

# RAPPORT DE STAGE

Effectué par

**Ulysse HERBACH**

au

Laboratoire de Biométrie et Biologie Evolutive (UMR CNRS 5558 )

Université Claude Bernard Lyon 1

Sujet du stage :

**Interaction algues-daphnies au sein d'un microcosme :  
modélisation dynamique, analyse de sensibilité  
et optimisation non linéaire**

juin-juillet 2012

Magistère de Mathématiques de Rennes, 1<sup>ère</sup> année

(Université Rennes 1 et ENS Cachan, antenne de Bretagne)

Maîtres de stage : Sandrine CHARLES et Christelle LOPES

Tuteur universitaire : Mihai GRADINARU



Université Claude Bernard



Lyon 1



## Résumé

Dans ce rapport, nous présentons un modèle mathématique décrivant l'évolution d'un microcosme algues-daphnies au cours du temps, dans un cadre écotoxicologique. Nous détaillons dans un premier temps les étapes de la construction du modèle, constitué d'équations différentielles ordinaires, puis nous effectuons une analyse de sensibilité globale basée sur les indices de Sobol, dont nous nous servons pour fixer certains paramètres du modèle. Enfin, nous abordons différentes techniques d'estimation des paramètres. Après avoir montré l'insuffisance d'une approche déterministe par des méthodes de type "descente de gradient", nous effectuons une inférence bayésienne grâce à deux méthodes de Monte Carlo : l'algorithme de Metropolis-Hastings, ainsi qu'un algorithme plus efficace basé sur le filtrage particulaire.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>8</b>
1.1	Les enjeux de la modélisation des écosystèmes . . . . .	8
1.2	Le cadre écotoxicologique . . . . .	8
1.3	Précisions sur le vocabulaire employé . . . . .	10
<b>2</b>	<b>Construction du modèle mathématique</b>	<b>11</b>
2.1	Caractéristiques du système algues-daphnies . . . . .	11
2.2	Données expérimentales . . . . .	12
2.3	Simplifications en vue de la modélisation . . . . .	13
2.4	Ecriture des équations différentielles . . . . .	14
2.4.1	Système sans daphnies . . . . .	14
2.4.2	Système avec daphnies . . . . .	17
2.5	Expression de la vitesse de sédimentation . . . . .	20
2.6	Modèle mathématique du système algues-daphnies . . . . .	22
2.6.1	Exemple de courbes obtenues . . . . .	22
2.6.2	Premières propriétés . . . . .	23
2.6.3	Précisions sur les constantes et les paramètres . . . . .	24
<b>3</b>	<b>Analyse de sensibilité</b>	<b>26</b>
3.1	Motivations et principe de l'analyse de sensibilité . . . . .	26
3.2	Choix de la méthode . . . . .	27
3.2.1	Application au modèle du système algues-daphnies . . . . .	27
3.2.2	Nécessité d'une décomposition de la variance . . . . .	28
3.3	Les indices de Sobol . . . . .	28
3.3.1	Définition des indices de Sobol dans un cadre général . . . . .	28
3.3.2	Les indices de Sobol dans un cadre probabiliste . . . . .	30

3.3.3	Indices de sensibilité totaux . . . . .	32
3.3.4	Application au modèle algues-daphnies . . . . .	32
3.4	Estimation des indices de Sobol . . . . .	34
3.4.1	Estimation par Monte Carlo . . . . .	34
3.4.2	Résultats . . . . .	34
<b>4</b>	<b>Estimation des paramètres</b>	<b>37</b>
4.1	Estimation ou optimisation? . . . . .	37
4.2	Notations en vue de l'utilisation des données . . . . .	37
4.3	Approche déterministe : méthode des moindres carrés . . . . .	38
4.3.1	Algorithmes de minimisation . . . . .	39
4.3.2	Problème fondamental : minima locaux . . . . .	40
4.4	Approche stochastique : méthodes de Monte Carlo . . . . .	41
4.4.1	Lien entre l'approche fréquentiste et la méthode des moindres carrés . . . . .	42
4.4.2	L'approche bayésienne . . . . .	42
4.4.3	Principe des méthodes de simulation de Monte Carlo . . . . .	44
4.5	L'algorithme de Metropolis-Hastings . . . . .	44
4.6	Un algorithme de filtrage particulaire . . . . .	46
4.6.1	Convergence de l'algorithme SIR . . . . .	46
4.6.2	En pratique : étapes de l'algorithme . . . . .	48
4.6.3	Résultats . . . . .	48
<b>5</b>	<b>Conclusion</b>	<b>50</b>
	<b>Références</b>	<b>52</b>
<b>A</b>	<b>Compléments du chapitre 2</b>	<b>53</b>
A.1	Valeurs des paramètres associés à la figure 12 . . . . .	53
A.2	Régularité de la fonction $F$ . . . . .	53
A.3	Preuve de la proposition 1 . . . . .	54
A.4	Preuve de la proposition 2 . . . . .	55
A.5	Distribution <i>a priori</i> pour les méthodes de Monte Carlo . . . . .	58
<b>B</b>	<b>Conditionnement dans le cas des variables à densité</b>	<b>59</b>

# Liste des figures

1	Schéma du microcosme complet ( <i>Image : B. Clément</i> ) . . . . .	9
2	Les deux espèces composant le microcosme étudié . . . . .	10
3	Représentation graphique des données sur le système algues-daphnies . . . . .	12
4	Schéma du système algues-daphnies . . . . .	13
5	Allure d'une solution de l'équation logistique . . . . .	15
6	Schéma du processus de sédimentation : passage de $t$ à $t + dt$ . . . . .	16
7	Graphe de la fonction $TI$ lorsque $ILL$ et $TI_{max}$ sont fixés . . . . .	18
8	Allure d'une solution de l'équation de von Bertalanffy ( $a = 0,2$ ) . . . . .	18
9	Graphe de la fonction $h(\cdot, a_3N_2)$ lorsque $a_3N_2$ est fixé . . . . .	19
10	Graphe de la fonction $\tilde{h}(\cdot, a_3N_2)$ lorsque $a_3N_2$ est fixé . . . . .	20
11	Evolution de la concentration algale en l'absence de lumière . . . . .	21
12	Exemple de solution pour le système algues-daphnies . . . . .	23
13	Représentation des sommes des indices de Sobol totaux . . . . .	36
14	Courbes ajustées par méthode de descente . . . . .	40
15	Projection de $h_1$ sur l'axe $r_1$ , les autres paramètres étant fixés . . . . .	41
16	Courbes correspondant à un tirage aléatoire selon une approximation de $\pi$ . . . . .	45
17	Exemple de projection du nuage de particules . . . . .	49
18	Histogrammes obtenus par l'algorithme de filtrage particulaire . . . . .	49

# Liste des tableaux

1	Quantités intervenant dans le modèle du système algues-daphnies . . . . .	14
2	Constantes du modèle du système algues-daphnies . . . . .	25
3	Paramètres du modèle du système algues-daphnies . . . . .	25
4	Intervalles de variation de chaque paramètre pour l'analyse de sensibilité . . . . .	33
5	Estimation des indices de Sobol totaux pour chaque paramètre . . . . .	35
6	Valeurs des paramètres considérés comme peu influents . . . . .	36
7	Bornes de la distribution <i>a priori</i> pour les méthodes de Monte Carlo . . . . .	58

# Chapitre 1

## Introduction

### 1.1 Les enjeux de la modélisation des écosystèmes

La préservation des écosystèmes est devenue un enjeu majeur, notamment à cause de leur lien étroit avec la disponibilité de certaines ressources indispensables à la survie de l'humanité, ne serait-ce que l'eau potable. Leur fragilité intrinsèque et l'irréversibilité de certaines perturbations incitent l'Homme à vouloir mesurer l'impact qu'il exerce dessus. Mais pour cela, il est nécessaire de mieux comprendre leur fonctionnement.

Bien sûr, comme c'est le cas avec tout ce que l'on trouve dans la nature, on est incapable d'avoir une compréhension *totale* d'un écosystème, du fait de la complexité des processus qui le caractérisent. En revanche, on peut tenter de le décrire par un modèle qui soit *relativement proche* de la réalité, c'est-à-dire en se plaçant à une échelle de précision adaptée à la problématique posée.

