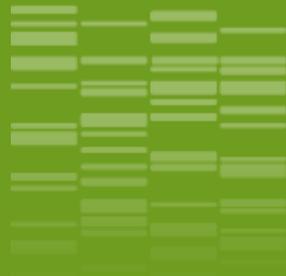


## Apports de la modélisation pour l'étude des interactions microbiennes

*J. Harmand, SAMI, LBE-INRA*

# CONTEXTE

- ❖ Quelques mots sur le LBE
- ❖ Problèmes de classification « fonctionnelle d'espèces »
- ❖ Approches combinatoires « a posteriori » pour l'étude des interactions entre microorganismes dans les écosystèmes microbiens
- ❖ Exemples
- ❖ Perspectives



01

Quelques mots sur le LBE

# Le LBE



Le site du Quatourze  
à Narbonne



Parc Méditerranéen de  
l'Innovation à Narbonne



Le site SUPAGRO Montpellier  
(groupe ELSA)

**WEB OF SCIENCE™**

Search

**Results: ...**  
(from Web of Science Core Collection)

You searched for: TOPIC: (anaerobic digestion) ...More

Create Alert

Organizations-Enhanced Refine Exclude Cancel Sort these by: Re

The first 100 Organizations-Enhanced (by record count) are shown. For advanced refine options, u

<input type="checkbox"/> INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE INRA (313)	<input type="checkbox"/> UNIVERSITY OF JYVASKYLA (65)
<input type="checkbox"/> WAGENINGEN UNIVERSITY RESEARCH CENTER (215)	<input type="checkbox"/> UNIVERSITY OF SOUTHAMPTON (63)

1<sup>er</sup> laboratoire au monde publiant sur la méthanisation

Equipe de Direction



Le collectif du **Lbe** au 1<sup>er</sup> janvier 2017



SERVICES COMMUNS



OBJETS THÉMATIQUES DE RECHERCHE

80 ETPs,  
36 INRA

**Groupe elsa**  
Responsable: Inesad ELIAS  
Analyse de Cycle de Vie, Environnement, Economie circulaire

**UMR mistea**  
Responsable: Pascal FERRY  
Systèmes agricoles, Modèles agroécologiques, Développement durable, Sécurité alimentaire, Sécurité de l'Alimentation

**EPI BioCo<sub>2</sub>re**  
Responsable: Inesad ELIAS  
Modélisation, Caractérisation, Diagnostic des systèmes, Optimisation de production, Bioéconomie

**Système, Analyse, Modélisation et Informatique SAMI**

**Caractérisation de la Matière complexe et Ingénierie des bioProcédés MaXIP**

**Diversité et Ingénierie des Ecosystèmes Microbiens DIEM**

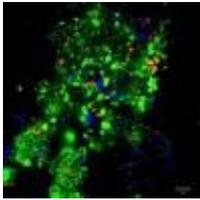
**Développement Technologique et Innovation DTI**

