



**HAL**  
open science

# Contribution of modeling for a better understanding of microbial ecosystems

Jérôme Harmand

► **To cite this version:**

Jérôme Harmand. Contribution of modeling for a better understanding of microbial ecosystems. Ecological Engineering and Environment Protection, 2019, pp.14-22. 10.32006/eep.2019.2.1422 . hal-04172850

**HAL Id: hal-04172850**

**<https://hal.inrae.fr/hal-04172850>**

Submitted on 24 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## II. BIOAUTOMATICS AND BIOINFORMATICS

### CONTRIBUTION OF MODELING FOR A BETTER UNDERSTANDING OF MICROBIAL ECOSYSTEMS

Jérôme Harmand

**Abstract.** This paper aims at presenting how a number of recent modeling approaches can be used for better understanding microbial ecosystems dynamics. In first part, an important question – the ability of certain ecosystems to exhibit overyielding – is investigated using a model-based approach. It is shown that classical competition theory cannot explain such phenomenon, thus invalidating a large class of classical mass-balance-based models, Rapaport et al. (2019). In second part, we show how new combinatorial approaches can be used to find the best combination of species of a functional ecosystem with limited complexity. More precisely classification approaches inspired from the work by Jaillard et al. (2018) are used and illustrated with simulations. Their robustness with respect to a number of experimental parameters (investigated in simulation) is studied. For ecosystems with higher richness, we show how another probabilistic approach proposed by Jaillard et al. (2014) may be useful.

**Keywords:** Modeling, microbial ecosystems, microbial interactions, mathematical ecology, diversity, community assemblage.

#### 1. INTRODUCTION

The capacity of certain ecosystems to exhibit better performances than when the species which are assembled are cultivated alone is called overyielding. Such a phenomenon – very interesting from an applied viewpoint - has been studied for a long time in plant ecology, cf. for instance Tilman et al. (2006) or Schmid et al. (2008). Because they can be easily manipulated, microbial ecosystems are now more and more used for studying ecological-related questions, Jessup et al. (2004). In addition to being easier to manage than macro-systems, microbial ecosystems have a practical interest, in particular since bioreactors with natural ecosystems are widely used in industry. It has been shown that we observe such phenomenon with the anaerobic digestion. In systems fed with complex substrates, complementary and interaction effects are usually invoked to explain overyielding, cf. Hamelin and Milferstedt, 2015. This paper aims at illustrating how the use of mathematical modelling may help to investigate important questions for the optimization of biological processes such as the transgressive overyielding. More specifically, we consider batch experiments with inocula containing 1 to  $N$  species of a regional pool of species each of them being in competition for a limiting substrate. First, we recall that without mortality terms, the best biogas productivity is obtained with the reactor that has been inoculated only with the “best” species (there is no possibility of overyielding). Second, if other processus [here mortality, but another may be used cf. Rapaport et al. (2019) or Harmand et al. (2019)] are added it is established that overyielding is possible. We establish conditions under which

biodiversity promotes performance overyielding. It is important to notice that these results are not dependent on the growth rate functions used.

#### 2. BIODIVERSITY: AN ENGINE FOR OVERYIELDING

A key question in microbial ecology is to establish whether the performance of a process is dependent on the inoculum used during start-up or not (Forster-Carneiro et al., 2008). In particular, if it is assumed that we can build the inoculum from a combination of species assembled from a regional pool, the question we address is to study the performance of the associated process with respect to the biomass activity (for instance wrt the production of biogas in anaerobic systems: it is the performance index that will be considered latter in the document), Jaillard et al. (2014).

Transgressive overyielding refers here to the following property. Consider two different species A and B (here a “species”) is characterized by its kinetics parameters, its consumption yield (mass of substrate consumed by mass of biomass produced) and its conversion yield in biogas (volumetric mass of biogas by mass of biomass produced). Assume we proceed to three batch experiments. The first two refer to batch experiments in which the inoculum consists only of A or B with concentrations  $X(0)=X_0$  and  $S(0)=S_0$  in both experiments (where  $X$  and  $S$  are the biomass and the substrate concentrations). The third experiment consists in mixing both species (it is supposed here that there is as much A as B at the origin, that is  $X(0)/2$  for each). Let  $Q^A(T)$ ,  $Q^B(T)$  and  $Q^{AB}(T)$  be the total volume of biogas produced



at instant  $T$  in each experiment,  $T$  being supposed to be large enough such that  $S(T)$  is negligible with respect to  $S_0$  in all experiments. Experimental results are illustrated in Fig. 1.

