i)\}}(dg) = \sum_{g'=\{u', v'\} \in E} N_t^{u'v'} \delta_{\{u', v'\}}(dg)$$

$$N_t^{uv} = N_0^{uv} + \int_0^t (r^{uv}(Z_s) - d \times N_s^{uv}) ds + M_t^{uv}$$

$$\langle M^{uv} \rangle_t = \int_0^t (r^{uv}(Z_s) + d \times N_s^{uv}) ds.$$

Un nouveau regard probabiliste (avec Chi Tran)

$$\begin{aligned} r^{uv}(Z_t) = & \bar{r} \mathbb{1}_{u \neq v} \left[\frac{1}{4} \left(\sum_{u' \neq u} \sum_{v' \neq v} N_t^{uu'} p^{uu'}(v, v') + N_t^{vv'} p^{vv'}(u, u') \right) \right. \\ & + \frac{1}{2} \left(\sum_{v' \neq v} N_t^{uu} p^{uu}(v, v') + \sum_{u' \neq u} N_t^{uu'} p^{uu'}(v, v) \right) \\ & + \frac{1}{2} \left(\sum_{u' \neq u} N_t^{vv} p^{vv}(u, u') + \sum_{v' \neq v} N_t^{vv'} p^{vv'}(u, u) \right) \\ & \left. + (N_t^{uu} p^{uu}(v, v) + N_t^{vv} p^{vv}(u, u)) \right] \\ & + \bar{r} \mathbb{1}_{u=v} \left[\frac{1}{4} \sum_{u', v' \in \llbracket 1, n \rrbracket \setminus \{u\}} N_t^{uu'} p^{uu'}(u, v') \right. \\ & + \frac{1}{2} \left(\sum_{u' \neq u} N_t^{uu'} p^{uu'}(u, u) + \sum_{v' \neq u} N_t^{uu} p^{uu}(u, v') \right) \\ & \left. + N_t^{uu} p^{uu}(u, u) \right]. \end{aligned}$$

Un problème classique : sous représentation dans les îles ?



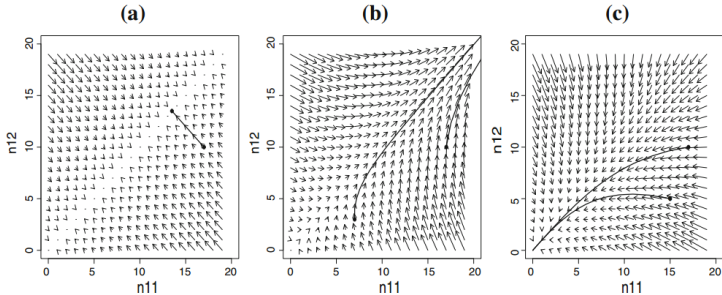
3/34 espèces endémiques sont auto-incompatibles (Chamorro et al. 2012)

Un cas particulier : Distylie (primevère, lilas, jasmin, ...)

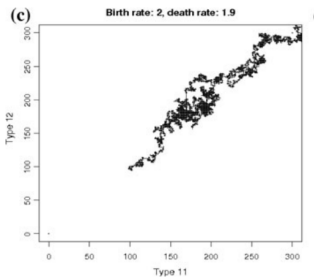
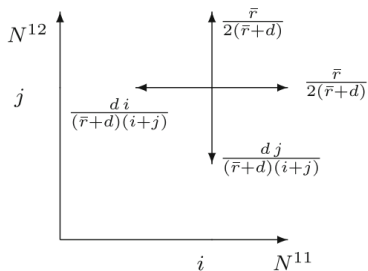


Approximation en grande population

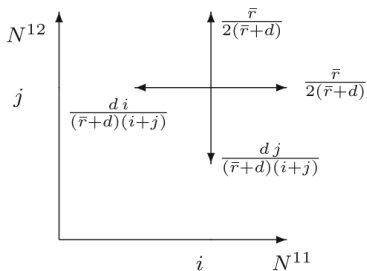
$$\frac{dn_t^{uv}}{dt} = r^{uv}(\xi_t) - d n_t^{uv}$$



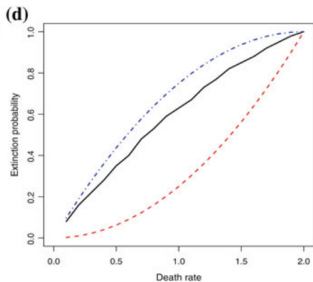
En petite population : marche aléatoire dans un quart de plan



En petite population : colonisation, extinction



$$\left(\frac{d}{\bar{r}}\right)^{i+j} \leq p_{i,j} \leq \left(\frac{d}{\bar{r}}\right)^i + \left(\frac{d}{\bar{r}}\right)^j - \left(\frac{d}{\bar{r}}\right)^{i+j}$$



Avec Sylvain : marche aléatoire pour les populations distyles

Kilian Raschel et les marches aléatoires dans le quart de plan

Sylvain et Laurence : marches aléatoires sur des triangles ?

Transitions pour le modèle distyle

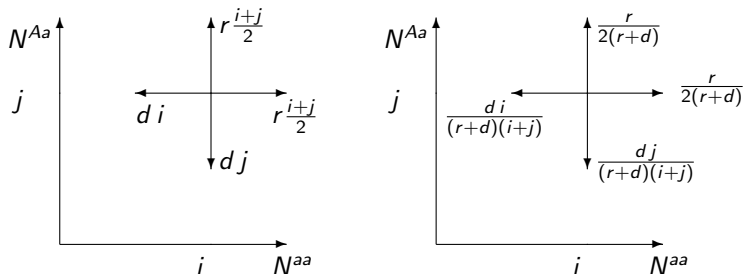


Figure: Evolution de $(N_t^{aa}, N_t^{Aa})_{t \in \mathbb{R}_+}$ pour le modèle de Wright. Les taux d'événements sont à gauche et les transitions de probabilité de la chaîne de Markov induite à droite.

Equation pour le calcul de la probabilité d'extinction

★ $\tau_0 = \inf\{t \geq 0 : N_t^{Aa} = 0 \text{ or } N_t^{aa} = 0\}$

★ $p_{i,j} = \mathbb{P}_{i,j}(\tau_0 < +\infty)$

Prop $(p_{i,j})_{i,j \geq 1}$ est la plus petite solution positive de :

$$p_{i,j} = \frac{di}{(r+d)(i+j)} p_{i-1,j} + \frac{dj}{(r+d)(i+j)} p_{i,j-1} \\ + \frac{r}{2(r+d)} p_{i+1,j} + \frac{r}{2(r+d)} p_{i,j+1}$$

avec les conditions au bord :

$$p_{i,0} = p_{0,j} = 1.$$

★ Problèmes avec réflexions au bord : files d'attente.

1. Foddy. PhD, 1984.

2. Guillemin, Simonian, Nasri, Rodriguez, arXiv:2009.13254, 2020

Simulations de la marche aléatoire

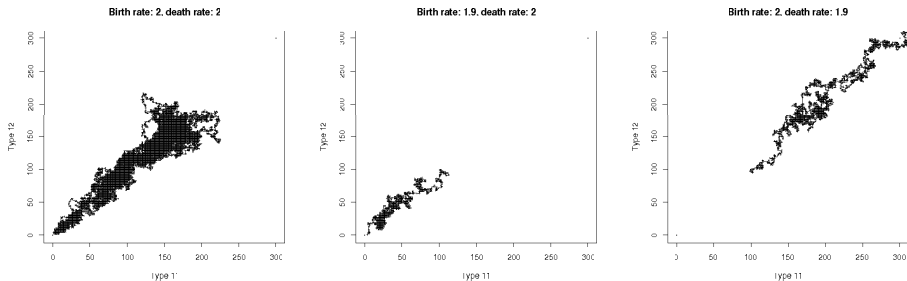


Figure: Simulation des trajectoires $(N_t^{aa}, N_t^{Aa})_{t \geq 0}$ dans les cas critique, sous-critique et sur-critique. Le nombre d'individus $\{a, a\}$ est en abscisse et le nombre d'individus $\{A, a\}$ en ordonnées. Pour les trois simulations, la condition initiale est (100, 100).

Probabilité d'extinction

★ Comparaison avec un Galton-Watson (bi-type) :

Si $r \leq d$, on a extinction presque-sûre.