Or, les mathématiques ont depuis longtemps montré leur force lorsqu'il s'agit de modéliser des phénomènes biologiques (Fibonacci, XIII<sup>ème</sup> siècle). Au fil du temps, la biologie a connu pour cette science un intérêt croissant, décuplé avec l'apparition de l'informatique. A présent, la modélisation mathématique apparaît comme un outil essentiel pour décrire et comprendre le fonctionnement des écosystèmes.

### 1.2 Le cadre écotoxicologique

Ce stage s'inscrit dans une problématique d'*écotoxicologie* : on s'intéresse aux effets d'une pression chimique sur un écosystème. Plus précisément, on cherche ici à savoir comment se comporte un écosystème aquatique continental lorsqu'il est soumis à une substance toxique.



On peut baser ce type d'étude sur l'expérimentation, en considérant une version simplifiée de l'écosystème, appelée *mésocosme* ou *microcosme* selon le nombre d'espèces : on réalise alors des expériences sur un mésocosme/microcosme schématique qu'on appelle *système* (cf. partie 1.3), en mesurant des quantités qui caractérisent l'état de ce système à différents instants.

Dans le cas d'un écosystème aquatique continental, un microcosme adapté à l'expérimentation (fig. 1) a été mis au point par Clément *et al.* [1]. Il est composé de cinq espèces :

- *Pseudokirchneriella subcapitata* : algue unicellulaire microscopique (taille  $< 6\mu\text{m}$ ), souvent utilisée en écotoxicologie car facilement cultivable en laboratoire ;
- *Lemna minor* : lentille d'eau se développant à la surface de la colonne d'eau ;
- *Daphnia magna* : crustacé zooplanctonique (taille  $\approx 5\text{mm}$ ) qui consomme les algues par broutage et se déplace grâce à ses grandes antennes plumeuses ;
- *Chironomus riparius* : insecte diptère dont la larve aquatique, de couleur rouge, est communément appelée "ver du pêcheur" ;
- *Hyalella azteca* : crustacé épibenthique, *i.e.* qui évolue à la fois dans le sédiment et la colonne d'eau.

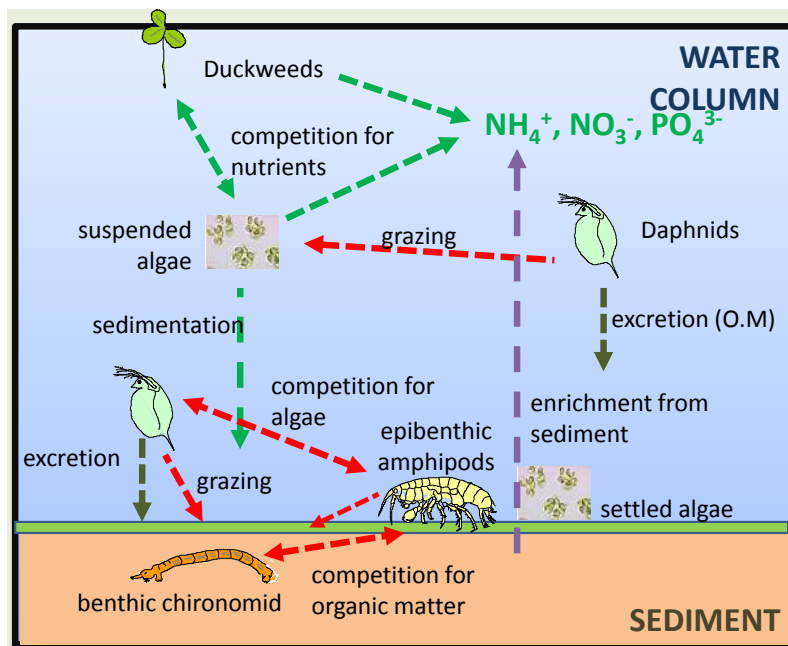


FIGURE 1 – Schéma du microcosme complet (Image : B. Clément)

Bien que simpliste comparé à un écosystème réel, ce microcosme constitue déjà un système complexe avec plusieurs interactions : on a notamment les relations trophiques, *i.e.* relatives à l'alimentation des différentes espèces (flèches rouges et vertes sur la figure 1).

On pourrait tenter de modéliser d'emblée ce microcosme en entier, mais dans un premier temps il est plus simple de décrire les interactions séparément. Ainsi, dans ce stage nous avons considéré un microcosme plus petit, constitué seulement de deux espèces parmi celles du microcosme complet : la daphnie *Daphnia magna* (fig. 2a) et l'algue *Pseudokirchneriella subcapitata* (fig. 2b). Nous avons pris en compte les relations trophiques suivantes : les daphnies se nourrissent d'algues, et les algues absorbent les nutriments contenus dans l'eau. Enfin, on se limite à l'étude de ce microcosme algues-daphnies sans contaminant.



(a) *Daphnia magna*



(b) *Pseudokirchneriella subcapitata*

FIGURE 2 – Les deux espèces composant le microcosme étudié

### 1.3 Précisions sur le vocabulaire employé

Certains mots sont utilisés dans tant de domaines que leur sens devient flou : c'est le cas du mot *système*. Il convient de préciser le sens de certains mots dans ce rapport :

- **microcosme algues-daphnies** : les deux espèces et leurs interactions, lorsqu'on les considère de manière générale (à voir comme une partie de la figure 1).
- **système algues-daphnies** : microcosme schématique dans les conditions précises d'expérimentation et sous certaines hypothèses (cf. section 2.1). Dans la suite, on parlera également du sous-système “sans daphnies”.
- **modèle** : équations différentielles décrivant l'évolution du système algues-daphnies au cours du temps. Dans un contexte mathématique, on parlera de *système d'équations différentielles*.

## Chapitre 2

# Construction du modèle mathématique

### 2.1 Caractéristiques du système algues-daphnies

L'étude du microcosme algues-daphnies est basée sur l'expérimentation, et pour cela on construit d'abord un microcosme schématique appelé *système algues-daphnies*. On utilise un bécher de 2L qui contient de l'eau (avec des nutriments propices au développement des algues) ainsi qu'un sédiment sableux. A l'instant  $t = 0$ , on introduit une quantité fixée d'algues ( $N_1^0 = 4 \times 10^7$ ) et de daphnies ( $D = 10$ ), et le milieu est remué de façon à homogénéiser la concentration en algues. Le système algues-daphnies évolue ensuite ( $t > 0$ ) selon différentes hypothèses :

- la population d'algues croît suivant le modèle logistique (cf. p. 15) ;
- les daphnies consomment les algues avec un taux qui dépend à la fois de leur âge et de la concentration locale en algues ;
- après l'homogénéisation initiale, l'eau n'est plus remuée et les algues tombent au fond avec une certaine vitesse (non constante) ;
- le nombre de daphnies est supposé constant, les jeunes étant retirés régulièrement pour une raison pratique (éviter la surpopulation des daphnies).

Le but est de construire un modèle mathématique décrivant ce système algues-daphnies, de manière à mieux comprendre son évolution, voire la prédire. Les mesures de l'état du système à différents instants, constituant les *données*, permettront de confronter le modèle à la réalité.

## 2.2 Données expérimentales

Les données proviennent d'expériences réalisées à l'ENTPE par Bernard Clément et ses stagiaires [2, 3, 4]. On utilise les résultats témoins (*i.e.* sans cadmium) de trois essais, chaque essai comprenant en fait deux expériences : une sur le système algues-daphnies, et une sur le système sans daphnies, contenant seulement les algues (pour plus de détails sur les caractéristiques des trois essais, on renvoie à [4]). On dispose pour chaque expérience des mesures de deux quantités : le nombre d'algues *en suspension* dans la colonne d'eau, noté  $N_1$ , et le nombre d'algues qui se sont déposées sur le sédiment, dites *sédimentées*, noté  $N_2$ .

Ces quantités ont été mesurées à différents instants entre 0 et 21 jours, et sont représentées graphiquement sur la figure 3. Les figures à gauche représentent  $N_1$  sur 21 jours, sans daphnies (en haut) et avec daphnies (en bas). De même, les figures à droite représentent le nombre d'algues sédimentées  $N_2$ , sans et avec daphnies. Il convient de noter les différences d'échelle, dues au fait que les daphnies consomment une grande partie des algues. On remarquera également les variations et le nombre des instants de mesure, dont il faudra tenir compte lors de l'utilisation des données.