Excluding the cases where the performance of the mixture AB would be lower than the lowest performance obtained with the single specie A, three cases of interest here may arise: either none overyielding is observed (i.e. case #1 which the produced gas is denoted by  $Q^{AB}(T)^1$ , the performance of AB equals the mean performance of A and B alone), or an overyielding is observed. In this last case, the overyielding may be transgressive (case #3 where the performance of AB, denoted as  $Q^{AB}(T)^3$ , is better than the best performance observed for single species) or not (case #2 where the performance  $Q^{AB}(T)^2$  is better than the expected mean value of A and B but is not better than the best single batch).

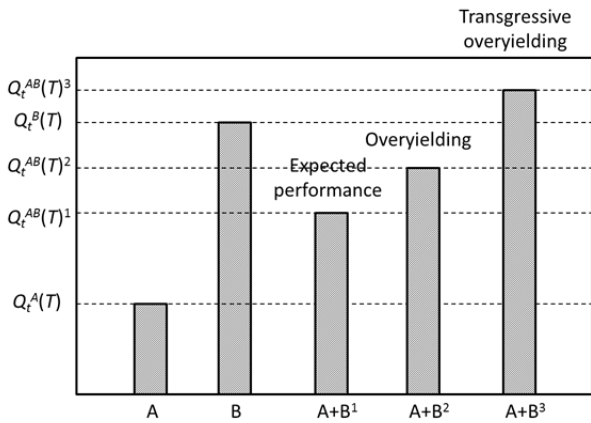


Fig. 1. Performances of pure and mixed cultures of microorganisms in batch experiments.

In the next section, we investigate, via modelling, the links between biodiversity and overyielding in batch bioreactors. We shall mainly focus on transgressive overyielding even though we shall not always precise the adjective “transgressive”.

### 3. PROCESS MODELING

The most common model used in biotechnology to describe the competition of  $n$  species for a single limiting substrate is written as:

$$\begin{cases} \dot{X}_i = \mu_i(\cdot)X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(\cdot)}{Y_i} X_i \end{cases} \quad (1)$$

where  $X_i$  and  $S$  stand for biomasses and substrate concentrations, respectively,  $Y_i$  is the biomass yield, and  $\mu_i$  is the specific growth rate of the  $i^{\text{th}}$  biomass.

The output gas flow rate is supposed to be a linear function of the biomass activity and can thus be written as:

$$SQ_{gaz}(T) = \int_0^T \sum_i k_i \mu_i(\cdot) X_i(\tau) d\tau \quad (2)$$

Does this model exhibit overyielding? Equation (1) allows us to establish the following invariant:

$$\sum_i \frac{\dot{X}_i}{Y_i} + \dot{S} = 0 \implies \sum_i \frac{X_i(T) - X_i(0)}{Y_i} \quad (3)$$

If it is further assumed that  $T$  is large enough such that  $S(T)$  can be neglected with respect to  $S(0)$  (from now, we will refer to this hypothesis as H1) then one has:

$$\sum_i \frac{(X_i(T) - X_i(0))}{Y_i} \simeq S(0) \quad (4)$$

Posing:

$$p_i = \frac{X_i(T) - X_i(0)}{Y_i S(0)} \quad (5)$$

one has:

$$\sum_i p_i = 1 \quad (6)$$

Using these notations, one can rewrite the biogas production as:

$$\begin{aligned} SQ_{gaz}(T) &= \int_0^T \sum_i k_i \mu_i(S(\tau)) X_i(\tau) d\tau \\ &= \int_0^T \sum_i k_i \dot{X}_i(\tau) d\tau \\ &= \sum_i k_i (X_i(T) - X_i(0)) \\ &= S(0) \sum_i k_i Y_i p_i \end{aligned} \quad (7)$$

Under H1, one can then establish that:

$$SQ_{gaz}(T) = S(0) \sum_i k_i Y_i p_i \leq S(0) \sum_i \max_i (k_i Y_i) p_i$$

or

$$SQ_{gaz}(T) \leq S(0) k_{i^*} Y_{i^*} \sum_i p_i = S(0) k_{i^*} Y_{i^*}$$

where  $i^*$  is such that  $k_{i^*} Y_{i^*} = \max_i k_i Y_i$ .