Si $r > d$, la probabilité d'extinction d'un Galton-Watson issu de (i, j) individus est :

$$\left(\frac{d}{r}\right)^{i+j}.$$

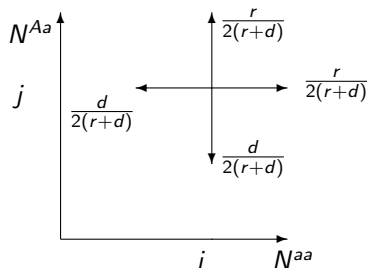
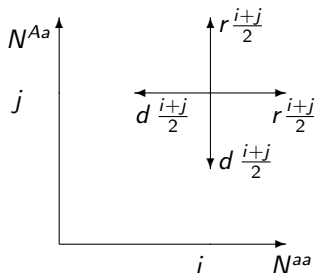
★ Alors :

$$p_{ij} \geq \left(\frac{d}{r}\right)^{i+j}.$$

Probabilité d'extinction

★ Comparaison avec une marche aléatoire homogène sur le quart de plan

:



$$\mathbb{P}_{ij}(\exists t \in \mathbb{R}_+, N_t^{Aa} = 0 \text{ or } N_t^{aa} = 0) = \left(\frac{d}{r}\right)^i + \left(\frac{d}{r}\right)^j - \left(\frac{d}{r}\right)^{i+j}.$$

Probabilité d'extinction

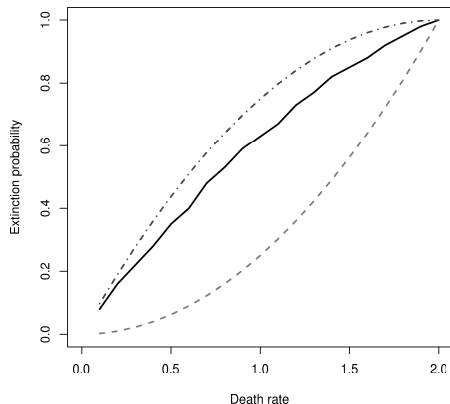


Figure: Simulation de la probabilité d'extinction de (N^{aa}, N^{Aa}) . La condition initiale est $(N_0^{aa}, N_0^{Aa}) = (1, 1)$.

Avec Sylvain : marche aléatoire pour les populations distyles

Kilian Raschel et les marches aléatoires dans le quart de plan

Sylvain et Laurence : marches aléatoires sur des triangles ?

Marches aléatoires homogènes

★ On considère les fonctions

$$Q(x, y) = \sum_{n \geq 0} \left(\sum_{i, j \geq 0} q_G(n; i, j) x^i y^j \right) t^n.$$

où $q_G(n; i, j)$ est la probabilité d'arriver en (i, j) partant de $(0, 0)$ avec n pas.

★ Pour des marches aléatoires homogènes sur le quart de plan :

$$R(x, y)Q(x, y) = R(x, 0)Q(x, 0) + R(0, y)Q(0, y) - R(0, 0)Q(0, 0) - xy.$$

On s'intéresse alors au lieu où $R(x, y) = 0$ (courbe algébrique).

De l'équation des $p_{i,j}$ à une EDP

★ $P(x, y) = \sum_{i,j \geq 1} p_{i,j} x^i y^j$, for $x, y \in [0, 1)$.

★ $P(x, y)$ vérifie $AP(x, y) = h(x, y, P)$ où

$$AP(x, y) = Q(x, y)\partial_x P(x, y) + Q(y, x)\partial_y P(x, y) + R(x, y)P(x, y)$$

$$Q(x, y) = (r + d)x - \frac{r}{2} - \frac{r}{2} \frac{x}{y} - dx^2$$

$$R(x, y) = \frac{r}{2x} + \frac{r}{2y} - dx - dy$$

$$\begin{aligned} h(x, y, P) &= -\frac{r}{2} \left(\sum_{i \geq 1} p_{i,1} i x^i + \sum_{j \geq 1} p_{1,j} j y^j \right) + d \left(\sum_{i \geq 1} x^i y + \sum_{j \geq 1} y^j x \right) \\ &= -\frac{r}{2} \left(x \frac{\partial^2 P}{\partial x \partial y}(x, 0) + y \frac{\partial^2 P}{\partial y \partial x}(0, y) \right) + dxy \left(\frac{1}{1-x} + \frac{1}{1-y} \right) \end{aligned}$$

Résolution

★ Equations des caractéristiques:

$$\begin{cases} \dot{x}_s = \frac{dx}{ds}(s) = Q(x_s, y_s), \\ \dot{y}_s = \frac{dy}{ds}(s) = Q(y_s, x_s), \end{cases}$$

★ Résolution :

$$x_s = \frac{\lambda r e^{rs} + \mu d e^{ds}}{d(\lambda e^{rs} + \mu e^{ds} + 1)}, \quad y_s = \frac{\lambda r e^{rs} + \mu d e^{ds}}{d(\lambda e^{rs} + \mu e^{ds} - 1)},$$

où :

$$\lambda = \frac{2d x_0 y_0 - d(x_0 + y_0)}{(x_0 - y_0)(r - d)}, \quad \mu = \frac{-2d x_0 y_0 + r(x_0 + y_0)}{(x_0 - y_0)(r - d)}.$$

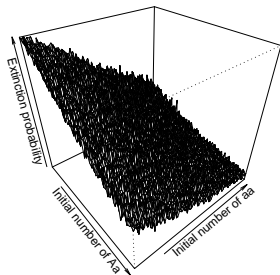
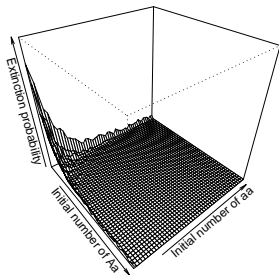
★ Enfin :

$$\begin{aligned} P(x_0, y_0) &= \frac{r}{2} \sum_{i \geq 1} \int_0^{s_0} (p_{i,1} i(x_u)^i + p_{1,i} i(y_u)^i) e^{\int_0^u R(x_\alpha, y_\alpha) d\alpha} du \\ &\quad - d \int_0^{s_0} x_u y_u \left(\frac{1}{1 - x_u} + \frac{1}{1 - y_u} \right) e^{\int_0^u R(x_\alpha, y_\alpha) d\alpha} du. \end{aligned}$$

Développements

★ Pour les bords, on utilise que pour i fixé, lorsque $N \rightarrow +\infty$,

$$p_{i,N} \sim \left(\frac{2d}{r}\right)^i \frac{i!}{N^i}.$$



Gauche : $r = 3$ et $d = 2$; Droite : $r = 2.002$ et $d = 2$

Prolongements par Kilian (et Gerald)

★ On définit :

$$\kappa_* := \liminf_{i \rightarrow +\infty} p_{i,i}^{1/i}, \quad \kappa^* := \limsup_{i \rightarrow +\infty} p_{i,i}^{1/i}$$

La détermination des bornes κ_* et κ^* pose elle-même un problème.

★ Si $d < r$, alors il existe $\nu \in (0, d/r)$ tel que lorsque $i \rightarrow +\infty$,

$$p_{i,i} = o(\nu^i).$$

★ On montre :

$$\left(\frac{d}{r}\right)^2 \leq \kappa_* \leq \kappa^* \leq \frac{d}{r}.$$

★ En utilisant de l'analyse harmonique :

$$\begin{aligned} \left(\frac{d}{r}\right)^{i+j} \vee \left[\left(1 + \frac{d}{2(r+d)}\right)^{i \wedge j} \frac{i!j!}{(i+j)!} \left(\left(\frac{d}{r}\right)^i + \left(\frac{d}{r}\right)^j \right) \right] \\ \leq p_{i,j} \leq \left(\frac{d}{r}\right)^i + \left(\frac{d}{r}\right)^j - \left(\frac{d}{r}\right)^{i+j}. \end{aligned}$$

Avec Sylvain : marche aléatoire pour les populations distyles

Kilian Raschel et les marches aléatoires dans le quart de plan

Sylvain et Laurence : marches aléatoires sur des triangles ?

Et que se passe-t-il avec 3 allèles ?

★ Génotypes et phénotype :

$$a \Rightarrow b \leftrightarrow c$$

Génotype	aa	bb	cc	ab	ac	bc
Phénotype	a	b	c	a	a	bc

★ Appariement possibles

$$\begin{array}{llll} ab & bb & \rightarrow & ab \left(\frac{1}{2}\right) + bb \left(\frac{1}{2}\right) \\ ab & cc & \rightarrow & ac \left(\frac{1}{2}\right) + bc \left(\frac{1}{2}\right) \\ ab & bc & \rightarrow & ab \left(\frac{1}{4}\right) + ac \left(\frac{1}{4}\right) + bb \left(\frac{1}{4}\right) + bc \left(\frac{1}{4}\right) \\ ac & bb & \rightarrow & ab \left(\frac{1}{2}\right) + bc \left(\frac{1}{2}\right) \\ ac & cc & \rightarrow & ac \left(\frac{1}{2}\right) + cc \left(\frac{1}{2}\right) \\ ac & bc & \rightarrow & ab \left(\frac{1}{4}\right) + ac \left(\frac{1}{4}\right) + bc \left(\frac{1}{4}\right) + cc \left(\frac{1}{4}\right) \\ bb & cc & \rightarrow & bc (1) \end{array}$$

1. Introduction

Historique et questionnements principaux

2. avec Chi Tran (Univ. G. Eiffel) :

Modélisation stochastique d'un système de reproduction chez les plantes à fleurs, et classes de problèmes mathématiques

3. avec Charline Smadi (INRAE, Grenoble) :

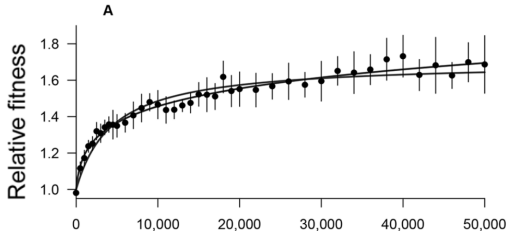
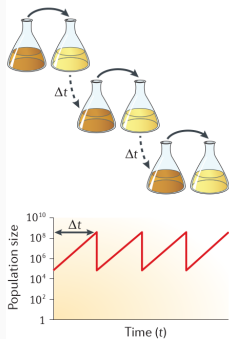
Dynamique d'une communauté à diverses échelles, et utilité des théorèmes