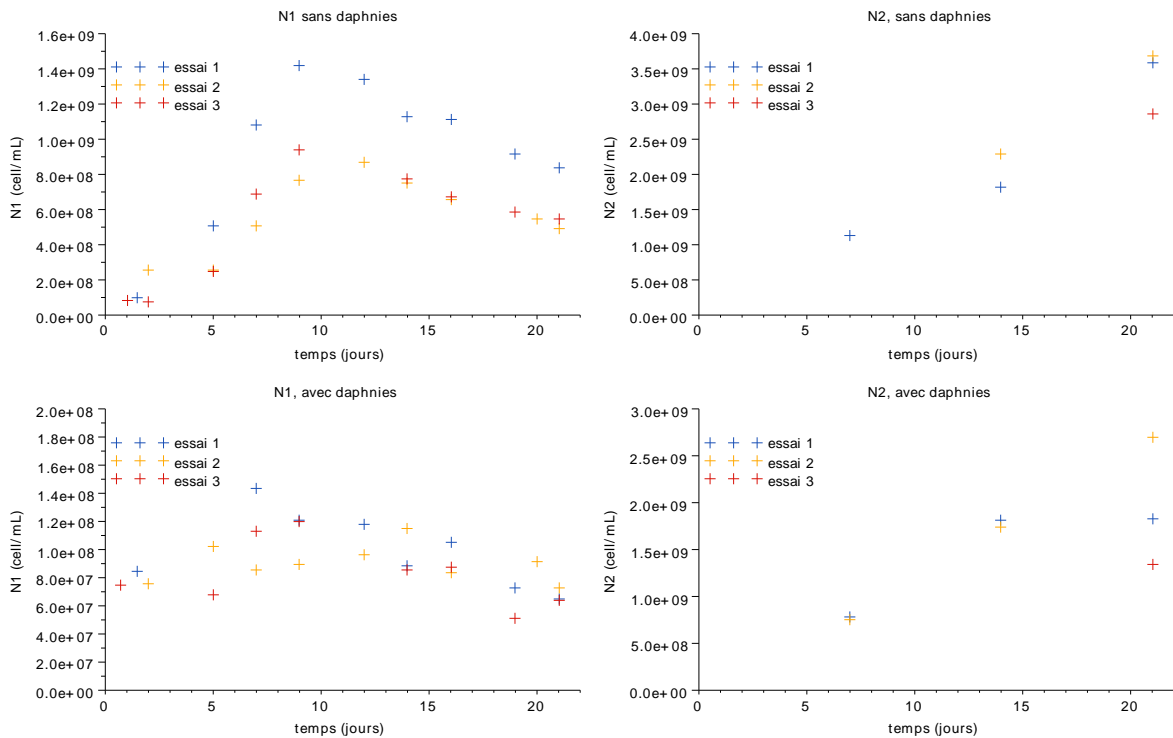


FIGURE 3 – Représentation graphique des données sur le système algues-daphnies

## 2.3 Simplifications en vue de la modélisation

Pour garder une cohérence avec les données, on distingue seulement deux catégories d’algues comme dans la section 2.2. On simplifie donc le système algues-daphnies en considérant uniquement deux compartiments, comme l’indique la figure 4. Les quantités respectives d’algues en suspension (compartiment 1) et sédimentées (compartiment 2) sont notées  $N_1(t)$  et  $N_2(t)$ .

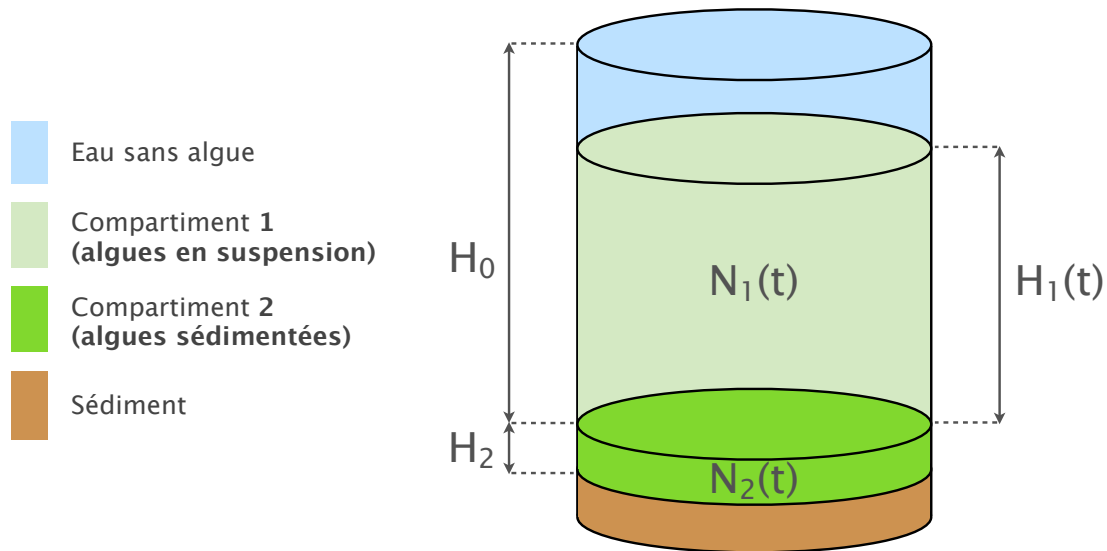


FIGURE 4 – Schéma du système algues-daphnies

On représente le b cher par un cylindre de section  $S$ . Le premier compartiment est de hauteur variable  $H_1(t)$ , et de volume  $V_1(t) = S \times H_1(t)$ . Le deuxi me est de hauteur constante  $H_2$ , et de volume constant  $V_2 = S \times H_2$ . En fait,  $H_2$  varie aussi (elle augmente avec le temps), mais elle reste n gligeable par rapport la hauteur totale  $H_0$  de la colonne d’eau. Quant   la taille du premier compartiment, elle diminue car l’eau n’est pas remu e : occupant initialement un volume  $V_1^0 = S \times H_0$ , les algues ne peuvent que descendre puisque leur densit  est plus grande que celle de l’eau. Enfin, on n glige toute infiltration possible des algues dans le s diment et on consid re que le deuxi me compartiment est immobile (il ne s’enfonce pas dans le s diment). Les quantit s introduites sont r pertori es dans le tableau 1.

Concernant les daphnies, on fait l’hypoth se qu’elles se situent en permanence dans une zone o  il y a des algues, c’est- -dire dans les compartiments 1 et 2 : cette hypoth se est vraisemblable puisqu’il s’agit de leur seule source de nourriture. Dans la suite de l’ tude, on ne consid rera plus que la r union de ces deux compartiments (oubliant ainsi les parties “eau sans algues” et “s diment”), en gardant l’appellation *syst me algues-daphnies*.

Notation	Quantité	Unité
$D$	nombre de daphnies	daphnies
$S$	section du cylindre	cm <sup>2</sup>
$N_1(t)$	nombre d'algues en suspension au temps $t$	cell
$N_2(t)$	nombre d'algues sédimentées au temps $t$	cell
$N_1^0$	nombre d'algues en suspension à $t = 0$	cell
$H_1(t)$	hauteur du compartiment 1 au temps $t$	cm
$H_2$	hauteur du compartiment 2	cm
$H_0$	hauteur totale de la colonne d'eau (= $H_1(0)$ )	cm
$V_1(t)$	volume du compartiment 1 au temps $t$	mL
$V_2$	volume du compartiment 2	mL
$V_1^0$	volume du compartiment 1 à $t = 0$	mL

TABLEAU 1 – Quantités intervenant dans le modèle du système algues-daphnies

## 2.4 Ecriture des équations différentielles

On fait l'hypothèse d'une croissance algale continue dans le temps, et on souhaite décrire l'état du système comme une fonction continue du temps : on opte donc pour un modèle basé sur des équations différentielles. On pourrait songer à utiliser des équations aux dérivées partielles, mais on les juge trop complexes dans le cas présent, par rapport au niveau de connaissance physique du microcosme. De plus, comme les données ne concernent que deux classes d'algues, un modèle basé uniquement sur ces deux classes semble pertinent.