From this calculations, we can claim the following (Rapaport et al., 2019 or Harmand et al., 2019):

- Under H1, the maximum of biogas produced at instant  $T$  is obtained with a pure culture of the species  $i^*$  having its parameters which verify  $k_{i^*} Y_{i^*} = \max_i k_i Y_i$ .

- Regardless of the dependence of  $\mu_i$  with respect to any state of the system (or even of any external input), as long as some growth rates are not zero while others remain positive, the previous result remains valid.

#### 4. CONDITIONS FOR OVERYIELDING

Now, instead of considering the model (1), one introduces mortality terms in dynamical equations:

$$\begin{cases} \dot{X}_i = (\mu_i(\cdot) - m_i) X_i \\ \dot{S} = -\sum_{i=1}^n \frac{\mu_i(\cdot)}{Y_i} X_i \end{cases} \quad (8)$$

where  $m_i$  are mortality terms.

In this case, it may be shown that under some conditions on the model parameters, the ecosystem may exhibit overyielding (Rapaport et al., 2019). The explanation is as follows. Consider a system with dynamics given by (8) with only one biomass. If  $m_1$  is large enough with respect to growth parameters, it is expected that the biomass concentration may converge towards zero before – and possibly a long time before – the substrate is close to zero. In other words, it may happen that some substrate remains in the system. The

question is to identify if the remaining substrate at a given time  $T$  is negligible or not (that is Assumption H1). Consider now a system with two biomasses. Assume one biomass is potentially the best (it has the largest  $kY$  parameter) but it has a high mortality. The other performs less but has a lower mortality. The first species (with high mortality) will produce less biogas than if it was alone but will leave some substrate available to the second species after the first one has died because of its high mortality (Fig. 2). The second will then use this remaining substrate to produce more biogas since it will live longer (Fig. 2 and Fig. 3). There exists then situations where the sum of the biogas produced by both species will be higher than that produced in pure cultures of each species (Fig. 3). In the present theoretical example, the overyielding is of about 20 %.

As mentioned in the introduction, other phenomena may explain overyielding. In fact any process where a species is penalized – at least ponstually – while another is not, may lead to overyielding. For instance, if species exhibit growthrates with thresholds (Harmand et al., 2019).

These results largely depend on the model used and thus on its parameters. In practice, it is rather rare to have a model whatever the ecosystem under interest. Thus, we need other approaches to study complex ecosystems and to be able to predict the performances of a given mixture of microorganisms. In the next sections, we show how new combinatorial approaches may be used to deal with such problems for different ecosystems depending on their complexity.

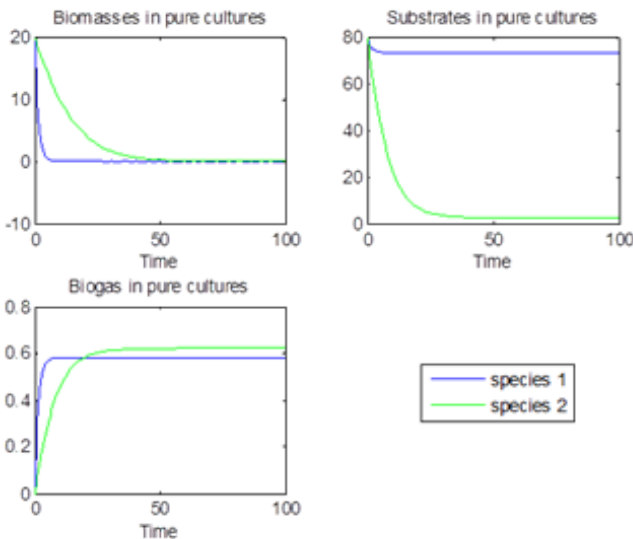


Fig. 2. Pure cultures of two species with mortality in batch experiments.