4. avec Vincent Bansaye (E. Polytechnique) :

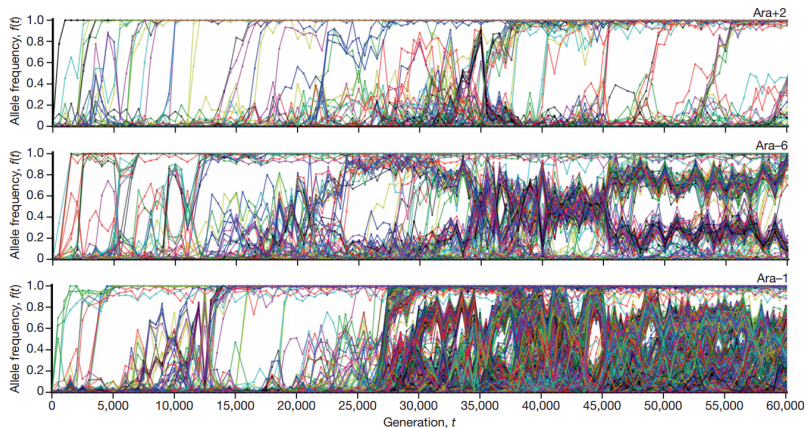
Comment les individus interagissent ? Convergence, approximations et inférence

Long-Term Dynamics of Adaptation in Asexual Populations

Michael J. Wiser,^{1,2} Noah Ribeck,^{1,3} Richard E. Lenski^{1,2,3*}



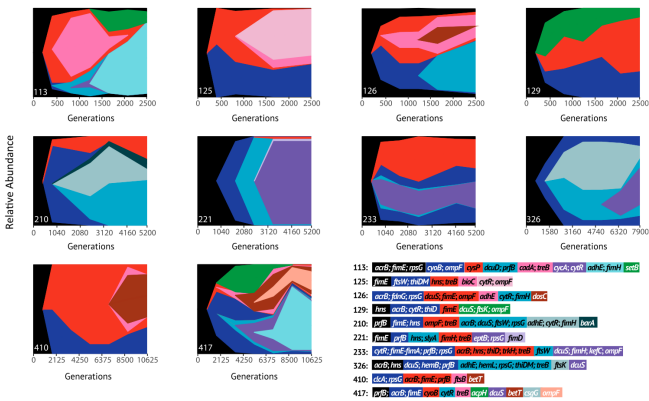
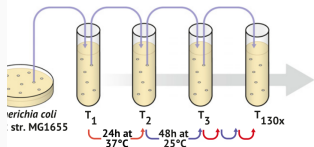
Good, Benjamin H., Michael J. McDonald, Jeffrey E. Barrick, Richard E. Lenski, and Michael M. Desai. 2017. "The Dynamics of Molecular Evolution Over 60,000 Generations." *Nature* 551 (7678): 45-50.



The dynamics of molecular evolution. Allele frequency trajectories of all *de novo* mutations detected

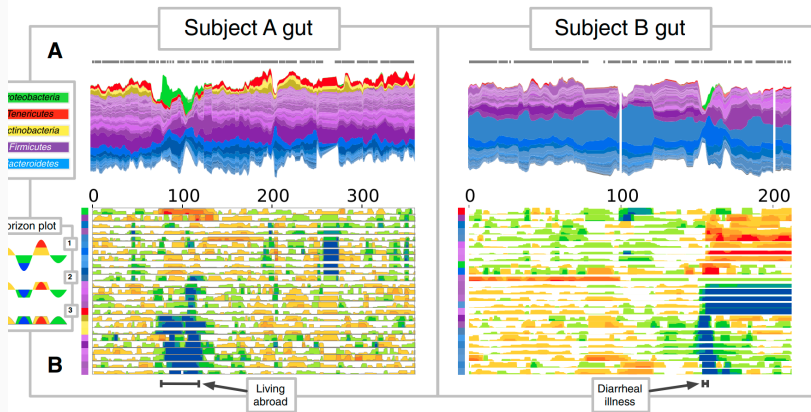
Escherichia coli cultures maintain stable subpopulation structure during long-term evolution

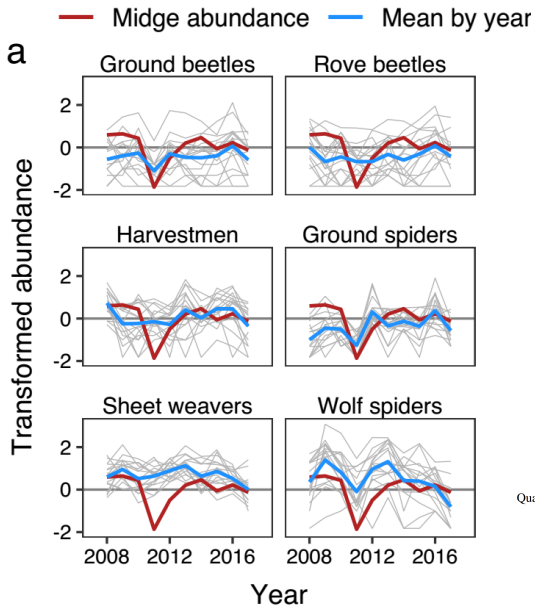
Megan G. Behringer^{a,b,1}, Brian I. Choi^a, Samuel F. Miller^b, Thomas G. Doak^{a,c}, Jonathan A. Karty^d, Wanfeng Guo^{a,b}, and Michael Lynch^{a,b,1}



lifestyle affects human microbiota on daily scales

A David^{1,2,11}, Arne C Materna³, Jonathan Friedman⁴, Maria I Campos-Baptista⁵, Matthew C Blackburn⁶,
rota⁷, Susan E Erdman⁸ and Eric J Alm^{4,7,9,10*}





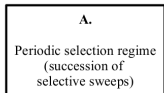
Quantifying community responses to environmental
variation from replicate time series

Joseph S. Phillips^{1,2,3,†}

Lucas A. Nell^{1,4,†}

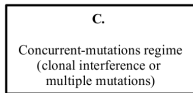
Jamieson C. Botsch¹

Regimes of communities dynamics



Two interacting clones
(≤ 2)

$$\frac{1}{K U S/\beta} \gg \frac{\ln(K S/\beta)}{S}$$



Large number of interacting clones
($\gg 2$)

$$\frac{1}{K U S/\beta} \ll \frac{\ln(K S/\beta)}{S}$$

$$\nu \propto K U S^2$$

(e.g. Crow and Kimura 1970)

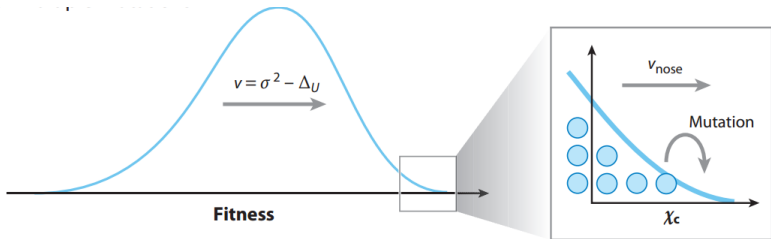
$$\nu \propto \ln(K U)^\alpha$$

(e.g. Gerrish and Lenski 1998)

$$\nu \propto S^2 \frac{\ln(K S)}{\ln^2(U/S)}$$

(e.g. Desai and Fisher 2007)

+ Tons of deterministic of population ecology and population genetics models with frequency dependent selection, or species coexistence, etc. (reviewed in Laurens et al. 2017, Chesson 2018)



Genetic Draft, Selective
Interference, and Population
Genetics of Rapid Adaptation

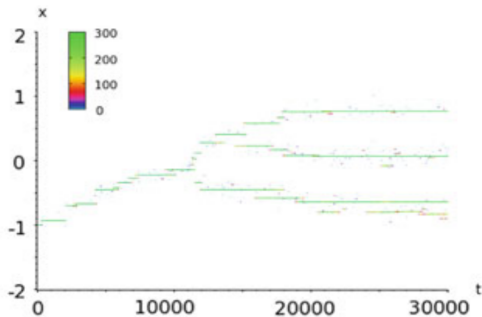
Richard A. Neher

Echelle 1 : mutations rares

Séparation des échelles de temps.

$K\mu_K \rightarrow 0$, $\ln K \ll \frac{1}{K\mu_K} \ll e^{KV}$ pour tout V quand K est grand

(Champagnat 2006, Champagnat et Méléard 2011)



Taux d'adaptation proportionnel à la taille de population et au taux de mutation.

Echelle 2 : mutations fréquentes

Plutôt étudié dans un cadre de génétique des populations.

- Population de taille N
- Accumulation de mutations favorables d'effet s_N avec un taux μ_N
- $X_j(t)$ nombre d'individus avec j mutations au temps t
- $M(t) := \frac{1}{N} \sum_{j=0}^{\infty} jX_j(t)$ nombre moyen de mutations
- Fitness d'un individu avec j mutations : $0 \vee (1 + s_N(j - M(t)))$
- Chaque individu vit un temps exponentiel de paramètre 1, et le remplaçant est choisi proportionnellement aux fitnesses.

Echelle 2 : mutations fréquentes

Différentes échelles étudiées ; point commun : de nombreuses mutations présentes en même temps dans la population

(Rouzine, Wakeley et Coffin 2003, Desai et Fisher 2007, Rouzine, Brunet et Wilke 2008 Yu, Etheridge et Cuthbertson 2010, Durrett et Mayberry 2011, Kelly 2014, ...)