On va donc construire un système de deux équations différentielles ordinaires décrivant l'évolution de  $N_1$  et  $N_2$ , en prenant le **jour** comme unité de temps. Pour cela, on va d'abord s'intéresser séparément aux différents phénomènes impliqués. L'idée est de décrire chaque phénomène indépendamment des autres sur un intervalle de temps  $dt$  supposé très petit, de manière à pouvoir ensuite additionner les variations de  $N_1$  et  $N_2$  dues à chaque phénomène.

### 2.4.1 Système sans daphnies

Commençons par le système le plus simple, à savoir les algues seules. Deux phénomènes interviennent : la croissance de la population, que l'on décrira par un modèle logistique, et la sédimentation, lors de laquelle une certaine quantité d'algues passe du compartiment 1 au 2.

## Croissance logistique

Considérons une population de  $N(t)$  algues. On combine la reproduction et la mortalité des algues en un *taux de croissance intrinsèque* noté  $r$  (exprimé en  $\text{jour}^{-1}$ ). De plus, on prend en compte le phénomène de compétition intra-spécifique : la croissance des algues est limitée par la quantité maximale d'algues pouvant vivre dans le milieu, appelée *capacité limite*  $K$ . On obtient le modèle logistique (cf. fig. 5), introduit par Pierre François Verhulst vers 1840 [5] :

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \frac{N(t)}{K}\right),$$

et la variation de  $N$  sur un petit intervalle de temps  $dt$  s'écrit :

$$dN = rN \left(1 - \frac{N}{K}\right) dt. \quad (2.1)$$

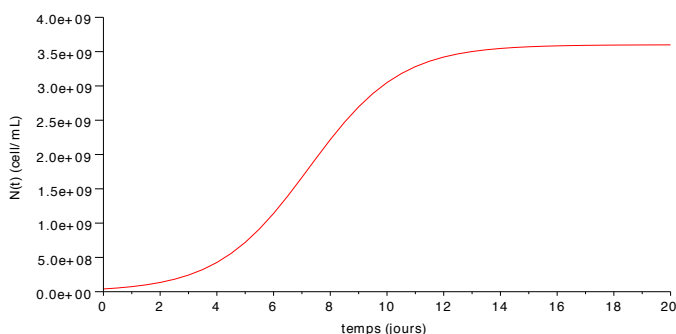


FIGURE 5 – Allure d'une solution de l'équation logistique ( $r = 0,6$ ,  $K = 3,6 \times 10^9$ ,  $N(0) = 4 \times 10^7$ )

On considère que les algues ont une croissance logistique dans les deux compartiments. On écrit donc (2.1) pour chaque compartiment avec l'indice qui lui correspond :

$$\begin{cases} dN_1 = r_1 N_1 \left(1 - \frac{N_1}{K_1}\right) dt \\ dN_2 = r_2 N_2 \left(1 - \frac{N_2}{K_2}\right) dt. \end{cases}$$

Mais il reste un aspect à prendre en compte : la capacité limite dépend de la quantité de nutriments disponibles, et donc du volume que les algues peuvent occuper. En effet, on considère que les nutriments contenus dans l'eau diffusent très peu et que les algues ne peuvent absorber que les nutriments situés à leur hauteur. Or, le volume du compartiment 2 est supposé constant, mais pas celui du compartiment 1. Il faut donc définir  $K_1$  comme une fonction de  $V_1$ . On fait l'hypothèse que  $K_1$  est proportionnelle à  $V_1$  :

$$K_1(t) = K_1^0 \frac{V_1(t)}{V_1(0)}$$

où  $K_1^0 = K_1(0)$  est la capacité limite des algues en suspension lorsqu'elles occupent toute la colonne d'eau. On obtient immédiatement :

$$K_1(t) = K_1^0 \frac{H_1(t)S}{H_0S} = K_1^0 \frac{H_1(t)}{H_0}.$$

### Prise en compte de la sédimentation

On s'intéresse maintenant au phénomène de sédimentation. On fait d'abord deux hypothèses simplificatrices :

- à tout instant, les algues de chaque compartiment sont uniformément réparties ;
- la vitesse de sédimentation des algues en suspension est la même partout.

Pour modéliser la sédimentation, on effectue un “bilan de quantité d'algues” par analogie à la thermodynamique (cf. fig. 6). On se place à l'instant  $t$ . Soit  $dN_2$  le nombre d'algues transférées du premier compartiment au deuxième pendant  $dt$ , on a :

$$dN_2 = C_1(t)dV$$

où  $C_1(t)$  est la concentration en algues du compartiment 1 au temps  $t$  et  $dV$  le volume d'algues en suspension qui sédimentent pendant  $dt$ . Avec les hypothèses que l'on vient de faire, on obtient <sup>1</sup> :

$$\begin{cases} dV = -H_1'(t)Sdt \\ C_1(t) = \frac{N_1(t)}{H_1(t)S} \end{cases}$$

et par conséquent

$$dN_2 = N_1(t) \frac{-H_1'(t)}{H_1(t)}.$$

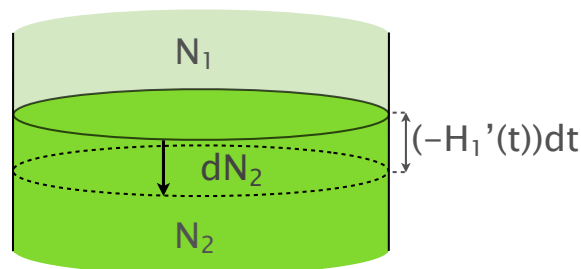


FIGURE 6 – Schéma du processus de sédimentation : passage de  $t$  à  $t + dt$

1. Il convient de noter le signe “-”, qui vient de la convention que l'on a pris : ici  $H_1$  est décroissante. Cette convention peut paraître étrange mais elle permet d'alléger l'écriture par la suite.



Enfin, sachant que les algues qui se déposent sur le sédiment sont exactement celles qui sortent du compartiment 1, on a :

$$dN_1 = -dN_2 = N_1(t) \frac{H_1'(t)}{H_1(t)}.$$

On obtient ainsi les équations différentielles du système sans daphnies :

$$(E_1) \begin{cases} N_1'(t) = r_1 N_1(t) \left(1 - \frac{N_1(t)}{K_1(t)}\right) + N_1(t) \frac{H_1'(t)}{H_1(t)} \\ N_2'(t) = r_2 N_2(t) \left(1 - \frac{N_2(t)}{K_2}\right) - N_1(t) \frac{H_1'(t)}{H_1(t)}. \end{cases}$$

### 2.4.2 Système avec daphnies

On va maintenant prendre en compte la présence des daphnies. Là encore, deux phénomènes entrent en jeu : la consommation d'algues par les daphnies dans chaque compartiment, et la "stratégie" de consommation, c'est-à-dire le fait que les daphnies mangent plutôt les algues de tel ou tel compartiment.

#### Consommation des daphnies

*Daphnia magna* est régulièrement étudiée et modélisée [6] : les caractéristiques individuelles des daphnies sont plutôt bien connues, notamment l'évolution de leur taille au cours du temps [7].

Le **taux d'ingestion**  $TI$  (cell/jour) d'une daphnie en fonction de la concentration algale  $x$  (cell/mL) est modélisé par une fonction affine par morceaux :

$$TI(x) = \begin{cases} \frac{TI_{max}}{ILL} x & \text{si } 0 \leq x \leq ILL \\ TI_{max} & \text{si } x > ILL \end{cases} \quad (2.2)$$

où  $ILL$  est l'*Incipient Limit Level* (concentration au-delà de laquelle la daphnie ne peut pas consommer plus) et  $TI_{max}$  est le taux d'ingestion maximum (cf. fig. 7).