**Inra Transfert Environnement**

PARTENAIRES HÉBERGÉS

# Objets d'étude et compétences du LBE

## Deux échelles d'étude



*Ecosystèmes*



*Bioprocédés*



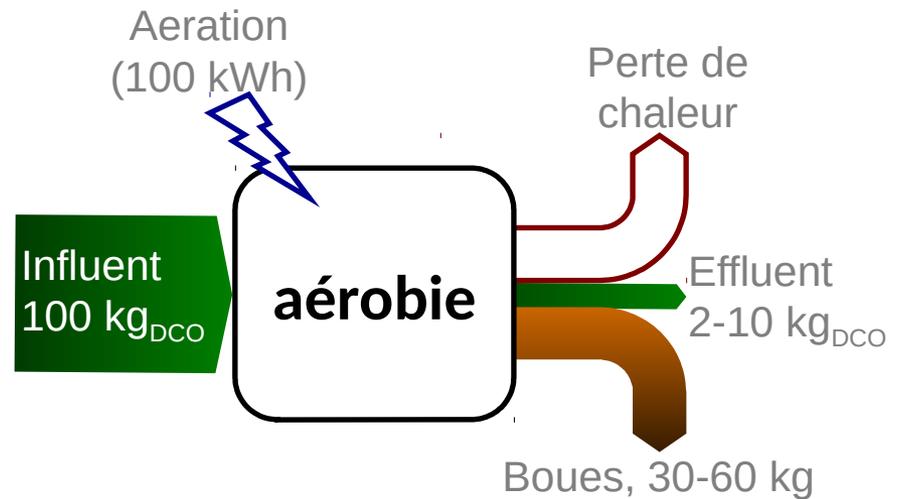
## Approche interdisciplinaire



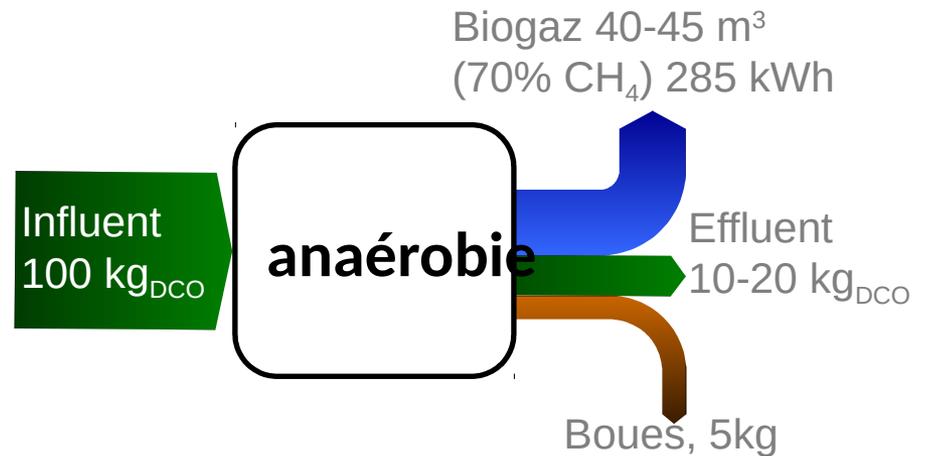
# Le traitement des eaux et des déchets

Hier et aujourd'hui

## Traitement aérobie

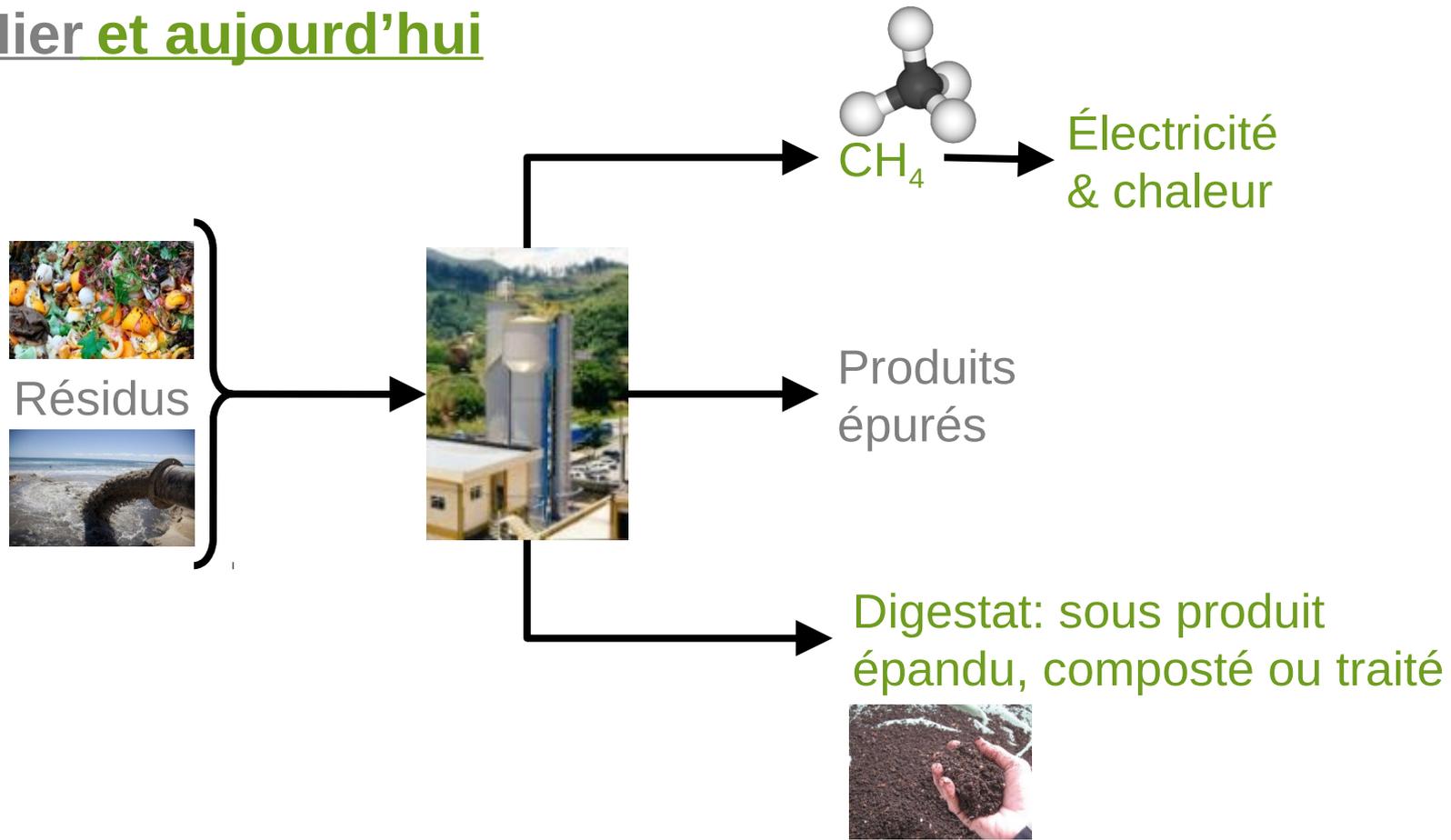


## Traitement anaérobie



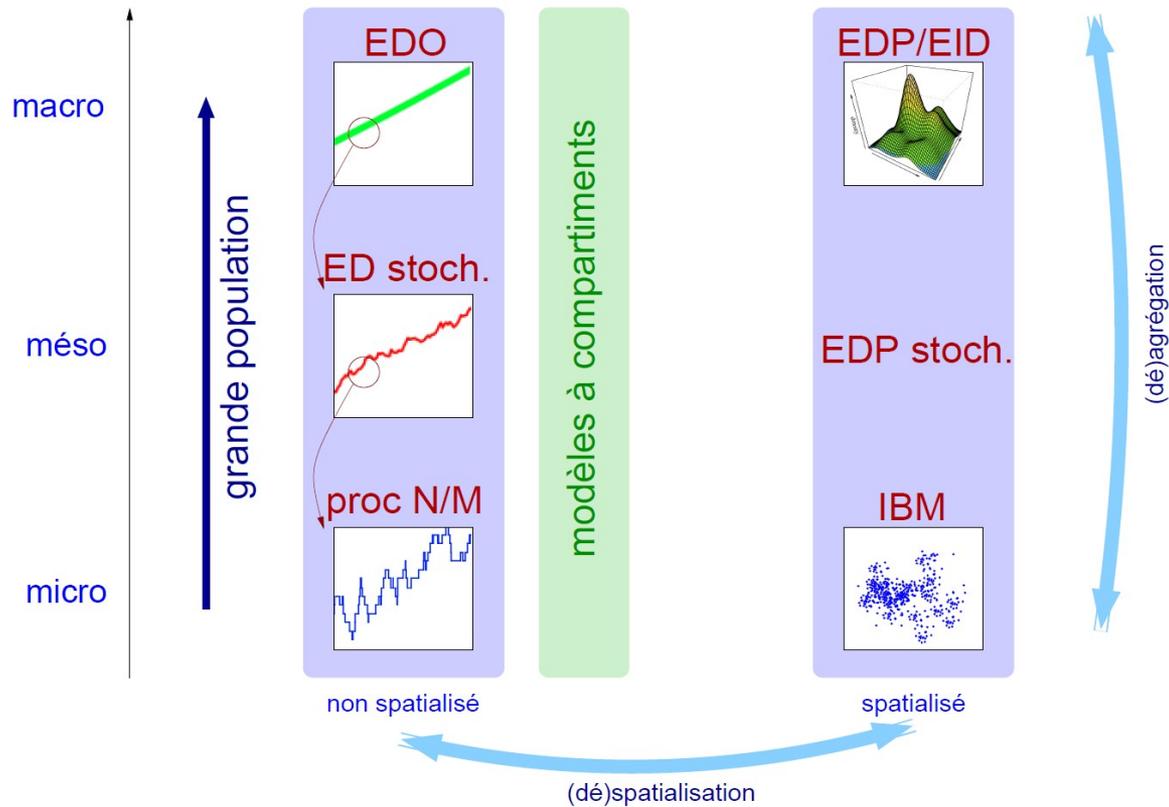
# La méthanisation ou digestion anaérobie

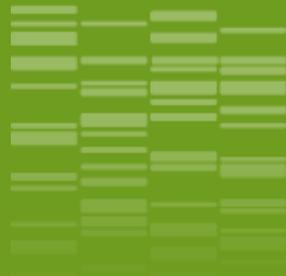
## Hier et aujourd'hui



# Pas d'a priori concernant la modélisation

- Modélisation (mise en équations) statique/dynamique, déterministe et/ou stochastique, informatique...

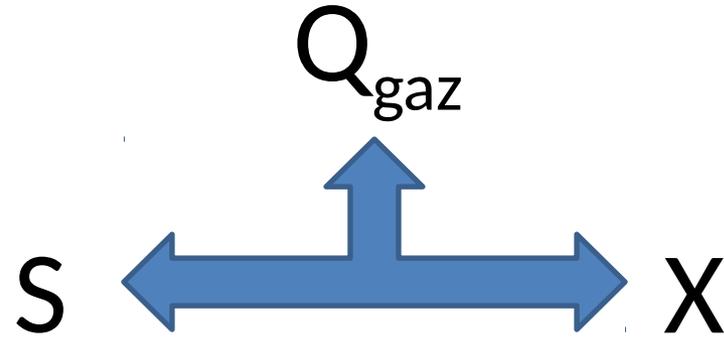




02

## Problèmes de classification « fonctionnelle d'espèces »

# Réacteurs biologiques



# Modélisation des réacteurs biologiques

$$\begin{cases} \frac{dS}{dt} = -S \frac{Q_{in}}{V} + \frac{Q_{in}}{V} S_{in} - \frac{\mu(S)}{Y} X \\ \frac{dX}{dt} = -X \frac{Q_{in}}{V} + \mu(S) X \\ \frac{dV}{dt} = Q_{in} - Q_{out} \end{cases}$$