Quantités étudiées :

- Premier temps où un individu a j mutations
- $\mathbb{E}[M(t) - M(t - 1)]$, $\mathbb{E}[M(t)]$
- Différence de nombre de mutations entre l'individu le 'fittest' et un individu typique
- Nombre de mutations d'un individu typique

Résultats probabilistes rigoureux sous les hypothèses :

A1: We have $\lim_{N \rightarrow \infty} \frac{\log N}{\log(s_N/\mu_N) \log(1/s_N)} = \infty$.

A2: We have $\lim_{N \rightarrow \infty} \frac{\log N}{[\log(s_N/\mu_N)]^2} \log\left(\frac{\log N}{\log(s_N/\mu_N)}\right) = 0$.

A3: We have $\lim_{N \rightarrow \infty} \frac{s_N \log N}{\log(s_N/\mu_N)} = 0$.

En particulier, le taux d'adaptation est sous linéaire en la taille de population et le taux de mutation.

Cas particulier de l'échelle $\mu_N = 1/N^\alpha, \alpha \in (0, 1)$

Regardé dans un cadre éco-évolutif, cf cours de Sylvie

(Smadi 2017, Bovier, Coquille et Smadi 2019, Champagnat, Méléard et Tran 2021, Coquille, Kraut et Smadi 2021+, Blath et Tóbiás 2021+, Blath, Paul et Tóbiás 2021+)

Regimes of communities dynamics

A.
Periodic selection regime
(succession of selective sweeps)

Two interacting clones
(≤ 2)

B.
Intermediate regime

Few interacting clones
(≥ 2)

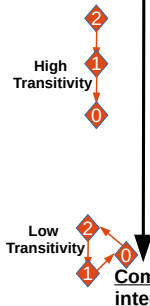
C.
Concurrent-mutations regime
(clonal interference or multiple mutations)

Large number of interacting clones
($\gg 2$)

$$\frac{1}{KUS/\beta} \gg \frac{\ln(KS/\beta)}{S}$$

$$\frac{1}{KUS/\beta} \approx \frac{\ln(KS/\beta)}{S}$$

$$\frac{1}{KUS/\beta} \ll \frac{\ln(KS/\beta)}{S}$$



- Ecological time scale
- Dynamics with two or three competing clones
- Large mutations (or immigration)

- Individual-based stochastic model in continuous time
- Scaling factor $K \rightarrow \infty$ for population size, competition and time
- Three clones **0**, **1** and **2**
- **1** and **2** have a positive probability of invasion
- Clone **2** enters the population at time α

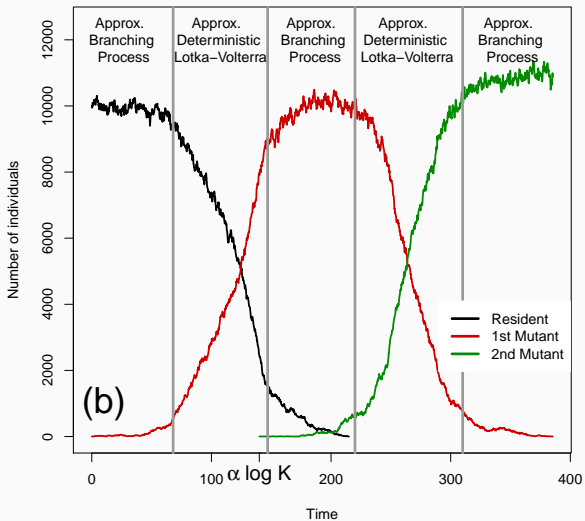
Rates of ecological processes

Birth β_i

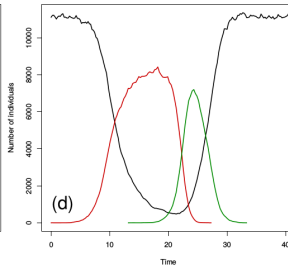
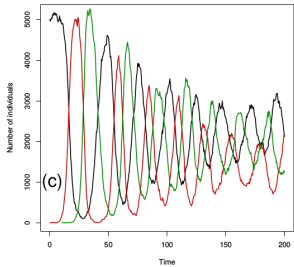
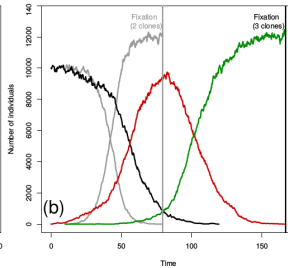
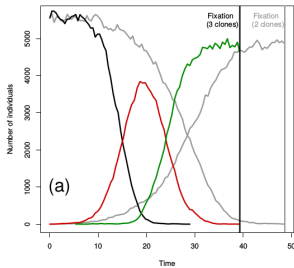
Death δ_i

Competition $\frac{C_{i,0}}{K} n_0 + \frac{C_{i,1}}{K} n_1 + \frac{C_{i,2}}{K} n_2, C_{ij} > 0$

Three clones dynamics approximation

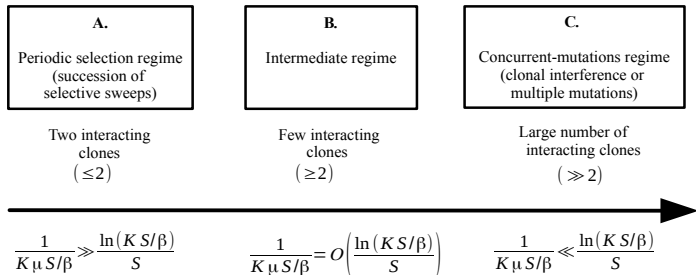


Possible dynamics: Examples

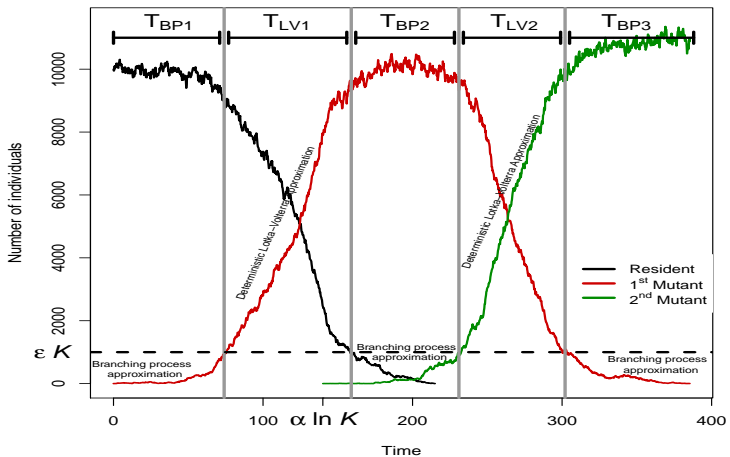


	Lower invasion probability ($S_{21} < S_{20}$ or $S_{201} < S_{20}$)	Higher invasion probability ($S_{21} < S_{20}$ or $S_{201} < S_{20}$)
Polymorphism	Soft clonal coexistence	Hard clonal coexistence
Faster fixation	Soft clonal assistance	Hard clonal assistance
Slower fixation	Soft clonal hindrance	Hard clonal hindrance

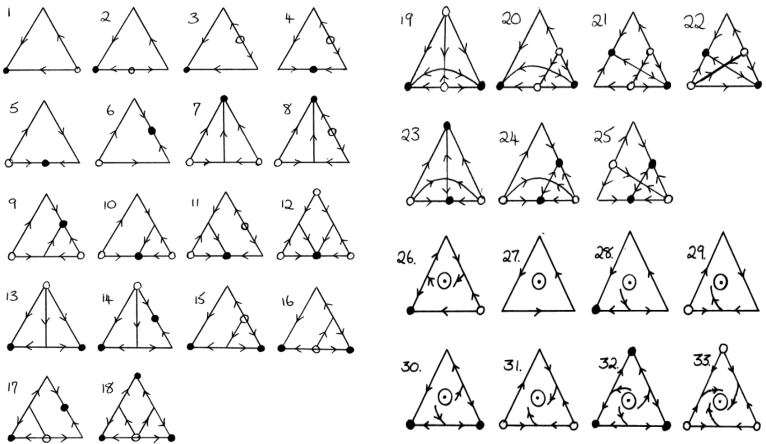
Une échelle intermédiaire



- On se restreint à une échelle de temps de l'ordre de $\ln K$.
- Echelon nécessaire si l'on veut comprendre les dynamiques complexes dans les expériences d'évolution.
- Permet de se passer d'hypothèses d'existence d'équilibres stables.
- En particulier ce qui nous intéressait en lien avec les expériences d'évolutions : les interactions non compétitives, i.e. $2 > 1$, $1 > 0$, et pourtant $0 > 2$.



(Zeeman 1990)



Des questions inspirées par la biologie

Publier dans un journal de biologie

Les théorèmes limites en pratique

Enoncé typique

Pour des paramètres donnés, il existe une constante c telle que pour tout $\varepsilon > 0$ assez petit,

$$\mathbb{P}\left(\left(1 - c\varepsilon\right)a \log K \leq T^{(\varepsilon, K)} < \left(1 + c\varepsilon\right)a \log K\right) = p + O_K(\varepsilon),$$

où $O_K(\varepsilon)$ est de l'ordre de ε pour K assez grand.