Or  $ILL$  et  $TI_{max}$  ne sont pas des constantes : ces grandeurs dépendent de l'âge des daphnies et donc du temps. On les modélise par des fonctions qui sont des solutions de l'équation de croissance de von Bertalanffy

$$y'(t) = a(y_\infty - y(t))$$

dont la solution vérifiant la condition initiale  $y(0) = 0$  est (cf. fig. 8) :

$$y(t) = y_\infty(1 - \exp(-at)).$$

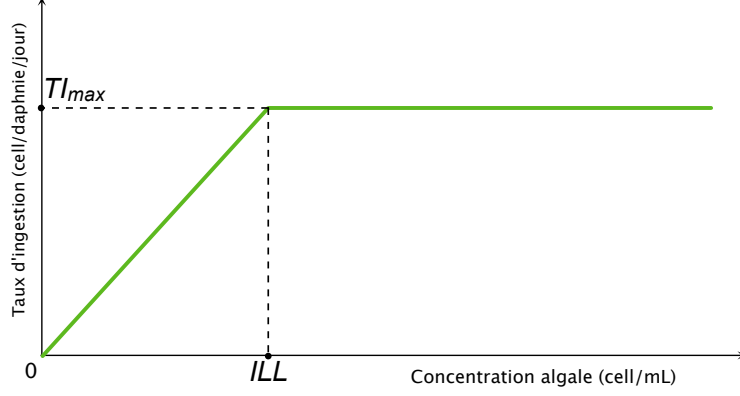


FIGURE 7 – Graphe de la fonction  $TI$  lorsque  $ILL$  et  $TI_{max}$  sont fixés

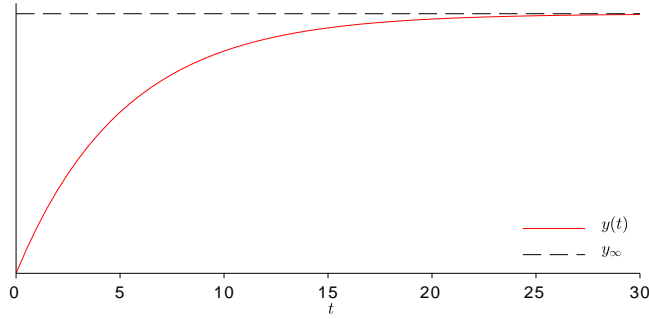


FIGURE 8 – Allure d'une solution de l'équation de von Bertalanffy ( $a = 0,2$ )

Cette équation, formulée par Ludwig von Bertalanffy en 1934 [8], permet en fait de décrire l'évolution de la taille au cours du temps. Cependant, la modélisation de  $ILL$  et  $TI_{max}$  par la même équation semble cohérente si l'on considère que ces quantités dépendent linéairement de la taille des daphnies. On définit donc les fonctions  $ILL$  et  $TI_{max}$  de la manière suivante :

$$\begin{cases} ILL(t) = ILL_{\infty} (1 - \exp(-a_1 t)) \\ TI_{max}(t) = TI_{\infty} (1 - \exp(-a_2 t)) \end{cases} \quad (2.3)$$

où  $ILL_{\infty}$ ,  $TI_{\infty}$ ,  $a_1$  et  $a_2$  sont des constantes.

De plus, on fait l'hypothèse que les daphnies sont introduites à l'âge d'un jour à  $t = 0$ . Pour prendre en compte ce décalage ainsi que les fonctions  $ILL$  et  $TI_{max}$ , on introduit la fonction  $g$  (en cell/daphnie) définie comme le taux d'ingestion à la concentration  $x$  et à l'âge  $t + 1$  :

$$g(x, t) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{TI_{max}(t+1)}{ILL(t+1)} x & \text{si } 0 < x \leq ILL(t+1) \\ TI_{max}(t+1) & \text{si } x > ILL(t+1). \end{cases} \quad (2.4)$$

## Stratégie d'alimentation des daphnies

Il faut maintenant modéliser la manière dont les daphnies se nourrissent. C'est l'un des aspects les plus problématiques : mangent-elles de préférence là où la concentration en algue est la plus élevée ? Là où il y a le plus d'algues indépendamment de leur concentration ?

Après comparaison de plusieurs hypothèses, on se base sur la suivante : les daphnies ont tendance à aller là où la quantité d'algues *disponible* est la plus grande, indépendamment de leur concentration. La totalité des algues du compartiment 1 est considérée comme *disponible*, et le nombre d'algues du compartiment 2 est affecté d'un *coefficient de disponibilité*, noté  $a_3$ . Ce coefficient illustre le fait qu'en réalité il y a infiltration des algues dans le sédiment, et les algues sédimentées ne sont donc pas toutes accessibles aux daphnies : seule la partie non infiltrée est accessible. Ainsi, si le modèle s'avère cohérent,  $a_3$  devrait correspondre à la proportion d'algues non infiltrées.

Enfin, sachant que la vitesse de nage des daphnies est très supérieure aux vitesses de croissance et de sédimentation des algues, on suppose que le choix du compartiment se fait instantanément. Cela nous amène à définir une "fonction de choix" notée  $h$  et basée sur la fonction de Heaviside, que l'on appliquera au couple  $(N_1, a_3N_2)$  (cf. fig. 9). Ainsi, pour tout  $(x, y) \in \mathbb{R}^2$  :

$$h(x, y) = \begin{cases} 0 & \text{si } x \leq y \\ 1 & \text{si } x > y. \end{cases}$$

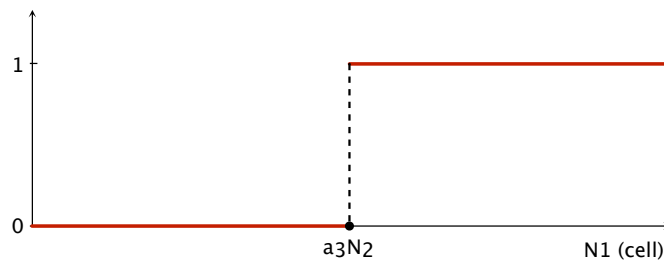


FIGURE 9 – Graphe de la fonction  $h(\cdot, a_3N_2)$  lorsque  $a_3N_2$  est fixé

Cette définition de  $h$  est facile à utiliser dans notre modèle à deux compartiments, en remarquant que  $h(y, x) = 1 - h(x, y)$  (cf. section 2.6). Cependant, elle soulève un problème important : la fonction  $h$  n'est pas continue, ce qui sera ennuyeux lorsqu'il s'agira de résoudre le système d'équations différentielles. On va donc contourner ce problème en approchant  $h$  par

une fonction  $\tilde{h}$  définie par :

$$\tilde{h}(x, y) = \frac{1}{1 + \exp\left(-\frac{100}{1+x+y}(x-y)\right)} \quad (2.5)$$

où le facteur  $\frac{100}{1+x+y}$  permet d'éviter le phénomène d'*overflow* (nombres trop grands pour le logiciel de résolution numérique, en l'occurrence Scilab) tout en gardant une approximation (arbitrairement) satisfaisante. On a encore  $\tilde{h}(y, x) = 1 - \tilde{h}(x, y)$ , et de plus  $\tilde{h}$  est de classe  $\mathcal{C}^1$ , comme l'illustre la figure 10.

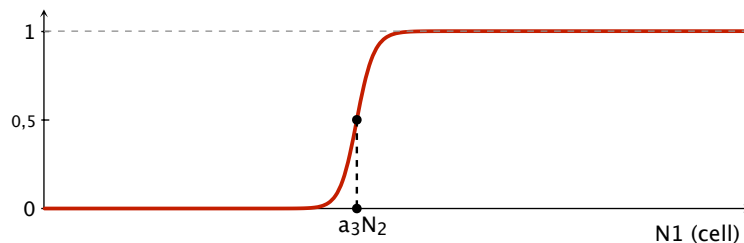


FIGURE 10 – Graphe de la fonction  $\tilde{h}(\cdot, a_3N_2)$  lorsque  $a_3N_2$  est fixé

Enfin, comme il n'est pas certain que les daphnies consomment les algues du sédiment au même taux que celles de la colonne d'eau, on introduit un coefficient d'*adaptation du taux d'ingestion*, noté  $a_4$ . Par conséquent, si  $D$  est le nombre de daphnies du système (supposé ici constant), celles-ci consomment pendant l'intervalle de temps  $dt$  :

- dans le compartiment 1 :  $D \cdot g\left(\frac{N_1(t)}{S \times H_1(t)}, t\right) \cdot \tilde{h}(N_1(t), a_3N_2(t))dt$
- dans le compartiment 2 :  $D \cdot a_4 \cdot g\left(\frac{N_2(t)}{V_2}, t\right) \cdot (1 - \tilde{h}(N_1(t), a_3N_2(t)))dt.$

## 2.5 Expression de la vitesse de sédimentation

Vient à présent la dernière étape de construction du modèle mathématique : la détermination des fonctions  $H_1$  et  $H_1'$ . On peut faire le lien avec la vitesse de sédimentation intuitive (vitesse de chute des algues en suspension) grâce à la définition suivante.