# Modélisation des réacteurs biologiques

- Le chemostat (système continu homogène)

$$\left\{ \begin{array}{l} \dot{X} = (\mu(\cdot) - D)X \\ \dot{S} = (S_{in} - S)D - \frac{\mu(\cdot)}{Y}X \end{array} \right. \quad \left\{ \begin{array}{l} \dot{X}_i = (\mu_i(\cdot) - D)X_i \\ \dot{S} = (S_{in} - S)D - \sum_i \frac{\mu_i(\cdot)}{Y_i}X_i \end{array} \right.$$

- Le batch (réacteur dit « fermé » homogène)

$$\left\{ \begin{array}{l} \dot{X} = \mu(\cdot)X \\ \dot{S} = -\frac{\mu(\cdot)}{Y}X \end{array} \right. \quad \left\{ \begin{array}{l} \dot{X}_i = \mu_i(\cdot)X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(\cdot)}{Y_i}X_i \end{array} \right.$$

# Pb 1 : « Efficacité d'espèces au sein d'un écosystème complexe continu »

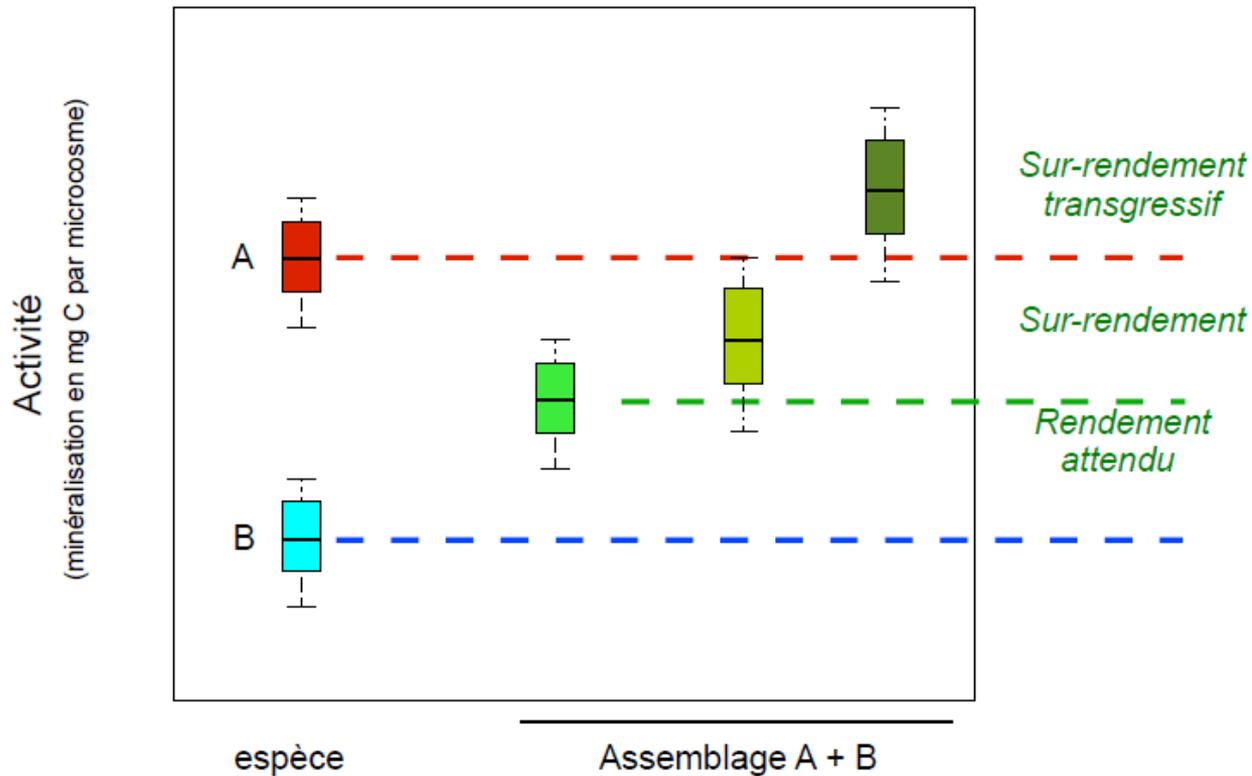
Etant donné un « indice de performance » (ici,  $\mu(\cdot)X$  appelée « l'activité ») et un mode de fonctionnement, on veut savoir quelle espèce (ou mélanges d'espèces) est(sont) le(les) plus efficace(s)...

Exemple 1 : dans un écosystème marin en chemostat, quelles espèces croissent le plus rapidement (!)?

$$\left\{ \begin{array}{l} \dot{X}_i = (\mu_i(\cdot) - D)X_i \\ \dot{S} = (S_{in} - S)D - \sum_i \frac{\mu_i(\cdot)}{Y_i} X_i \end{array} \right. \quad \begin{array}{l} \bullet \mu_i \text{ à } T? \\ \bullet SQ_{gaz}(T) = \int_0^T \sum_i k_i \mu_i(\cdot) X_i(\tau) d\tau \end{array}$$

# Pb 2 : « Pb de sur-rendement en batch »

Exemple 2 : étant donné un ensemble d'espèces, quels mélanges de ces espèces sont « les plus efficaces » ?



Pour des conditions initiales identiques (mêmes ressource et biomasse totales)!

## Pb 2 : « Pb de sur-rendement »

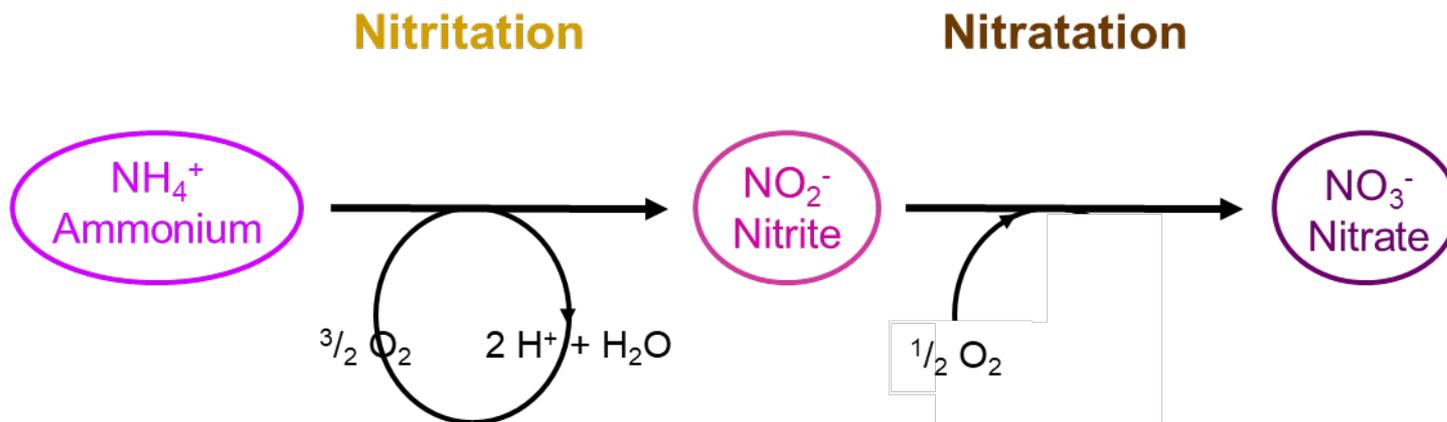
Même question mais modèles différents!