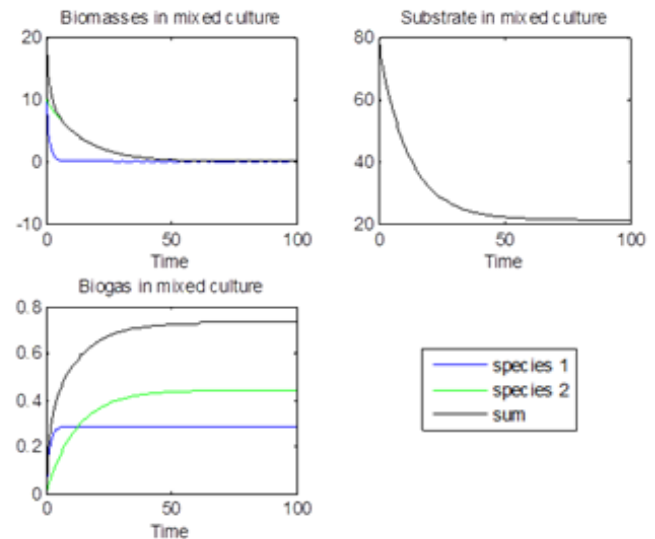


Fig. 3. Pure cultures of two species with mortality in batch experiments.

**5. CLASSIFYING SPECIES**

In this section, we show how we can classify species in order to be able to predict the performances of any assemblage of a limited number of species chosen in a strain library. We use an approach inspired from that proposed by Jaillard et al. (2018). This approach is particularly attracting since it only requires a list of experiments each comprising i) the composition of the ecosystems (which species of the library were used – assuming constant total initial substrate and biomass) and ii) a performance index as summarized in the following Table 1.

Here our objectives are:

- to classify species in specific families with respect to their good or bad influence the performance index when they are in interaction with other species;
- to predict the performance of any assemblage made of the co-culture of some species chosen in the strain library.

**Notations and definitions**

Let us denote  $X$  the number of available species and  $X_i$  the  $i^{th}$  species (in Table 1, Species 1 is then denoted by  $X_1$ , Species 2 by  $X_2$ , etc.)

- $C$  is the number of classes in which each strain must be assigned ( $C$  is the main degree of freedom for the user) and  $C_i$  the  $i^{th}$  class;
- $E$  is the number of available experiments and  $E_i$  the  $i^{th}$  experiment;

- Given an assignation, that is such species  $X_i, i=1..X$ , belongs to such class  $C_j, j=1..C$ , one can define motifs  $M_i$  which allows us to classify the experiments. Given  $C$ , the number of possible motifs is given by  $2^C-1$ .

Example with  $C=2$ :

- Either the experiment  $i$  contains only species of the same class 1: it will then classified in  $M_1$ ;
- Either the experiment  $i$  contains only species of the same class 2: it will then be classified in  $M_2$ ;
- Either the experiment  $i$  contains species of both classes: it will then be classified in  $M_3$ ;
- To do so, Jaillard et al. (2018) propose a general algorithm which may be difficult to encode since it uses many tricks, notably at the initial steps. If it has been shown on some experiments to give very interesting classification results, it may not be guaranteed to provide the global optimum of the classification since it uses a supervised algorithm. It is for sure useful to process data when the number of species or of classes is important but in the case, as in this section, where we consider ecosystems of limited complexity, we rather suggest to use a *brute force* algorithm which corresponds to testing all possible combinations as explained in the following algorithm.

Table 1: Experiments with  $X=6$  species

Performance	Presence / absence of species i in the assemblage					
	Species 1	Species 2	Species 3	Species 4	Species 5	Species 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

**Classification with a brute force algorithm**

The proposed algorithm is as follows:

**Initialization:** Get data and fix  $C$

**Step #1:** Assign each species  $X_i$  to a class  $C_i$  to which they have not yet been assigned in previous iterations (for instance  $X_1 \in C_1$  while  $X_2, X_3, X_4, X_5, X_6 \in C_2$ ).