**Définition 1.** On note  $v$  la **vitesse de sédimentation**, définie par :

$$v(t) = -H_1'(t).$$

En pratique, on va travailler avec  $H_1$  pour simplifier les calculs. On opte pour une démarche empirique : à partir d'expériences adaptées, on détermine l'équation de décroissance de  $N_1$ , et on en déduit les expressions de  $H_1$  et  $H_1'$ .

Cette approche est basée sur les expériences réalisées par Gauci [3]. On introduit des algues dans un b cher contenant 2 litres d'eau, puis on remue l'eau et on place le b cher dans le noir : les algues ne pouvant pas r aliser la photosynth se, la croissance est nulle. Ainsi, on observe uniquement le ph nom ne de s dimentation. A partir des observations (cf. fig. 11), on peut proposer le mod le suivant pour d crire le ph nom ne :

$$N_1(t) = N_1(0) \exp(-at) \quad (2.6)$$

o   $a$  est un coefficient r el positif. On a alors :

$$\begin{cases} H_1(t) = H_0 \exp(-at) \\ H'_1(t) = -aH_0 \exp(-at). \end{cases} \quad (2.7)$$

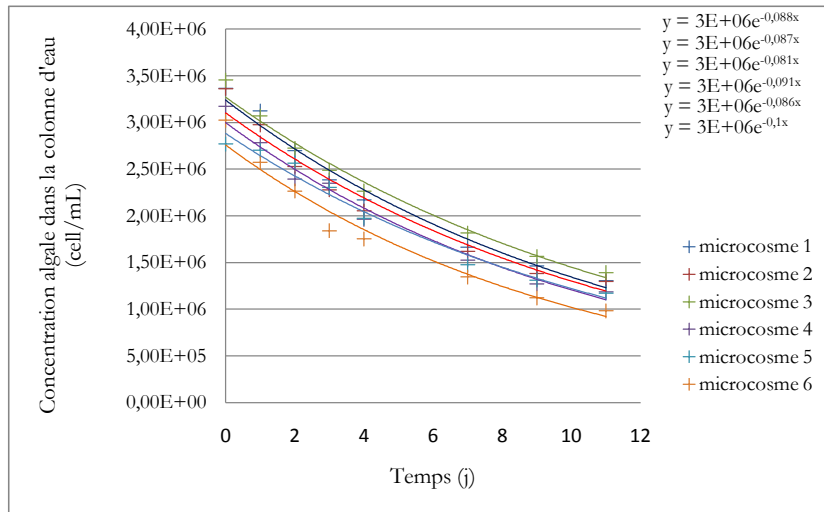


FIGURE 11 – Evolution de la concentration algale dans la colonne d'eau en l'absence de lumi re, et approximation exponentielle (6 essais en parall le). (Image : F.-X. Gauci)

*D monstration.* Sans croissance et sans daphnie, le syst me devient :

$$\begin{cases} N'_1(t) = N_1(t) \frac{H'_1(t)}{H_1(t)} \\ N'_2(t) = -N_1(t) \frac{H'_1(t)}{H_1(t)}. \end{cases}$$

Comme pour tout  $t \in \mathbb{R}_+$ ,  $N_1(t) > 0$ , il vient directement

$$\frac{N'_1(t)}{N_1(t)} = \frac{H'_1(t)}{H_1(t)},$$

et en int grant (d riv e logarithmique) :

$$\frac{N_1(t)}{N_1(0)} = \frac{H_1(t)}{H_0}.$$

On obtient immédiatement d'après (2.6) :

$$H_1(t) = \frac{H_0}{N_1(0)} N_1(t) = H_0 \exp(-at),$$

d'où le résultat (2.7). □

On peut donc inclure dans notre modèle :

$$\begin{cases} H_1(t) = H_0 \exp(-a_5 t) \\ H_1'(t) = -a_5 H_0 \exp(-a_5 t) \end{cases} \quad (2.8)$$

avec  $a_5$  le *coefficient de sédimentation* (réel strictement positif). Avant d'incorporer ces expressions dans le modèle, on pose  $V_1^0 = V_1(0) = S \times H_0$  pour alléger l'écriture.

## 2.6 Modèle mathématique du système algues-daphnies

En faisant une synthèse de ce qui précède, on obtient notre modèle mathématique.

$$(E_2) \begin{cases} N_1'(t) = r_1 N_1(t) \left(1 - \frac{N_1(t)}{K_1^0 \exp(-a_5 t)}\right) - a_5 N_1(t) - D \cdot g\left(\frac{N_1(t)}{V_1^0 \exp(-a_5 t)}, t\right) \cdot \tilde{h}(N_1(t), a_3 N_2(t)) \\ N_2'(t) = r_2 N_2(t) \left(1 - \frac{N_2(t)}{K_2}\right) + a_5 N_1(t) - D \cdot a_4 \cdot g\left(\frac{N_2(t)}{V_2}, t\right) \cdot (1 - \tilde{h}(N_1(t), a_3 N_2(t))) \end{cases}$$

De plus, on fixe les conditions initiales  $N_1(0) = N_1^0$  et  $N_2(0) = 0$ , et on a ainsi un problème de Cauchy. Nous verrons qu'il vérifie les conditions de Cauchy-Lipschitz, ce qui permet d'affirmer, pour des valeurs des paramètres fixées, l'existence et l'unicité d'une solution maximale.

En pratique, nous allons étudier à la fois les solutions de  $(E_2)$  et celles de  $(E_1)$  (système sans daphnies), chacune étant à valeurs dans  $\mathbb{R}^2$ . Nous appellerons ainsi *solution* un quadruplet de fonctions, constitué des solutions respectives de  $(E_1)$  et  $(E_2)$  pour un certain jeu de valeurs des paramètres. Avant d'aborder l'aspect mathématique du problème et afin d'avoir une première idée du comportement qualitatif de ces solutions, nous allons d'abord présenter un exemple en résolvant numériquement l'équation différentielle à l'aide du logiciel Scilab.

### 2.6.1 Exemple de courbes obtenues

En fixant les valeurs des paramètres (cf. annexe A.1), on obtient une solution dont on trace les composantes sur la figure 12, superposées aux données présentées dans la section 2.2. Les figures à gauche représentent  $N_1$  sur 21 jours, sans daphnies (en haut) et avec daphnies (en bas). De même, les figures à droite représentent le nombre d'algues sédimentées  $N_2$ .

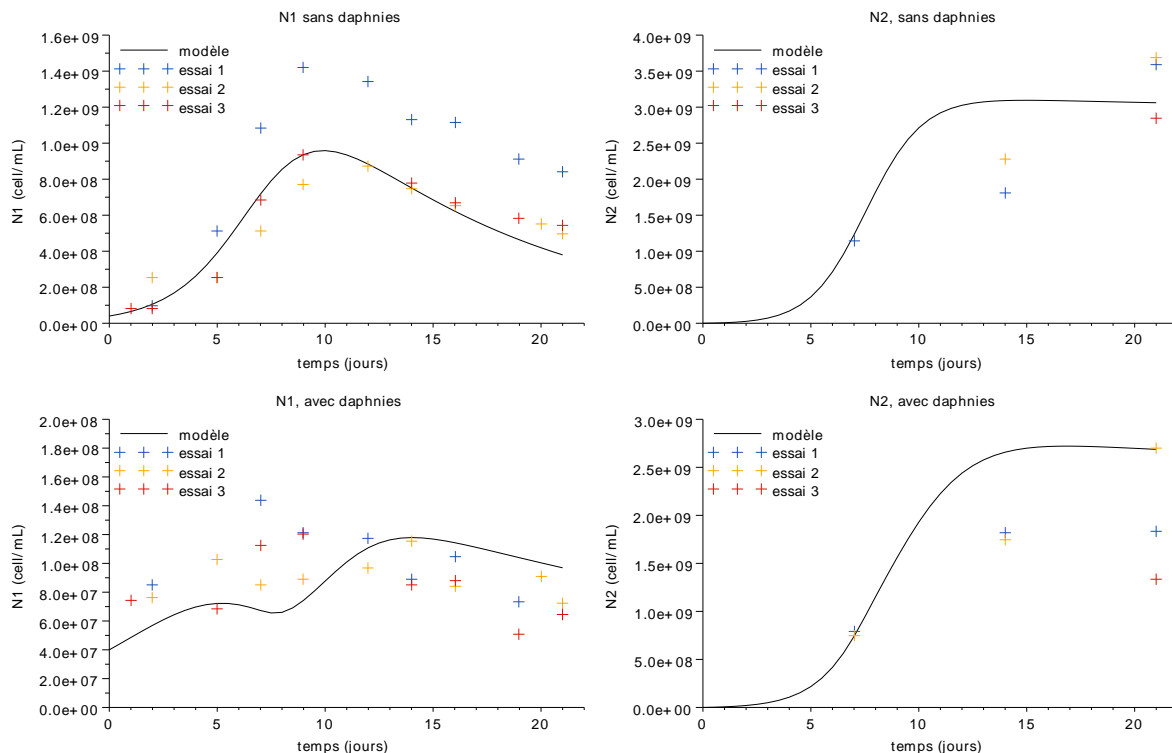


FIGURE 12 – Exemple de solution pour le système algues-daphnies

Il convient de noter la différence d'échelle lors de la présence des daphnies, qui illustre le fait que ces dernières consomment une grande partie des algues.