$$\left\{ \begin{array}{l} \dot{X}_i = \mu_i(S)X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(S)}{Y_i} X_i \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} \dot{X}_i = \mu_i(\cdot)X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(\cdot)}{Y_i} X_i \end{array} \right.$$

Pas de sur-rendement possible par rapport à l'activité...

$$\left\{ \begin{array}{l} \dot{X}_i = (\mu_i(\cdot) - k_i)X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(\cdot)}{Y_i} X_i \end{array} \right.$$

# Pb 3 : « Classement fonctionnel »



**Bactéries ammoniac oxydantes (AOB)**  
*Nitrosomonas, Nitrosococcus, Nitrospira*

**Bactéries nitrite oxydantes (NOB)**  
*Nitrobacter, Nitrococcus, Nitrospira*

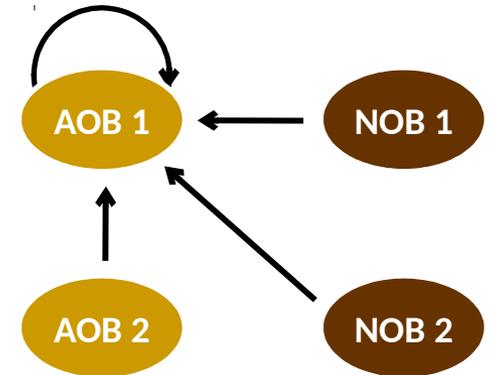
# Pb 3 : « Classement fonctionnel »

AOB (avec  $i=1;2$ )

$$\frac{dX_i}{dt} = (\mu_i(S_1) - D)X_i + \left( \sum_{j=1}^4 \alpha_{ij} X_j \right) X_i$$

NOB (avec  $i=3;4$ )

$$\frac{dX_i}{dt} = (\mu_i(S_2) - D)X_i + \left( \sum_{j=1}^4 \alpha_{ij} X_j \right) X_i$$



$\text{NH}_4^+$  res

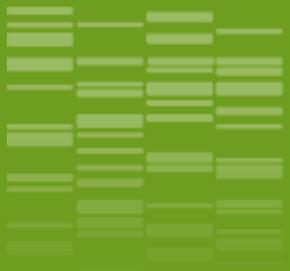
$$\frac{dS_1}{dt} = (S_{in} - S_1)D - \frac{1}{Y_A} \sum_{i=1}^2 (\mu_i(S_1) + \sum_{j=1}^4 \alpha_{ij} X_j) X_i$$

$\text{NO}_2^-$

$$\frac{dS_2}{dt} = \frac{1}{Y_A} \sum_{i=1}^2 (\mu_i(S_1) + \sum_{j=1}^4 \alpha_{ij} X_j) X_i - \frac{1}{Y_B} \sum_{i=3}^4 (\mu_i(S_2) + \sum_{j=1}^4 \alpha_{ij} X_j) X_i - S_2 D$$

$\text{NO}_3^-$

$$\frac{dS_2}{dt} = \frac{1}{Y_B} \sum_{i=3}^4 (\mu_i(S_2) + \sum_{j=1}^4 \alpha_{ij} X_j) X_i - S_3 D$$



# 03

Approches combinatoires « a posteriori » pour l'étude des interactions entre microorganismes dans les écosystèmes microbiens

# Classification combinatoire d'espèces au sein d'un écosystème

## Functional Ecology



British Ecological Society

*Functional Ecology* 2014

doi: 10.1111/1365-2435.12267

## Community assembly effects shape the biodiversity-ecosystem functioning relationships

Benoît Jaillard<sup>\*,1</sup>, Alain Rapaport<sup>2,3</sup>, Jérôme Harmand<sup>3,4</sup>, Alain Brauman<sup>5</sup> and Naoise Nunan<sup>6</sup>

<sup>1</sup>Inra, UMR1222 Ecologie fonctionnelle et Biogéochimie des Sols et Agrosystèmes (Eco&Sols), 2 place Pierre Viala, 34060 Montpellier, France; <sup>2</sup>Inra, UMR0729 Mathématiques, Informatique et STatistique pour l'Environnement et l'Agronomie (MISTEA), 2 place Pierre Viala, 34060 Montpellier, France; <sup>3</sup>INRA-INRIA MODEMIC Project-team, INRIA Sophia-Antipolis Méditerranée, Sophia-Antipolis 06902, France; <sup>4</sup>Inra, UR0050 Laboratoire de Biotechnologie de l'Environnement, avenue des Etangs, 11100 Narbonne, France; <sup>5</sup>IRD, UMR1222 Ecologie fonctionnelle et Biogéochimie des Sols et Agrosystèmes 2 place Pierre Viala, Montpellier 34060, France; and <sup>6</sup>CNRS, UMR7618 BioEMCo, Campus AgroParisTech, 781 n, France

# Classification combinatoire d'espèces au sein d'un écosystème



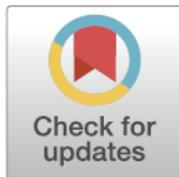
RESEARCH ARTICLE

## A combinatorial analysis using observational data identifies species that govern ecosystem functioning

**Benoît Jaillard<sup>1\*</sup>, Philippe Deleporte<sup>1</sup>, Michel Loreau<sup>2</sup>, Cyrille Violle<sup>3</sup>**

**1** Ecologie fonctionnelle et Biogéochimie des Sols et Agrosystèmes, CIRAD, INRA, IRD, SupAgro, Univ Montpellier, Montpellier, France, **2** Centre for Biodiversity Theory and Modelling, Theoretical and Experimental Ecology Station, CNRS and Paul Sabatier University, Moulis, France, **3** Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, Montpellier, France

\* [beno](#)



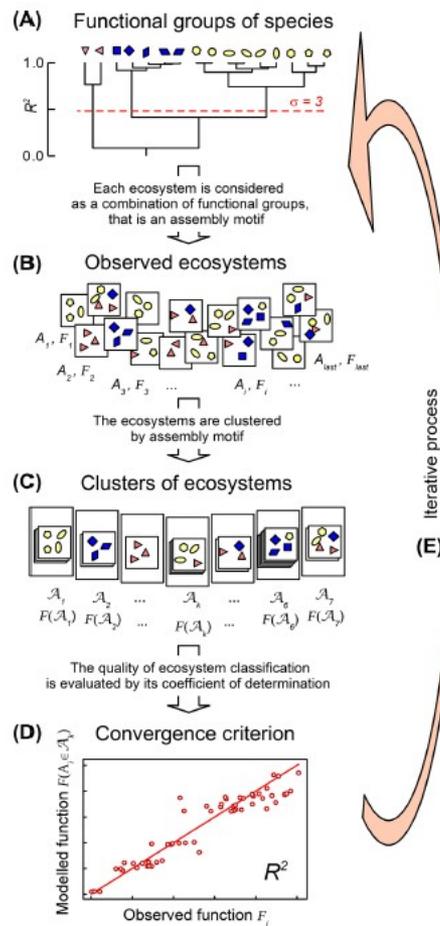
..