Question pertinente

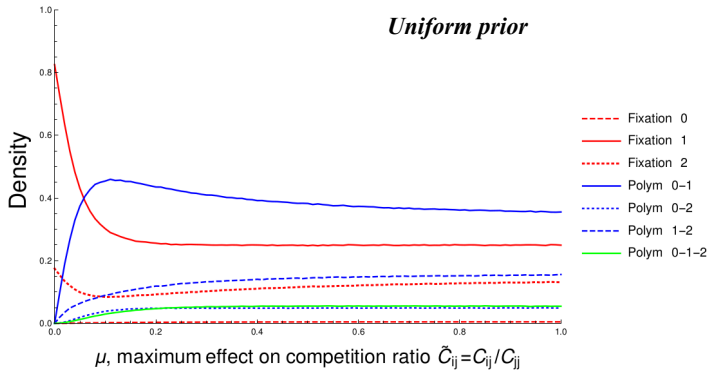
OK mais en pratique ???

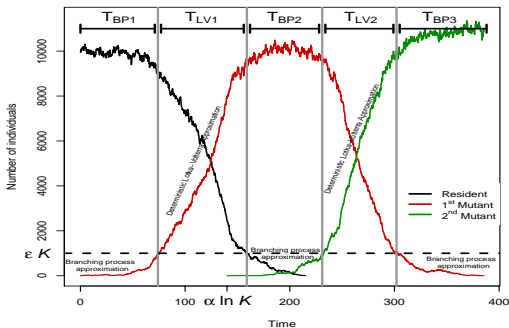
Les théorèmes limites en pratique

Contient (au moins) trois questions :

- A quelle fréquence observe-t-on les différents scénarios décrits ?
- Quelles sont les constantes qu'on néglige devant une grande quantité ?
- Pour quels K cette limite est proche de la réalité ?

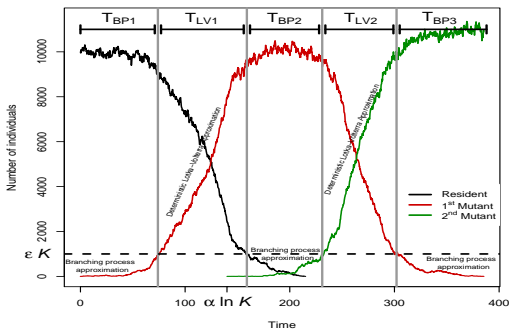
Likelihood of the final states and possible dynamics





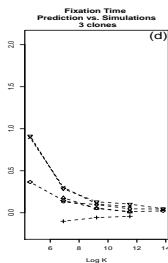
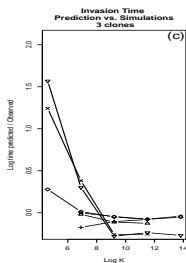
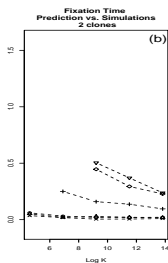
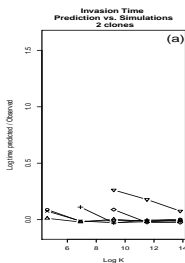
$$T_{BP1} \sim \frac{\ln K}{S_{10}} \Rightarrow T_{BP1} \sim \frac{1}{S_{10}} \left(\ln \left(\varepsilon K \frac{S_{10}}{\beta_1} \right) + \gamma \right)$$

$$T_{LV1} = O(1) \Rightarrow T_{LV1} \sim \frac{1}{S_{10}} \ln \left(\frac{(1 - \varepsilon)\bar{n}_1}{\varepsilon} \right)$$



$$T_{BP1} + T_{LV1} = \frac{1}{S_{10}} \left(\ln \left(K \frac{S_{10}}{\beta_1} \right) + \gamma \right) + \frac{1}{S_{10}} \ln \left((1 - \varepsilon) \bar{n}_1 \right)$$

Faible dépendance en ε .



Sources d'erreur

- Lorsque K est petit : population résidente affectée par compétition avec population mutante → temps surestimés.
- Lorsque S est petit : fluctuations aléatoires plus grandes de la taille de la population mutante → temps surestimés.
- Erreurs d'approximation de la durée de la phase déterministe qui se répercutent sur la durée de la phase stochastique suivante.



1. Introduction

Historique et questionnements principaux

2. avec Chi Tran (Univ. G. Eiffel) :

Modélisation stochastique d'un système de reproduction chez les plantes à fleurs, et classes de problèmes mathématiques

3. avec Charline Smadi (INRAE, Grenoble) :

Dynamique d'une communauté à diverses échelles, et utilité des théorèmes

4. avec Vincent Bansaye (E. Polytechnique) :

Comment les individus interagissent ? Convergence, approximations et inférence

From local interactions to population dynamics in site-based models of ecology

A. Johansson^a and D.J.T. Sumpter^{a,b,*}

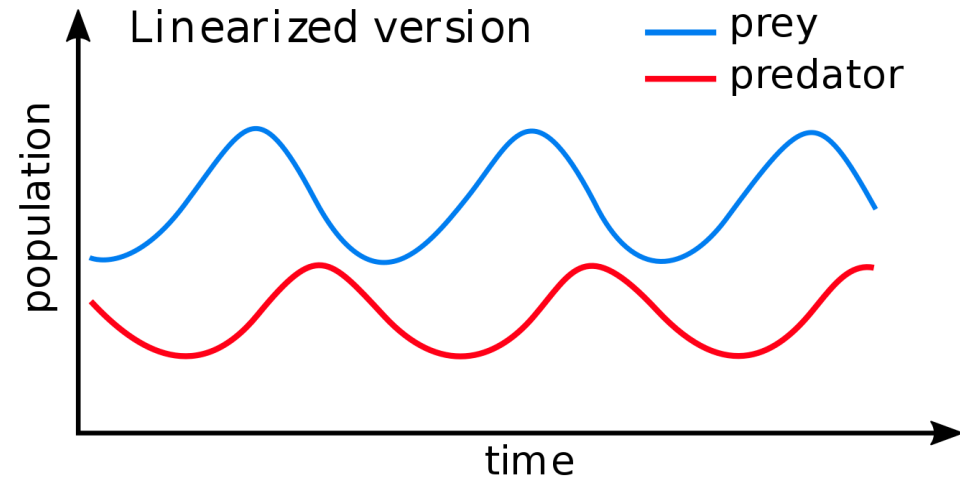
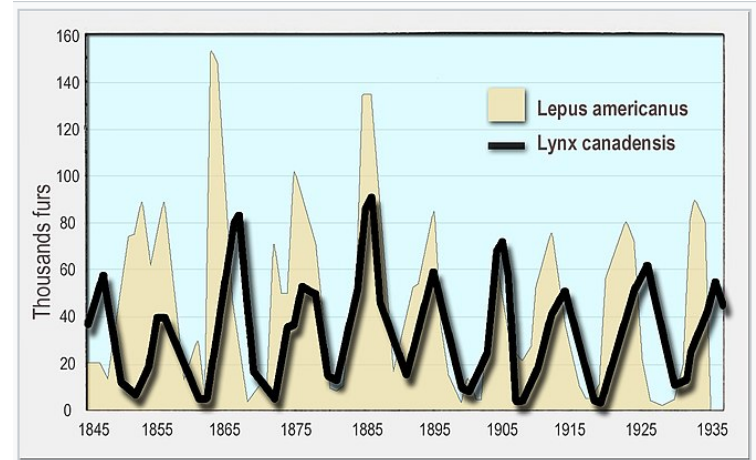
^a*Mathematics Department, Umeå University, SE-901 87 Umeå, Sweden*