### 2.6.2 Premières propriétés

Le système  $(E_2)$  que l'on considère est un système d'équations différentielles ordinaires, d'ordre 1, non linéaires. On peut l'écrire sous la forme :

$$\mathbf{X}'(t) = F(t, \mathbf{X}(t)) \quad (2.9)$$

où  $F : I \times \Omega \mapsto \mathbb{R}^2$  continue et localement lipschitzienne par rapport à sa deuxième variable, avec  $\mathbb{R}_+ \subset I$  et  $\mathbb{R}_+ \times \mathbb{R}_+ \subset \Omega$  (pour plus de précisions, cf. annexe A.2). Etant donné le vecteur

$$\mathbf{X}_0 = \begin{pmatrix} N_1^0 \\ 0 \end{pmatrix},$$

on sait d'après le théorème de Cauchy-Lipschitz qu'il existe une unique solution maximale vérifiant la condition initiale  $\mathbf{X}(0) = \mathbf{X}_0$ . On note

$$\mathbf{X}(t) = \begin{pmatrix} N_1(t) \\ N_2(t) \end{pmatrix}$$

cette solution, et  $J$  son intervalle de définition (ouvert de  $I$ ). On a déjà quelques résultats utiles pour appréhender le comportement de la solution, grâce à la proposition suivante.

**Proposition 1.** *On suppose que  $N_1^0 > 0$ . Alors :*

- (i)  $\forall t \in J, N_1(t) > 0$ ;
- (ii)  $\forall t \in J \cap \mathbb{R}_+^*, N_2(t) > 0$ ;
- (iii)  $\mathbb{R}_+ \subset J$ .

*Démonstration.* Cf. annexe A.3. □

### 2.6.3 Précisions sur les constantes et les paramètres

Dans la suite, on assimilera  $\mathbf{X}$  à sa restriction sur  $\mathbb{R}_+$ , la seule partie qui fait sens du point de vue biologique. De plus, les valeurs de  $N_1^0$ ,  $D$  et  $V_1^0$  sont fixées dans les expériences et sont considérées comme des **constantes** (cf. tableau 2). Il reste alors 12 **paramètres**, notés  $x_1, \dots, x_{12}$  (cf. tableau 3), qui peuvent varier sur des intervalles respectifs  $A_1, \dots, A_{12}$  à définir. Ces intervalles doivent être appropriés, c'est-à-dire : (1) ouverts et sur lesquels le problème de Cauchy est bien défini ; (2) biologiquement cohérents. Si les  $A_i$  sont des intervalles ouverts, on constate que la condition suivante est suffisante (cf. annexe A.2) :

$$\forall i \in \{1, \dots, 12\}, A_i \subset ]0, +\infty[ \text{ et } A_3 \subset ]0, 1[ \quad (2.10)$$

Pour fixer les idées, on définit dans un premier temps les  $A_i$  par :  $\forall i \neq 3, A_i = ]0, +\infty[$  et  $A_3 = ]0, 1[$ . Cependant, on verra dans la suite qu'en exploitant les informations fournies par les expériences et la littérature, il est possible de fixer une borne supérieure spécifique pour chaque intervalle, voire les deux bornes : dans les chapitres suivants, on parlera d'*information a priori sur les paramètres*. Ainsi, à  $t$  fixé, on peut voir  $\mathbf{X}(t)$  comme une fonction de  $x_1, \dots, x_{12}$ , que l'on note  $(x_1, \dots, x_{12}) \mapsto \mathbf{X}_{x_1, \dots, x_{12}}(t)$ , et la proposition suivante donne un résultat fondamental.

**Définition 2.** *On écrit le modèle sous la forme d'une fonction  $\phi$ , définie par :*

$$\begin{aligned} \phi : \mathbb{R}_+ \times A_1 \times \dots \times A_{12} &\rightarrow \mathbb{R}^2 \\ (t, x_1, \dots, x_{12}) &\mapsto \mathbf{X}_{x_1, \dots, x_{12}}(t). \end{aligned}$$

**Proposition 2.** *La fonction  $\phi$  est continue sur  $\mathbb{R}_+ \times A_1 \times \dots \times A_{12}$ .*

*Démonstration.* Cf. annexe A.4. □

**Remarque 1.** *On obtient ainsi une propriété de continuité de la sortie du modèle (i.e. les valeurs renvoyées par celui-ci) par rapport aux paramètres. Ce résultat de régularité sera important pour appliquer de manière licite les méthodes présentées dans les chapitres suivants.*



Constante	Valeur	Unité
$D$	10	Dm (daphnies)
$N_1^0$	$4 \times 10^7$	cell (cellules)
$V_1^0$	2000	mL

TABLEAU 2 – Constantes du modèle du système algues-daphnies

Indice	Paramètre	Signification	Unité
1	$a_1$	vitesse de convergence de $ILL$	jour <sup>-1</sup>
2	$a_2$	vitesse de convergence de $TI_{max}$	jour <sup>-1</sup>
3	$a_3$	disponibilité des algues sédimentées	[sans]
4	$a_4$	adaptation du taux d'ingestion	[sans]
5	$a_5$	coefficient de sédimentation	jour <sup>-1</sup>
6	$ILL_\infty$	seuil de concentration algale limite	cell/mL
7	$TI_\infty$	taux d'ingestion maximal limite	cell/Dm/jour
8	$V_2$	volume du compartiment 2	mL
9	$r_1$	taux de croissance des algues en suspension	jour <sup>-1</sup>
10	$r_2$	taux de croissance des algues sédimentées	jour <sup>-1</sup>
11	$K_1^0$	capacité limite initiale du compartiment 1	cell
12	$K_2$	capacité limite du compartiment 2	cell

TABLEAU 3 – Paramètres du modèle du système algues-daphnies

## Chapitre 3

# Analyse de sensibilité

### 3.1 Motivations et principe de l'analyse de sensibilité

Nous disposons à présent d'un modèle mathématique du système algues-daphnies : afin d'évaluer la qualité de ce modèle, il est nécessaire de savoir si celui-ci décrit bien les observations faites lors des expériences. Or, si l'on fixe les valeurs des paramètres d'après la littérature (du moins lorsque c'est possible), les fonctions retournées par le modèle ne semblent pas s'ajuster correctement aux données (cf. fig. 12). On est alors tenté de chercher “à la main” les valeurs des paramètres qui permettraient un ajustement idéal : un tel réglage empirique est ardu, et ce d'autant plus que le nombre de paramètres est grand.

En outre, la notion d'*ajustement aux données* est pour le moment très subjective. Il est donc nécessaire de la définir rigoureusement, et plusieurs approches sont possibles comme nous le verrons dans le chapitre 4. Quelle que soit l'approche choisie, il apparaît que la complexité de résolution croît très rapidement avec le nombre de paramètres. Enfin, nous n'avons qu'une connaissance partielle du système algues-daphnies : nous disposons de peu de données, et celles-ci comportent certainement des erreurs de mesure. Cela engendre une *incertitude* sur le système algues-daphnies, et *a fortiori* une incertitude sur notre modèle et ses paramètres.