**Step #2:** Given this species assignation, assign the experiments to the different motifs as shown in Table 2 for the example of assignation given in Step #1 with  $C=2$ ,  $X_1 \in C_1$  in written in red while  $X_2, X_3, X_4, X_5, X_6 \in C_2$  are written in green. Each experiment is then assigned to one of the three motifs.

**Step #3:** Using the performance indices of all experiments assigned in a given motif  $M_i$  compute the mean value and the dispersion of the performance index.

**Step #4:** Iterate on the species assignation until all possible combinations have been tested (notice that it is simply equivalent to count in the  $C$ -basis over the number of species).

*The best species assignation in the  $C$  classes is the one for which the sum of the dispersions within motifs is the smallest over all possible species combinations and the mean value of the performance within each motif will give us the best prediction of the assemblage corresponding to that defined by the motif.*

Table 2. Iteration  $i$  - assignation of each experiment to a given motif

Performance	Presence / absence of species $i$ in the assemblage						Motif
	Species 1	Species 2	Species 3	Species 4	Species 5	Species 6	
0.3970	1.0000	0	0	0	0	0	M1
1.3950	0	1.0000	0	0	0	0	M2
0.4200	0	0	1.0000	0	0	0	M2
0.7870	0	0	0	1.0000	0	0	M2
0.4180	0	0	0	0	1.0000	0	M2
0.3820	0	0	0	0	0	1.0000	M2
1.3100	1.0000	1.0000	0	0	0	0	M3
0.5530	1.0000	0	1.0000	0	0	0	M3
0.9250	1.0000	0	0	1.0000	0	0	M3
0.6720	1.0000	0	0	0	1.0000	0	M3
0.6940	1.0000	0	0	0	0	1.0000	M3
1.3400	0	1.0000	1.0000	0	0	0	M2
1.4080	0	1.0000	0	1.0000	0	0	M2
1.2330	0	1.0000	0	0	1.0000	0	M2
1.2540	0	1.0000	0	0	0	1.0000	M2

**Example: classifying species with respect to their growth rate**

An example of the effectiveness of this algorithm is given hereafter. To test the approach, we generated a dataset using combinations of 8 species. More precisely, we randomly selected 8 of the 22 species identified from Landa et al. (2018), characterized by their maximum specific growth rate. Then, the 255 possible combinations of species were simulated. The model considered to simulate them was the classical competition model in the chemostat with the growth rates given by  $\mu_i(S) = \mu_{i,max} S$  where  $\mu_{i,max}$ ,  $i=1,8$  are the specific growth rates of the species. The hydraulic retention was chosen to be large enough to guarantee that no washout occurs for the smallest growing species.

The system was simulated over a period  $T$  corresponding to 5 retention times (a period considered in practice to be sufficient for being close to the equilibrium). It is important to notice that we generated two distinct datasets: one was completely free of noise (dataset#1) while the second was noisy (dataset#2). For the noisy case, instead of choosing equally distributed initial conditions for the species present in a given assemblage, the initial ratios of each species were chosen randomly while the total mass was constant. For both datasets, the initial substrate concentration was constant. The eight maximum specific growth rates, ordered in the increasing order were 0.0001, 0.0205, 0.0237, 0.0294, 0.0343, 0.0372, 0.0406 and 0.0423. For a given experiment with  $n$  strains, the performance

index was the integral over a given period of time of the activity of the ecosystem that is:

$$\int_{t_1}^{t_2} \sum_{i=1}^{i=n} \mu_i(S) X_i$$

On the basis of these data, the idea of the test is to see if the algorithm can be used to classify the species according to their maximum specific growth rates.

**Results**

First, all data at the final time ( $t_1=0$  and  $t_2=T$ ), that is the dataset#1 comprising the 255 available

experiments, were used. The brute force algorithm was used first with  $C=2$ , and then with  $C=3$ . In both cases, the 8 strains were perfectly classified with respect to their maximum specific growth rates. Then, we investigated the robustness of the result with respect to time. To do so, we used data with the performance indexes taken at different instants: earlier the instant, lesser the information. The results remain quite good: unless for very small instants (two or less retention times) the classification remains good (Table 3).