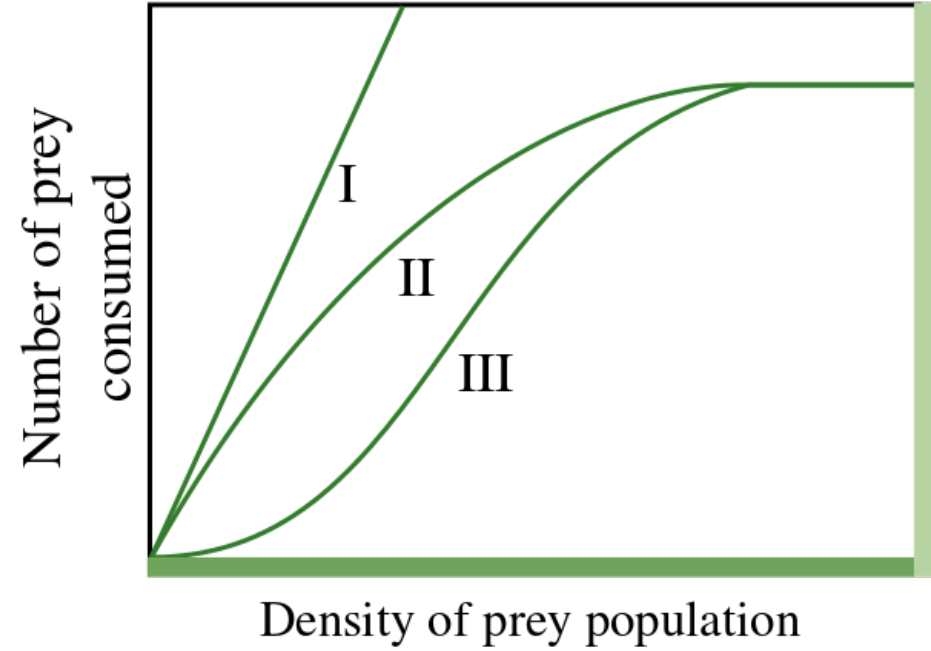
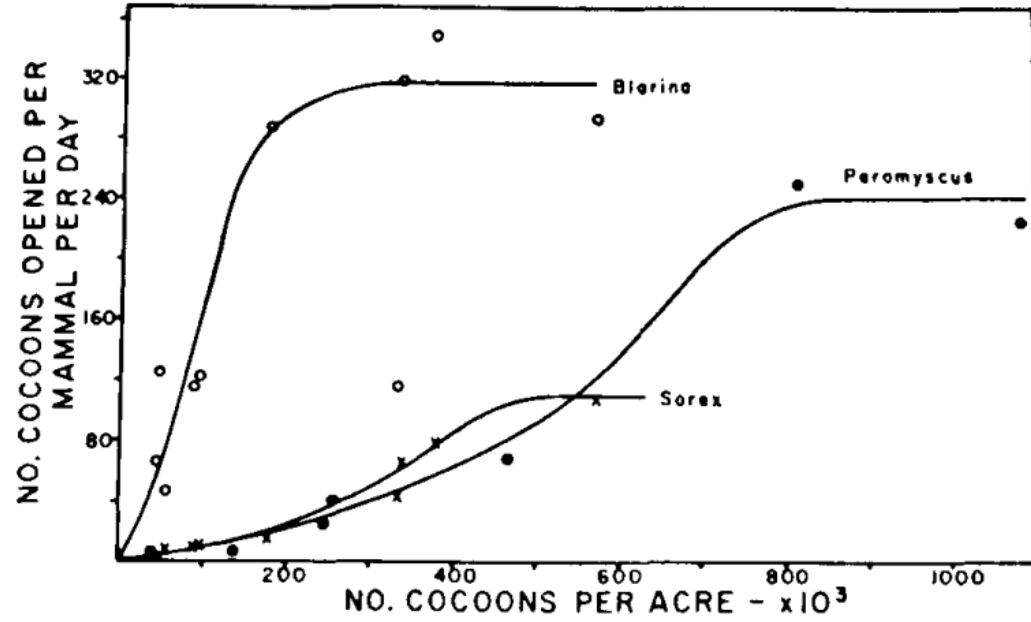
^b*Centre for Mathematical Biology, Mathematical Institute, Oxford University, 24–29 St Giles', Oxford OX1 2UT, Oxfordshire, UK*

Received 21 May 2003

$$\frac{dx}{dt} = \alpha x - \beta xy,$$

$$\frac{dy}{dt} = \delta xy - \gamma y,$$





Holling 1959

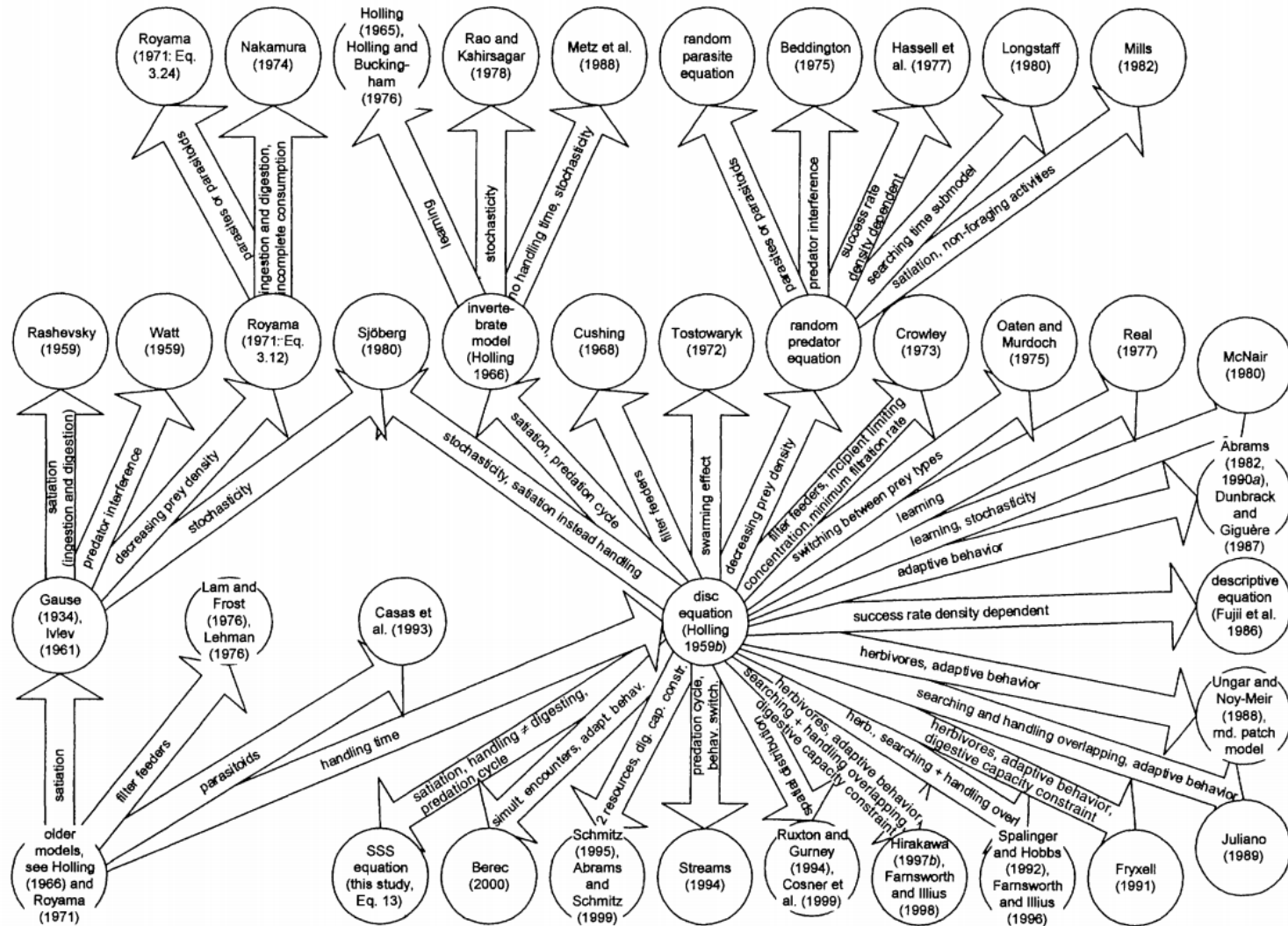
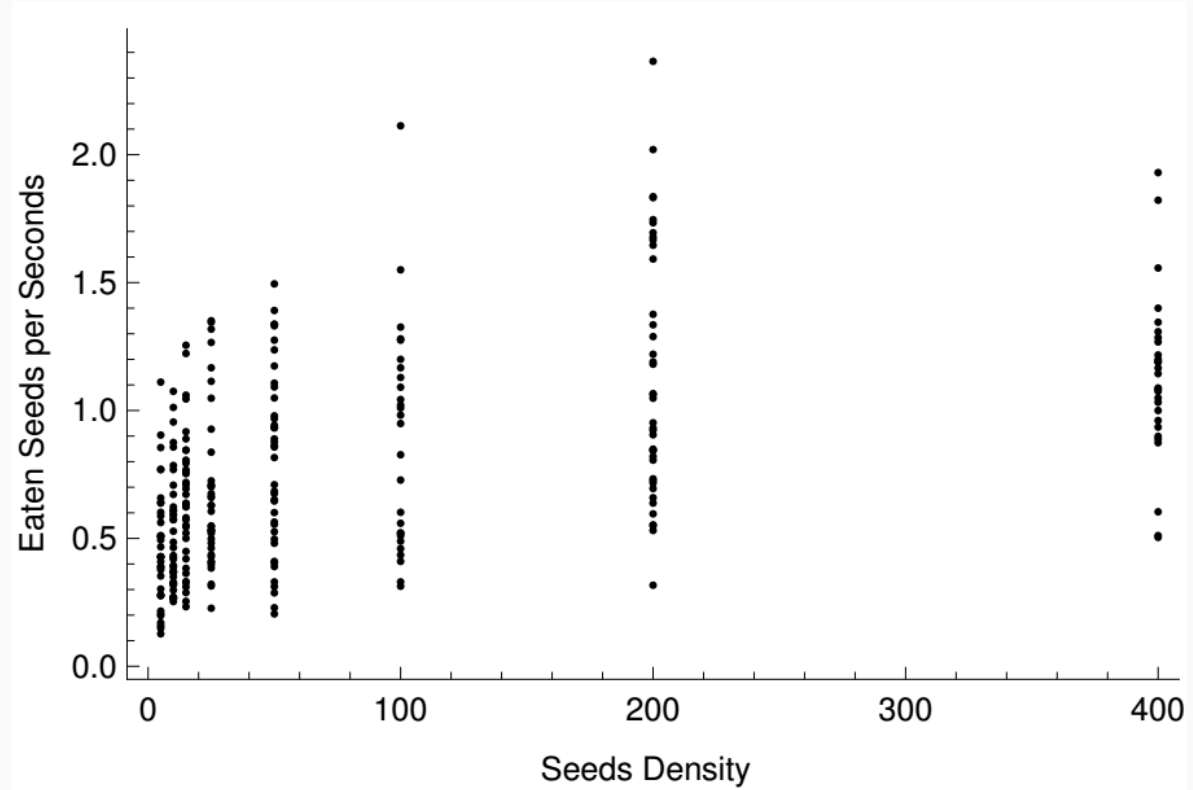


FIG. 1. A “family tree” of functional response models.