# Classification combinatoire d'espèces au sein d'un écosystème

Performance	Présence / absence de l'espèce i dans chaque mélange					
	Espèce 1	Espèce 2	Espèce 3	Espèce 4	Espèce 5	Espèce 6
0.3970	1.0000	0	0	0	0	0
1.3950	0	1.0000	0	0	0	0
0.4200	0	0	1.0000	0	0	0
0.7870	0	0	0	1.0000	0	0
0.4180	0	0	0	0	1.0000	0
0.3820	0	0	0	0	0	1.0000
1.3100	1.0000	1.0000	0	0	0	0
0.5530	1.0000	0	1.0000	0	0	0
0.9250	1.0000	0	0	1.0000	0	0
0.6720	1.0000	0	0	0	1.0000	0
0.6940	1.0000	0	0	0	0	1.0000
1.3400	0	1.0000	1.0000	0	0	0
1.4080	0	1.0000	0	1.0000	0	0
1.2330	0	1.0000	0	0	1.0000	0
1.2540	0	1.0000	0	0	0	1.0000
0.7180	0	0	1.0000	1.0000	0	0
0.4110	0	0	1.0000	0	1.0000	0
0.4020	0	0	1.0000	0	0	1.0000
0.6490	0	0	0	1.0000	1.0000	0
0.7450	0	0	0	1.0000	0	1.0000

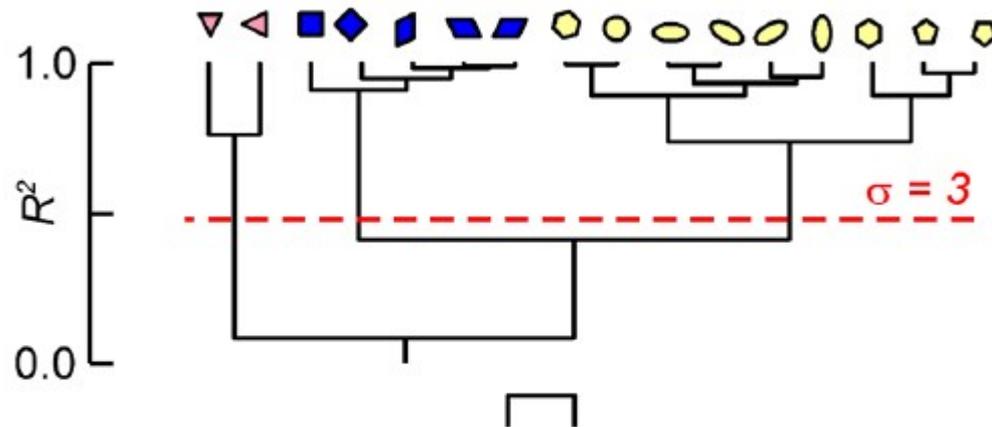
Je cherche un algorithme capable de prédire les performances d'un assemblage donné...

# Classification combinatoire d'espèces au sein d'un écosystème



# Classification combinatoire d'espèces au sein d'un écosystème

## (A) Functional groups of species

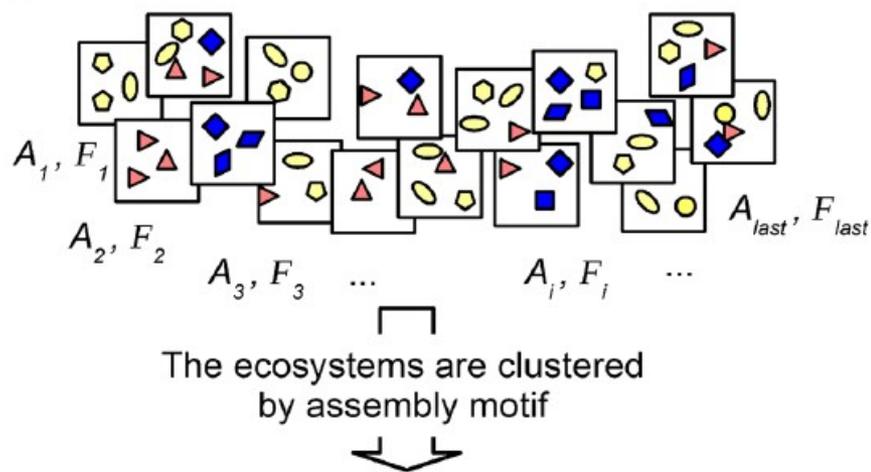


Each ecosystem is considered as a combination of functional groups, that is an assembly motif



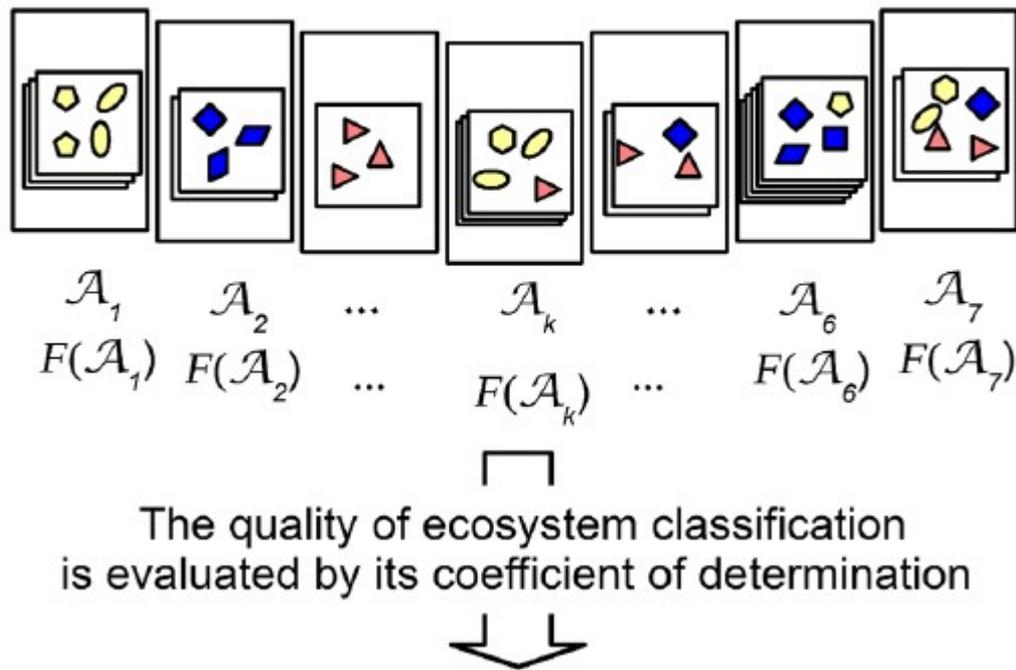
# Classification combinatoire d'espèces au sein d'un écosystème