Table 3. Assignment results for dataset#1 when using all data at different instants -  $t=1, 2, 3, 4$  or  $5$  retention times

X	C	M	Nb experiments / % of total	t	Best classification	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
X	C	M	Nb experiments / % of total	t	Best classification	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

In addition, the robustness of the approach was investigated with respect to the number of available experiments. Indeed, in practice, it is the rule rather than the exception that only few experiments can be performed. In the case of 8 strains, it is probable that only a limited number of all combinations will be made. Assuming only 50, 25 or 10% of experiments (randomly selected) are available, the optimization is re-run, the results of which are shown in Table 4. It is noticed that results really degrade only when both a limited number of experiments realized over a small period of time are used. Less than 50% of species are badly assigned as long as more than 25% of the total possible experimental combinations. The algorithm can then be said to be quite robust with

respect to the fractions of total experimental combinations used.

Finally, we tested the approche using the noisy dataset (dataset#2). Results are reported in Table 5. From these results, it is postulated that the method exhibit high robustness with respect to initial conditions of the species. In addition, there are less species badly assigned when dealing with uncertainty on the initial conditions than with respect to a limited number of data.

*Thus, it seem that it is better for the the user to perform as many different experiments as possible instead of repeating many times the same experiment.*

Table 4. Assignment results for dataset#1 when using only 50, 25 and 10% of data at instants -  $t=1$  and 5 retention times

X	C	M	Nb experiments / % of total	Tf	Best classification	R2	Recall with all experiments	Nb species badly assigned
8	2	3	125 / 50%	1	00000001	0.23	00000101	1
8	2	3	125 / 50%	5	00000111	0.51	00000111	0
8	3	7	125 / 25%	1	00111212	0.44	01111212	1
8	3	7	125 / 25%	5	00000112	0.64	00001122	2
8	2	3	25 / 10%	1	00001111	0.40	00000101	2
8	2	3	25 / 10%	5	00110000	0.53	00000111	5
8	3	7	25 / 10%	1	01100200	0.69	01111212	4
8	3	7	25 / 10%	5	01000222	0.84	00001122	3

Table 5. Assignment results for dataset#2 when using only 100, 50, 25 and 10% of data at instants -  $t=1$  and 5 retention times

X	C	M	Nb experiments / % of total	Tf	Best classificat'n	R2	Recall with all experiments	Nb species badly assigned	Recall with noise-free data
8	2	3	255 / 100%	1	00001111	0.44	-	-	00000101
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	00111212
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	00110000
8	3	7	25 / 10%	1	00001102	0.76	00111112	3	01100200
8	3	7	25 / 10%	5	00010112	0.95	00001112	2	01000222

## 6. A PROBABILISTIC APPROACH

In the previous section, the approach used allowed us to classify species in such a way we can predict any assemblage taken in a strain library. Such an approach seems quite well adapted for ecosystems of limited complexity since it is necessary to perform very specific experiments mixing together the different species. It may happen

that we have to deal with library including hundreds of species. In such a way, performing enough experiments to classify each species become untractable. In such a case, other approaches may be used to characterize the ecosystem. Obviously, information we can easily get from data will be much more qualitative.



Assume we can perform experiments allowing us to statistically characterize the performance of the ecosystem with respect to its diversity. Fig. 4 reports what could be such results.

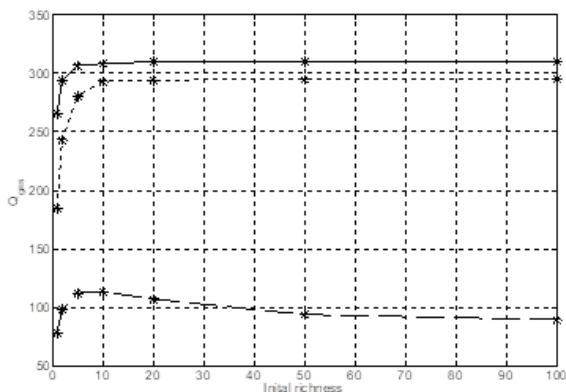


Fig. 4. Volume of biogas produced over a given period of time by an anaerobic digester as a function of the richness of the inoculum. Full and dotted lines were obtained with ecosystems only containing species interacting positively and neutrally, respectively, while the dashed line represents performances of an ecosystem containing some negative species.