Baker et al. 2010

Justification mécanistique

Causalité statistique

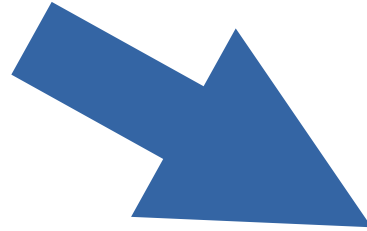
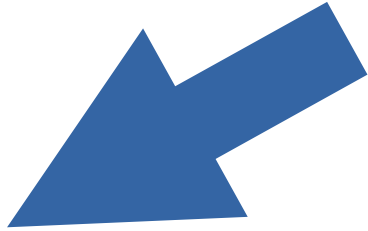
Forme

Propriété statistique

**Comportement des individus
dans un environnement**
(mouvement, manipulation,
interactions, etc.)



Réponse fonctionnelle
(taux de consommation par
individu par unité de temps.)



Rejuvenating functional responses with renewal theory

Sylvain Billiard¹, Vincent Bansaye² and J.-R. Chazottes³

$$R(x, y, z) \stackrel{\text{law}}{\approx} \frac{1}{\mathbb{E}(T(x, y, z))} + \mathcal{N} \left(0, \frac{1}{\Delta} \frac{\text{Var}(T(x, y, z))}{\mathbb{E}(T(x, y, z))^3} \right)$$

STOCHASTIC PREDATION MODEL WITH DEPLETION

GUY L. CURRY and RICHARD M. FELDMAN

Biosystems Research Division, Industrial Engineering Department, Texas A&M University,
College Station, Texas 77843

Abstract

Can. Ent. 111: 465–470 (1979)

$$N_e(t) = N_o (1 - e^{-\lambda(t - N_e(t)t_h)})$$

$$N_e(t) = \sum_{n=1}^{N_o} G_n(t)$$

$$F_t^{(d)} \sim \begin{cases} \sqrt{\frac{\pi}{8\Delta}} \frac{1}{\sqrt{t}} W^{(1)} & \text{if } d = 1 \\ \frac{\pi}{\Delta \ln t / \Delta} \left(1 - \frac{1}{\ln t / \Delta} (\gamma + \pi \tau_h / \Delta + 2\pi W^{(2)}) \right) & \text{if } d = 2 \\ \frac{\kappa}{\Delta + \kappa \tau_h} + \sqrt{\frac{\ln t / \Delta}{t \Delta (1 + \kappa \tau_h / \Delta)^3}} W^{(3)} & \text{if } d = 3 \end{cases}$$

(travail en cours avec JR Chazottes, G. Berthelot, E. Vergu)

$$F_t^{(d)} \sim \begin{cases} \sqrt{\frac{\pi}{8\Delta}} \frac{1}{\sqrt{t}} W^{(1)} & \text{if } d = 1 \\ \frac{\pi}{\Delta \ln t / \Delta} \left(1 - \frac{1}{\ln t / \Delta} \left(\gamma + \pi \tau_h / \Delta + 2\tau W^{(2)} \right) \right) & \text{if } d = 2 \\ \frac{\kappa}{\Delta + \kappa \tau_h} + \sqrt{\frac{\ln t / \Delta}{t \Delta (1 + \kappa \tau_h / \Delta)^3}} W^{(3)} & \text{if } d = 3 \end{cases}$$

$\max B_s^{(1)} - \min B_s^{(1)}$

$\mathcal{N}(0, \eta^2)$

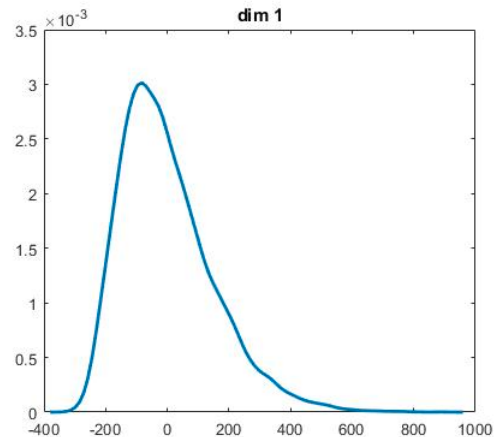
(travail en cours avec JR Chazottes, G. Berthelot, E. Vergu)

Quantifier les 4 sources de variabilité

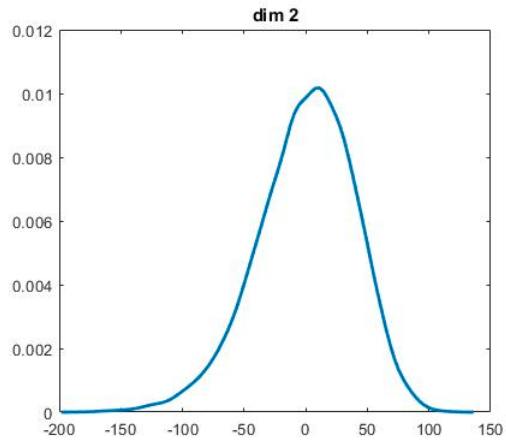
- $R_{ijk} | \psi_i \sim \mathcal{N}(m_{ij}, s_{ij}^2)$

with $m_{ij} = \frac{\sqrt{d_j - 1}}{C_1 \lambda_i + h_i (\sqrt{d_j - 1})}$ and $s_{ij}^2 = \frac{1}{\Delta} \frac{C_2 \lambda_i^2 (\sqrt{d_j - 1})}{(C_1 \lambda_i + h_i (\sqrt{d_j - 1}))^3}$

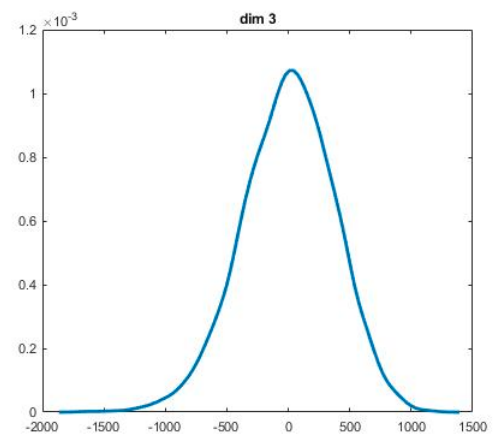
1D



2D

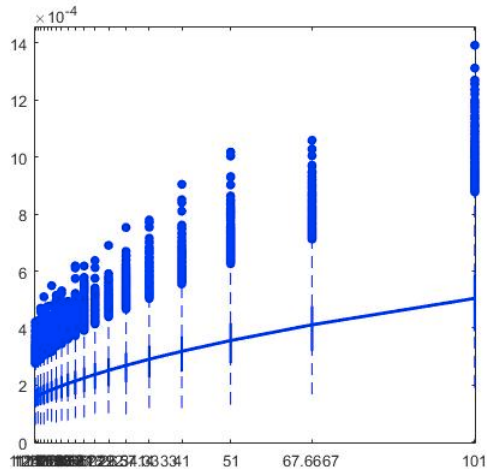


3D

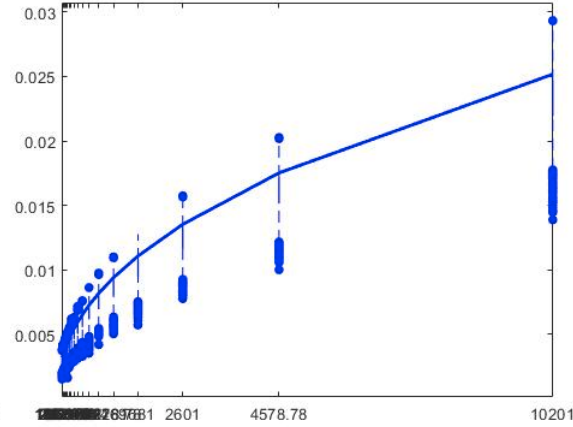


(simulations par Geoffroy Berthelot)