## (B) Observed ecosystems

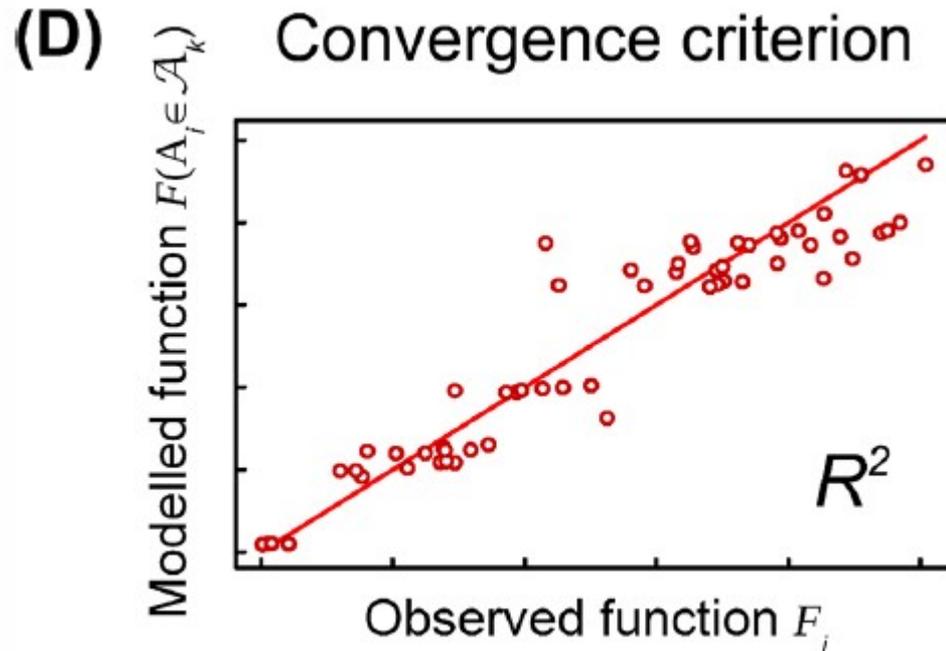


# Classification combinatoire d'espèces au sein d'un écosystème

## (C) Clusters of ecosystems



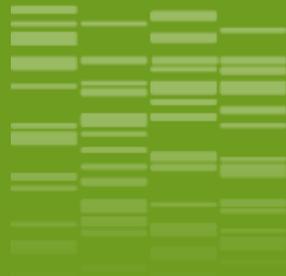
# Classification combinatoire d'espèces au sein d'un écosystème



# Classification combinatoire d'espèces au sein d'un écosystème

Algorithme « force brute »

- On assigne une classe à chaque espèce, par exemple A ou B (si deux classes)
- On regroupe les expériences disponibles en « motifs » comprenant soit des espèces de la classe A (motif 1), soit de la classe B (motif 2), soit des espèces de la classe A et de la classe B (motif 3)...
- A l'intérieur de chaque motif, on calcule la moyenne des performances et leur dispersion (ce qui dote l'approche d'une capacité prédictive!)
- On itère jusqu'à avoir testé la totalité des combinaisons possibles d'appartenance à des classes pour les espèces...
- On retient l'assignation qui minimise la somme des dispersions intra motifs



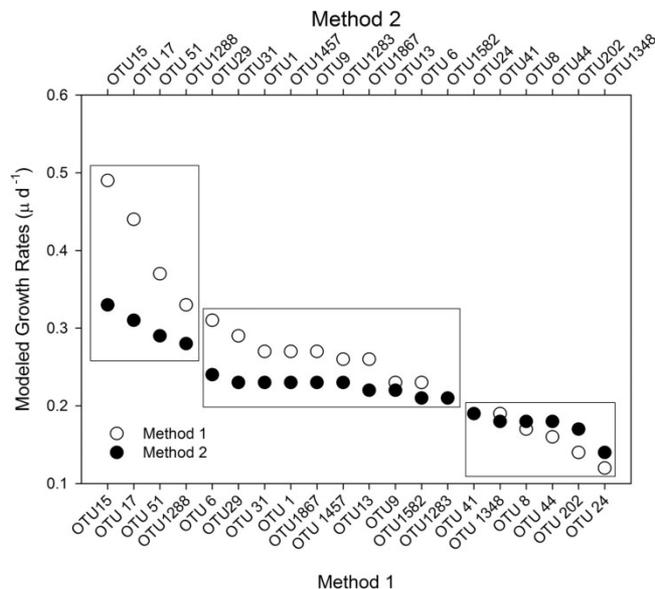
# 04 Exemples

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

## Major changes in the composition of a Southern Ocean bacterial community in response to diatom-derived dissolved organic matter

Marine Landa ✉, Stéphane Blain, Jérôme Harmand, Sébastien Monchy, Alain Rapaport, Ingrid Obernosterer Author Notes

*FEMS Microbiology Ecology*, Volume 94, Issue 4, 1 April 2018, fiy034,  
<https://doi.org/10.1093/femsec/fiy034>



# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

## Questions d'algorithmes

- Limitation combinatoire : si on retient 22 espèces, on a  $2^{22}$  manips possibles!!!!
- Classement supervisé vs non supervisé?

Pour cet exemple, on va :

- retenir 8 espèces tirées au hasard dans les 22 identifiées
- bruitez les données (on regarde l'activité en 5 instants correspondant à la valeur cumulée de l'activité sur 5 temps de séjour)
- considérer 2 ou 3 classes en considérant un nombre décroissant de données
- pour chaque classement, utiliser l'algorithme non supervisé

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

*Modèle pour générer les données*

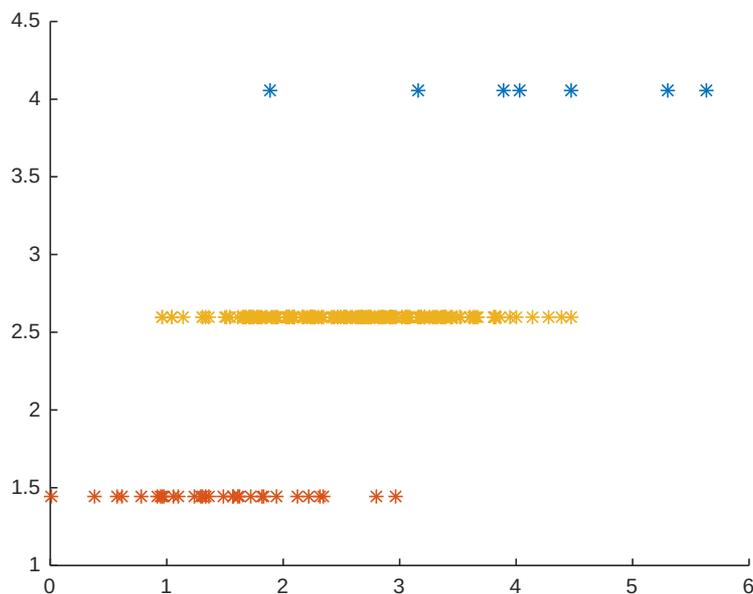
$$\begin{cases} \dot{X}_i &= (\mu_i(\cdot) - D)X_i \\ \dot{S} &= (S_{in} - S)D - \sum_i \frac{\mu_i(\cdot)}{Y_i} X_i \end{cases}$$

Et on suppose que l'on mesure l'activité au cours du temps (avec  $k_i=1/Y_i$ ) :

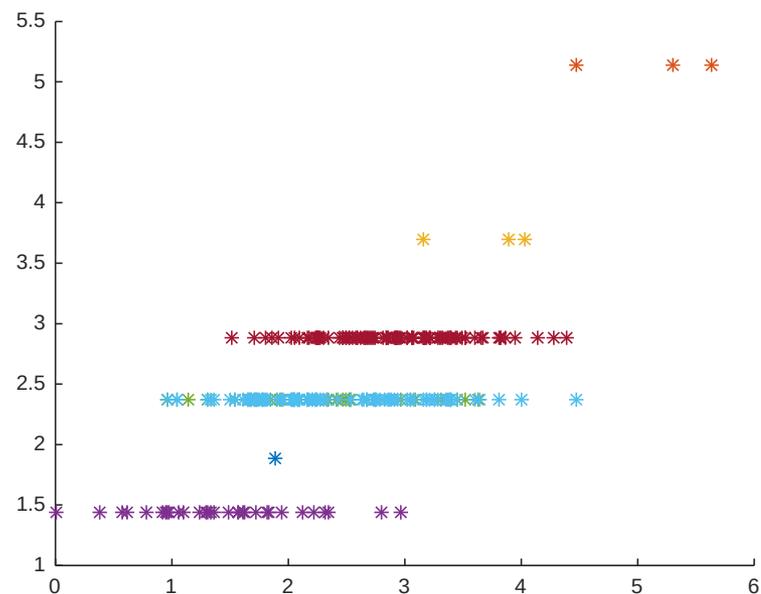
$$\int_0^T \sum_i k_i \mu_i(\cdot) X_i(\tau) d\tau$$