We can then consider how species interact. Classifying the available species in a number of classes (for instance in positive, neutral and negative groups) and assuming a number of rules when these species are together in the ecosystem (for instance, the majority rule states that if there is more positive species than negative ones, then the effect of the assemblage on the performance is positive), it is possible for each potential assemblage, to calculate the probability of observing a performance which will be greater than the expected one. The strength of the approach relies in the fact that, depending on the different classes and the way they interact, one can calculate analytically the above probabilities. Examples of the use of such rules are given in Jaillard et al. (2014). If it is not possible to go very far in the characterization of each species, one can typically deduce from the approach the ratio of positive versus negative species of the library.

## 7. CONCLUSIONS

In this paper, a number of modeling approaches were used to investigate the way species interact within an ecosystem. We have shown how a model-based approach could be used in order to investigate

overyielding. It was shown that the classical competition-model in the chemostat cannot exhibit overyielding. Instead it is necessary to include other processes such as mortality for the ecosystem to be able to exhibit overyielding. To classify species with respect to their growth rate from data obtained either in continuous or batch modes, we have used an approach based on a combinatorial approach. For low complexity ecosystem, such an approach has been shown to be quite robust with respect to both experimental uncertainty and to the number of data available. Finally, a probabilistic approach can be used in order to investigate the number of potential positive versus negative species within a strain library.

**ACKNOWLEDGEMENTS:** *The author gratefully acknowledges the Bulgarian Academy of Sciences.*

## REFERENCES

1. Jaillard, B., Deleporte, P., Loreau, M., Violle, C. (2018) A combinatorial analysis using observational data identifies species that govern ecosystem functioning, PLOS ONE, 13:8, Article Number: e0201135, DOI: 10.1371/journal.pone.0201135.
2. Jaillard B., A. Rapaport, J. Harmand, A. Brauman and N. Nunan (2014) Community assembly effects shape the biodiversity-ecosystem functioning, Functional Ecology, Functional Ecology, 28, 1523–1533.
3. Tilman, D., P. B. Reich and J. M. H. Knops (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment, Nature volume 441, 629–632.
4. Schmid, B. A. Hector, P. Saha and Michel Loreau (2008) Biodiversity effects and transgressive overyielding, Journal of Plant Ecology, Volume 1, Issue 2, 95–102.
5. Jessup, C. M., R. Kassen, S.E. Forde, B. Kerr, A. Buckling, P. B. Rainey and J. M. Bohannan (2004) Big questions, small world: microbial model systems in ecology, TRENDS in Ecology and Evolution, 19:4, 189–197.
6. Hamelin, J. and K. Milferstedt (2015) Mixing ecosystems during start-up of anaerobic digesters improves final performance, International Conference on Anaerobic Digestion, AD Technology and Microbial Ecology for Sustainable Development (ADTech2015), Chiang Mai, Thailand.
7. Rapaport, A., T. Nidelet, S. El Aida, J. Harmand (2019) About biomass overyielding of mixed cultures in batch processes, submitted to JTB.

8. Harmand, J., A. Rapaport, T. Nidelet (2019) About overyielding with mixed cultures in batch processes, FOSBE, 15-18 October, Valencia, Spain.

9. Forster-Carneiro, T., M. Perez, L. I. Romero (2008) Influence of total solid and inoculum contents on performance of anaerobic reactors treating food waste, *Bioresource Technology*, Volume: 99, Issue: 15, 6994-7002.

10. Landa, M., S. Blain, J. Harmand, S. Monchy, A. Rapaport, I. Obernosterer (2018) Major changes in the composition of a Southern Ocean bacterial community in response to diatom-derived dissolved organic matter, *FEMS Microbiology Ecology*, Vol. 94, Issue 4, DOI: 10.1093/femsec/fiy034

**Jérôme Harmand**

Laboratoire de Biotechnologie de l'Environnement,  
LBE, 11100 Narbonne, France

E-mail: [jerome.harmand@inra.fr](mailto:jerome.harmand@inra.fr)