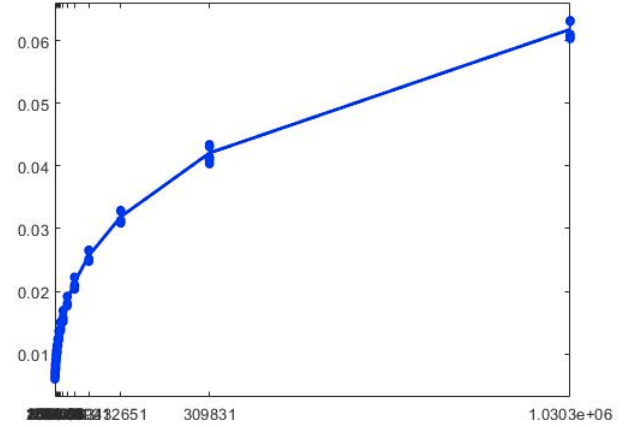
1D



2D

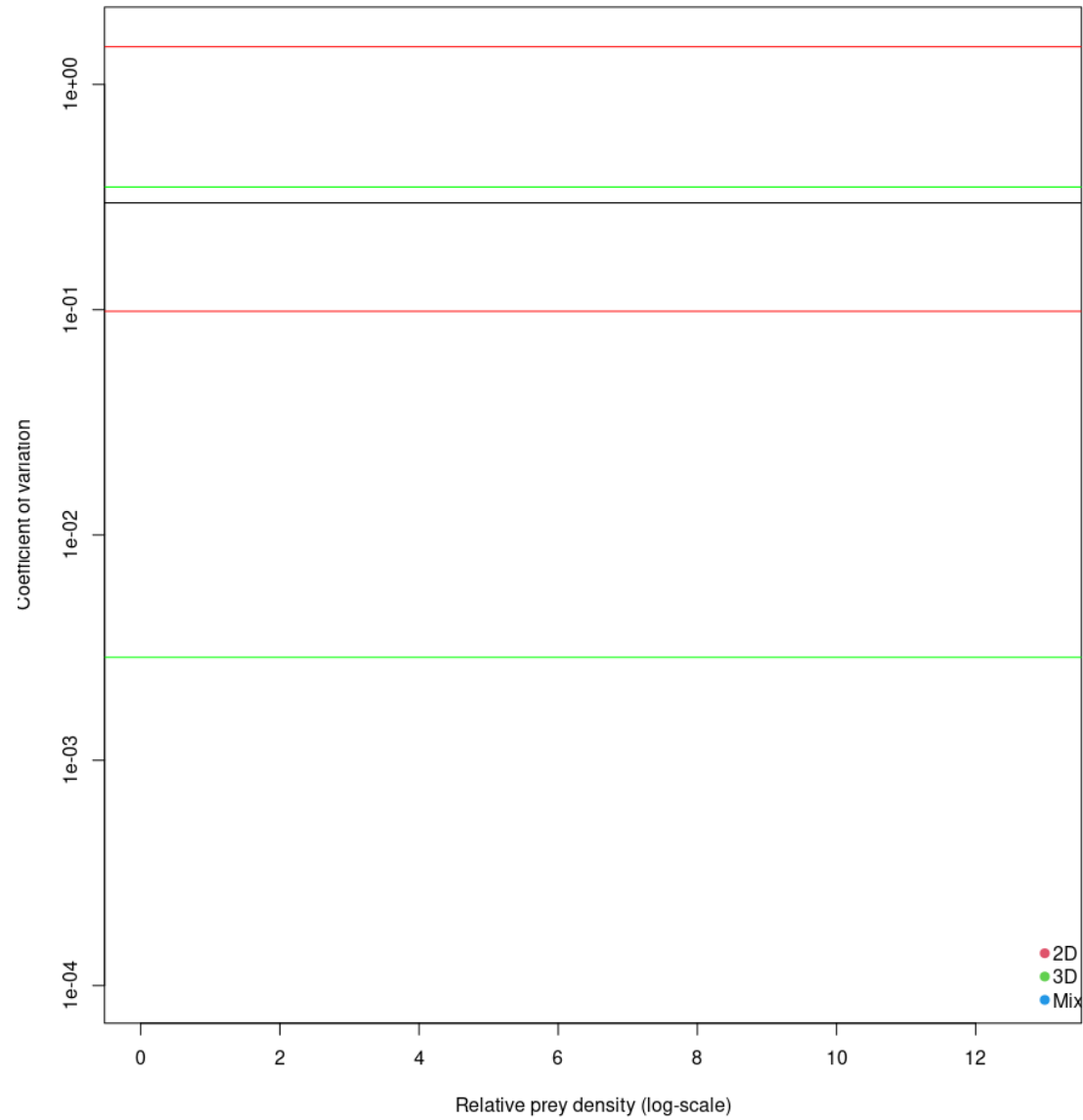


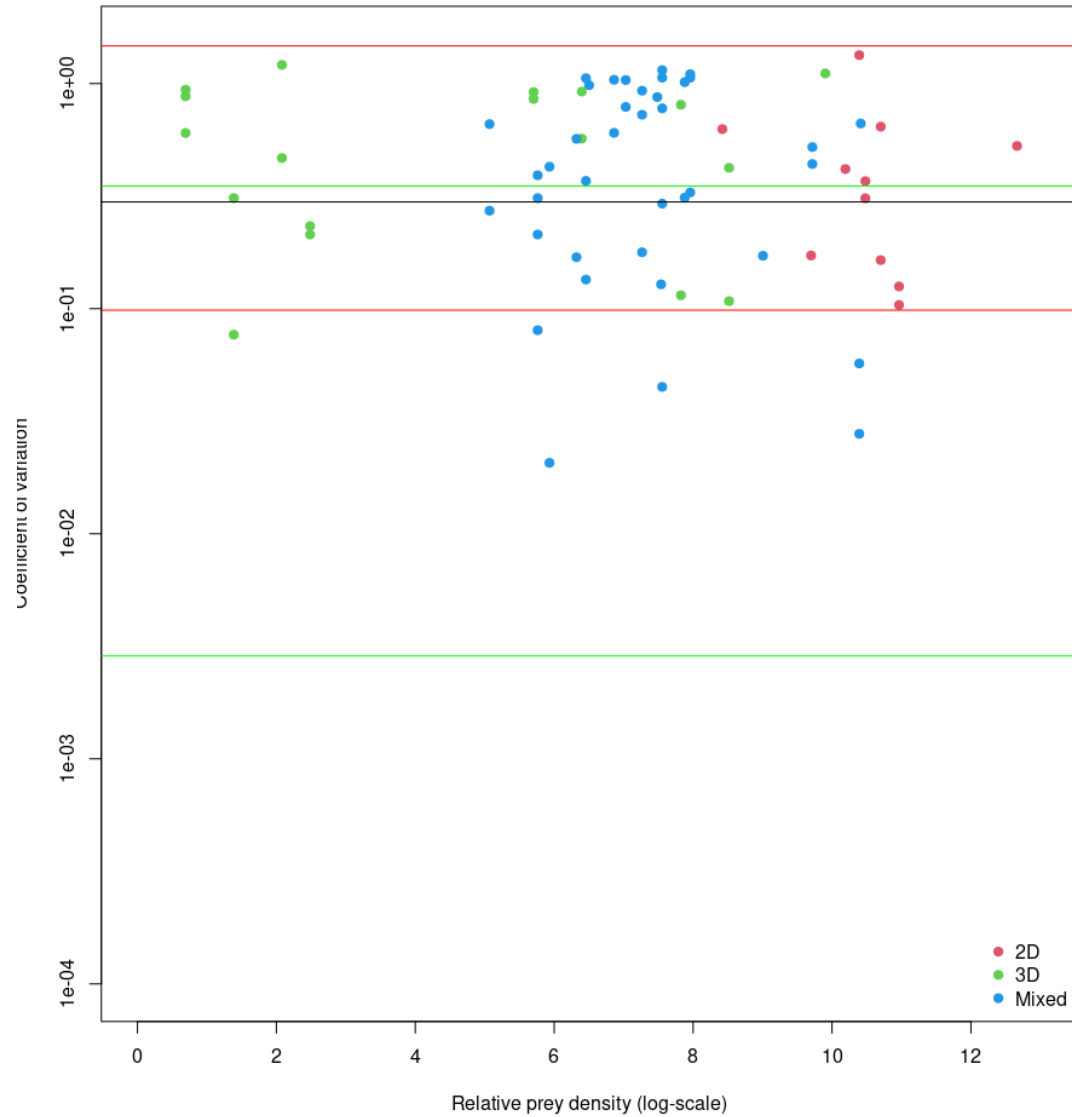
3D



(simulations par Geoffroy Berthelot)

Dimension	Expectation	Variance	Coefficient of variation
1D	$2\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{t\Delta}}$	$4 \frac{\ln 2 - 2/\pi}{t\Delta}$	$\sqrt{\frac{\pi}{2} \ln 2 - 1} \simeq 0.298$
2D	$\frac{\pi/\Delta}{\ln t/\Delta} \left(1 - \frac{\gamma + \tau_h \pi/\Delta}{\ln t/\Delta}\right)$	$\frac{2K\pi^2}{\Delta^2 (\ln t/\Delta)^4}$	$\frac{\sqrt{2K}}{\ln t/\Delta - \gamma - \tau_h \pi/\Delta} \simeq \frac{1.3}{\ln t/\Delta}$ <p>asuming that $\ln t \gg \tau_h$ i.e. total foraging time is much higher than the time taken to handle one prey</p>
3D	$\frac{\kappa}{\Delta + \kappa\tau_h}$	$\frac{27\kappa^4}{2\pi^2} \frac{\ln t/\Delta}{t\Delta(1 + \kappa\tau_h/\Delta)^3}$	$\frac{3\sqrt{3}}{\sqrt{2\pi}} \frac{\kappa}{\sqrt{1 + \kappa\tau_h/\Delta}} \sqrt{\frac{\ln t/\Delta}{t/\Delta}}$





[Data paper: FoRAGE \(Functional Responses from Around the Globe in all Ecosystems\) database: a compilation of functional responses for consumers and parasitoids](#)

Stella F. Uiterwaal, Ian T. Lagerstrom, Shelby R. Lyon, John P. DeLong

School of Biological Sciences, University of Nebraska – Lincoln, Lincoln, NE 68588 USA