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

## Résultats de classement



$T=t_2$ , 2 classes (3 motifs)



$T=t_2$ , 3 classes (7 motifs)

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

Résultats de classement (nbre d'expériences « non limitant »)

Nombre d'espèces	Nombre de classes	Nombre de motifs	Nombre de manipis dispo pour la classification	Tf	Meilleur classement	R2
8	2	3	255 / 100%	1	00000101	0.23
				2	00000111	0.28
				3	00000111	0.35
				4	00000111	0.43
				5	00000111	0.51
Nombre d'espèces	Nombre de classes	Nombre de motifs	Nombre de manipis dispo pour la classification	Tf	Meilleur classement	R2
8	3	7	255 / 100%	1	01111212	0.36
				2	00000121	0.41
				3	00000112	0.48
				4	00000112	0.56
				5	00001122	0.64

0.0001 0.0205 0.0237 0.0294 0.0343 0.0372 0.0406 0.0423

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

*Résultats de classement : robustesse par rapport au nombre d'expériences disponibles*

Nombre d'espèces	Nombre de classes	Nombre de motifs	Nombre de manip utilisées pour la classification	Tf	Meilleur classement	R2	Rappel en utilisant toutes les expériences possibles	Digit différents
8	2	3	125	1	00000001	0.23	00000101	1
8	2	3	125	5	00000111	0.51	00000111	0
8	3	7	125	1	00111212	0.44	01111212	1
8	3	7	125	5	00000112	0.64	00001122	2
8	2	3	25	1	00001111	0.40	00000101	2
8	2	3	25	5	00110000	0.53	00000111	5
8	3	7	25	1	01100200	0.69	01111212	4
8	3	7	25	5	01000222	0.84	00001122	3

**Attention : les résultats dépendent maintenant des données disponibles!**

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

Résultats de classement : robustesse par rapport aux CI avec  $\Sigma(X_0)=Cst$

Nombre d'espèces	Nombre de classes	Nombre de motifs	Nombre de manips utilisées pour la classification	Tf	Meilleur classement	R2	Rappel en utilisant toutes les expériences possibles	Digits différents	Rappel avec données bruitées
8	2	3	255 / 100%	1	00001111	0.44	-	-	00000 <b>101</b>
8	2	3	255 / 100%	5	00000001	0.71	-	-	00000111
8	3	7	255 / 100%	1	00111112	0.73	-	-	01111212
8	3	7	255 / 100%	5	00001112	0.92	-	-	00001122
8	2	3	125 / 50%	1	00111111	0.47	00001111	2	00000001
8	2	3	125 / 50%	5	00000001	0.71	00000001	0	00000111
8	3	7	125 / 50%	1	00111112	0.73	00111112	0	00111 <b>2</b> 12
8	3	7	125 / 50%	5	00001112	0.94	00001112	0	00000112
8	2	3	25 / 10%	1	00111111	0.75	00001111	2	00001111
8	2	3	25 / 10%	5	00000001	0.73	00000001	0	00 <b>11</b> 0000
8	3	7	25 / 10%	1	00001 <b>102</b>	0.76	00111112	3	<b>01100200</b>
8	3	7	25 / 10%	5	000 <b>10</b> 112	0.95	00001112	2	<b>01000222</b>

**Attention : les résultats dépendent maintenant des données disponibles!**

# Exemple 1 : Classification d'espèces en fonction de leur vitesse de croissance

*Si on considère la qualité du classement sur la base d'un ordonnancement des taux de croissance, on note :*

- Une assez grande robustesse du classement par rapport au bruit (CI du système dynamique pour générer les données)
- Une grande robustesse de l'approche par rapport au nombre de données disponibles : recherche des « meilleurs types » de données (en particulier s'il manque des expériences pour la prédiction des performances d'un motif!)

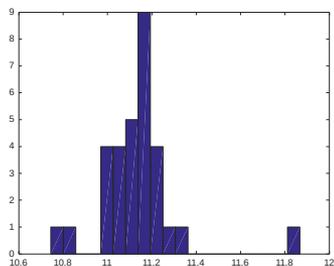
# Exemple 2 : Classification fonctionnelle d'espèces

## *Ecosystème considéré*

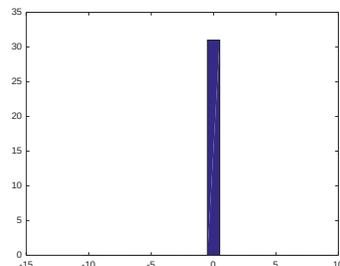
- 10 espèces (dont 5 AOB et 5 NOB) –  $0.7 < \mu_{max} < 0.9$ ,  $0.1 < K_s < 0.2$ ,  $-1 < \alpha_{hij} < 1$
- Toutes ou partie des manip
- CI non bruitées
- Simulation en batch sur un temps qui correspondrait à 5 TSH en continu
- Modèle incluant mortalité et interactions de type LVG
- Mesure de l'activité nitratante (seconde étape de la réaction)

# Exemple 2 : Classification fonctionnelle d'espèces

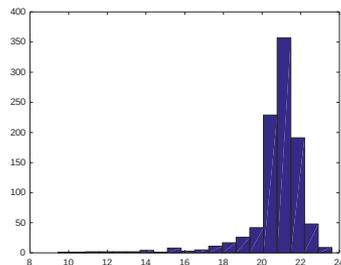
Résultats de classement (Utilisation de toutes les expériences possibles - exemple)



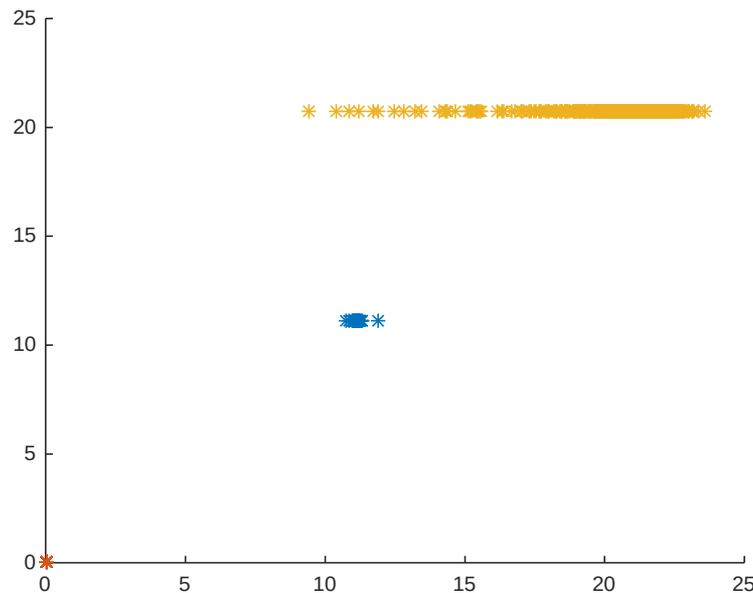
Distribution des expériences classées dans le motif {0 1} (activité du transitoire)



Distribution des expériences classées dans le motif {1 0} (absence de la seconde étape = aucune activité)



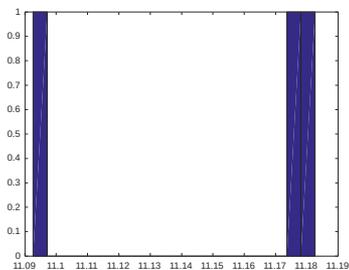
Distribution des expériences classées dans le motif {1 1}



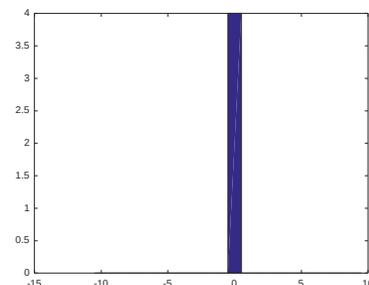
0000011111

# Exemple 2 : Classification fonctionnelle d'espèces

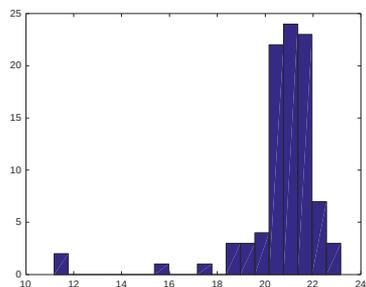
Résultats de classement (Utilisation de 10% des expériences possibles)



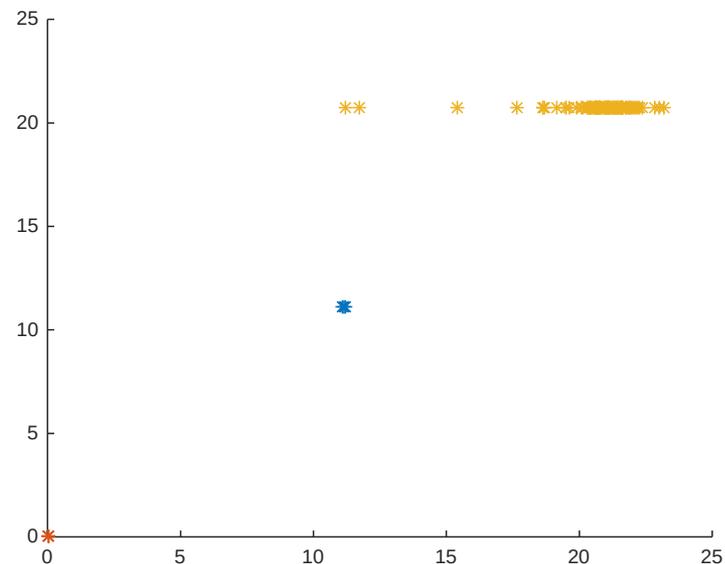
Distribution des expériences classées dans le motif {0 1} (activité du transitoire)



Distribution des expériences classées dans le motif {1 0} (absence de la seconde étape = aucune activité)



Distribution des expériences classées dans le motif {1 1}



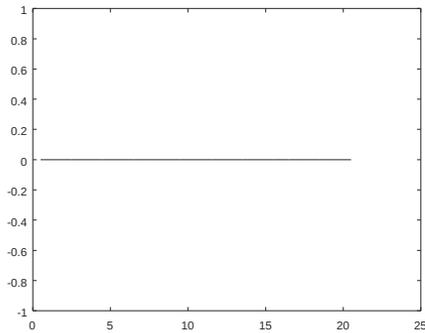
0000011111

## Exemple 2 : Classification fonctionnelle d'espèces

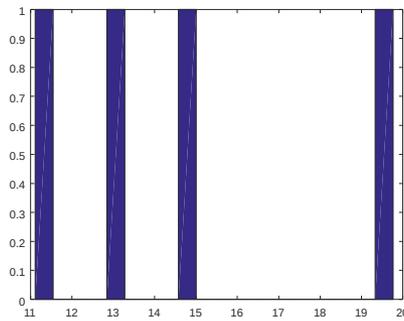
Nombre d'espèces	Nombre de classes	Nombre de motifs	Nombre de manip utilisées pour la calssification	Meilleur classement	R2
10	2	3	1023 / 100%	0000011111	0.94
			500 / 50%	0000011111	0.87
			50 / 5%	0000011111	0.83
			20 / 2%	0000011111	0.98
			20 / 2%	0011010000	0.93

# Exemple 2 : Classification fonctionnelle d'espèces

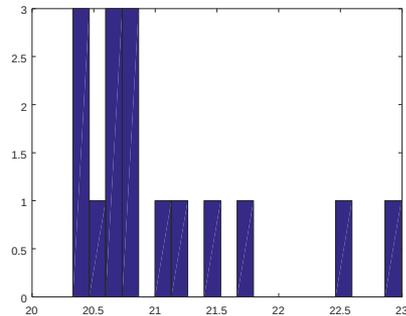
## Résultats de classement (Utilisation de 2% des expériences possibles)



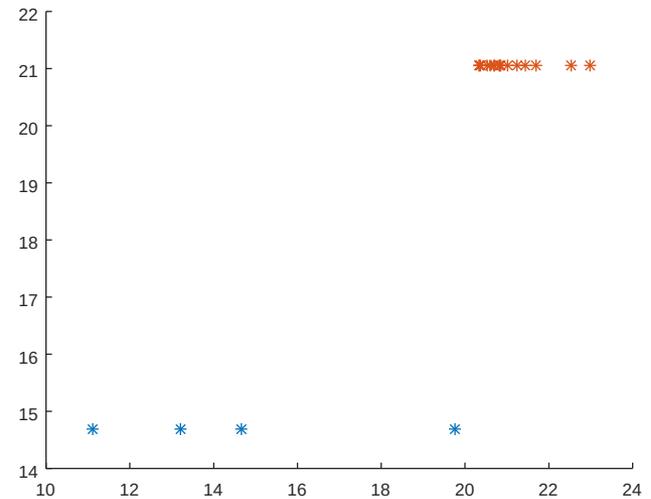
Distribution des expériences classées dans le motif {0 1} (activité du transitoire)



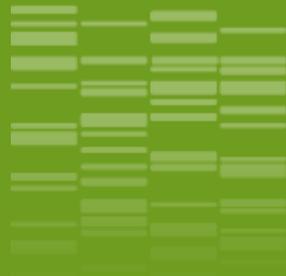
Distribution des expériences classées dans le motif {1 0} (absence de la seconde étape = aucune activité)



Distribution des expériences classées dans le motif {1 1}



0110100000



# 05 Perspectives

# Perspectives

- Approche efficace (performante et robuste) pour classer les espèces d'un écosystème de manière fonctionnelle
- Résultats dépendants des données disponibles...et de leur qualité!
- Les résultats suggèrent qu'il pourrait être « mieux » de réaliser le maximum d'expérimentations de mélanges d'un nombre d'espèces = nbre de classes (surtout si le nbre de classes devient important)... Stratégies optimales en termes d'information? Stratégies adaptatives en fonction des expérimentations disponibles?
- Peut-on se passer d'un algorithme non supervisé comme suggéré par Jaillard *et coll.*, 2018? OBLIGATOIRE!
- Application bien au-delà des assemblages microbiens (ex. mélange de substrats en codigestion)



Merci pour votre attention!