Dans ce contexte, il semble important de comprendre comment se comporte le modèle vis-à-vis des paramètres : existe-t-il des paramètres plus influents sur la sortie (*i.e.* les quatre courbes) que les autres ? Peut-on quantifier cette influence ? Est-elle due à chaque paramètre considéré de manière isolée ou à une interaction entre eux ? En fait, la *sensibilité* peut être définie comme *l'influence de l'incertitude* : l'analyse de sensibilité a donc pour ambition d'apporter une meilleure compréhension du modèle mathématique dans une perspective d'utilisation de données<sup>1</sup>.

---

1. Pour plus de précisions, on renvoie à l'article très didactique de Julien Jacques [9].

## 3.2 Choix de la méthode

Il existe trois types de méthodes d'analyse de sensibilité : les méthodes de *screening*, qui permettent de classer les paramètres de manière qualitative en fonction de leur influence, les méthodes d'analyse locale, qui évaluent quantitativement la sensibilité de la variable de sortie à une petite variation autour d'une valeur donnée (calcul des dérivées partielles par rapport à chaque paramètre), et enfin les méthodes d'analyse de sensibilité globale, qui étudient les variations de la sortie sur tout le domaine de définition des paramètres.

Les méthodes de *screening* apportent seulement un classement des paramètres par ordre d'importance, sans information sur la taille des écarts entre paramètres, et sont surtout utilisées lorsque le nombre de paramètre est très grand ( $> 100$  par ex.). Les méthodes d'analyse locale nécessitent une certaine valeur autour de laquelle on fait varier les paramètres, or c'est précisément ce type d'information qu'il nous manque. Par conséquent, nous nous sommes intéressés dans ce stage à l'analyse de sensibilité globale. Celle-ci est fondée sur une approche probabiliste : les paramètres sont vus comme une variable aléatoire  $X$ , et sous certaines conditions de régularité, la sortie du modèle est elle-même une variable aléatoire réelle  $Y$  : il s'agit alors de déterminer quelle part de la variance de  $Y$  est due à chaque paramètre.

### 3.2.1 Application au modèle du système algues-daphnies

On se place à  $t \in \mathbb{R}_+$  fixé, et on considère la fonction  $\theta_t$  définie sur  $A_1 \times \dots \times A_{12}$  par :

$$\forall (x_1, \dots, x_{12}) \in A_1 \times \dots \times A_{12}, \theta_t(x_1, \dots, x_{12}) = \phi(t, x_1, \dots, x_{12})$$

où  $\phi$  est la fonction de la définition 2. Soit  $X = (X_1, \dots, X_{12})$  une variable aléatoire à valeurs dans  $A_1 \times \dots \times A_{12}$  (vecteur aléatoire). Comme  $\theta_t$  est continue (proposition 2),  $\theta_t(X)$  est aussi un vecteur aléatoire, dans  $\mathbb{R}^2$ . En fait,  $\theta_t$  dépend par définition du nombre de daphnies  $D$  : on note  $(Y_1, Y_2) = \theta_t(X)$  pour  $D = 0$ , et  $(Y_3, Y_4) = \theta_t(X)$  pour  $D = 10$ . On obtient ainsi 4 variables aléatoires réelles (v.a.r.), dont on étudiera séparément les 4 variances respectives. Les résultats qui suivent ayant une portée générale, la notation  $Y$  désignera indifféremment  $Y_1, Y_2, Y_3$  ou  $Y_4$ .

On fait l'hypothèse que les v.a.r.  $X_1, \dots, X_{12}$  sont **indépendantes**. D'un point de vue biologique, cela revient à supposer que les paramètres sont *statiques*, même si les caractéristiques physiques qu'ils représentent ont par nature un caractère aléatoire, *i.e.* s'il y a de la *variabilité*. En d'autres termes, **la variabilité de la caractéristique physique représentée par un paramètre  $x$  n'a aucun lien avec la variabilité de la caractéristique physique représentée par un autre paramètre  $\hat{x}$** . Cela serait faux si, par exemple, il s'avérait que les algues

en suspension soient capables de modifier dynamiquement et de manière non négligeable leur *taux de croissance intrinsèque*  $r_1$  en fonction du taux d'ingestion des daphnies<sup>2</sup>.

### 3.2.2 Nécessité d'une décomposition de la variance

On pourrait être tenté, vu l'apparente simplicité des fonctions  $g$  et  $\tilde{h}$ , de résoudre analytiquement  $(E_2)$  sur chaque intervalle où ces fonctions sont "affines". Mais à cause de  $TI_{max}$  et  $ILL$  (définies à partir d'exponentielles), il apparaît que les équations de  $(E_2)$  sont des équations de Ricatti que l'on ne sait pas résoudre analytiquement : la forme analytique de la fonction  $\theta_t$  n'est donc pas connue. Ainsi, on ne peut pas compter sur un éventuel calcul analytique de  $\text{Var}(Y)$  en fonction des  $\text{Var}(X_i)$ .

Cette situation est en fait relativement courante, et plusieurs méthodes ont été développées pour la traiter<sup>3</sup>. L'idée générale est d'exhiber une certaine décomposition de  $\text{Var}(Y)$ , puis de fournir un moyen de *quantifier* l'influence des paramètres en définissant des *indices de sensibilité* basés sur cette décomposition. Nous nous basons ici sur les travaux du mathématicien Sobol [11].

## 3.3 Les indices de Sobol

La définition des indices de Sobol repose avant tout sur un résultat d'analyse dans les espaces de type  $L^1(\mu)$  où  $\mu$  est une mesure quelconque, que l'on peut ensuite appliquer dans un cadre probabiliste (avec  $\mu = P_X$  par exemple). Nous présentons d'abord la définition des indices de Sobol dans le cadre général, puis nous l'appliquons au problème qui nous intéresse, de manière à obtenir concrètement ces indices de sensibilité.

### 3.3.1 Définition des indices de Sobol dans un cadre général

Soient  $\mu_1, \dots, \mu_p$  des mesures sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . On note  $\mu = \mu_1 \otimes \dots \otimes \mu_p$  et on considère l'espace mesuré  $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), \mu)$ . Soit  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  telle que  $f \in L^1(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), \mu)$ . Pour  $x \in \mathbb{R}^p$ , on note  $x = (x_1, \dots, x_p)$ . On peut décomposer  $f$  en une somme de fonctions de dimensions croissantes :

$$f(x) = f_0 + \sum_{i=1}^p f_i(x_i) + \sum_{1 \leq i < j \leq p} f_{i,j}(x_i, x_j) + \dots + f_{1,\dots,p}(x_1, \dots, x_p)$$

---

2. L'hypothèse d'indépendance pourrait ainsi, en s'avérant fautive, être un bon moyen de révéler une éventuelle carence dans le modèle mathématique.

3. Pour un état de l'art sur les différentes méthodes d'analyse de sensibilité globale, on renvoie à [10].

où  $f_0$  est une constante et  $f_{i_1, \dots, i_s}$  ne dépend que de  $x_{i_1}, \dots, x_{i_s}$ . Ceci peut s'écrire de façon condensée :

$$f(x) = f_0 + \sum_{s=1}^p \sum_{i_1 < \dots < i_s} f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}). \quad (3.1)$$

**Définition 3** (Décomposition de Sobol). *Une décomposition de  $f$  de la forme (3.1) est appelée décomposition de Sobol si chaque terme  $f_{i_1, \dots, i_s}$  vérifie :*

$$\forall k \in \{1, \dots, s\}, \int_{\mathbb{R}} f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) \mu_{i_k}(dx_{i_k}) = 0.$$

**Remarque 2.** *Si  $f$  admet une décomposition de Sobol, on a directement :*

$$f_0 = \int_{\mathbb{R}^p} f(x) \mu(dx)$$

et pour tout  $(i_1, \dots, i_s) \neq (j_1, \dots, j_t)$ ,

$$\int_{\mathbb{R}^p} f_{i_1, \dots, i_s} f_{j_1, \dots, j_t} d\mu \stackrel{\text{déf}}{=} \int_{\mathbb{R}^p} f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) f_{j_1, \dots, j_t}(x_{j_1}, \dots, x_{j_t}) \mu(dx) = 0.$$

On a alors le théorème fondamental suivant, dont on trouve une preuve dans [12].

**Théorème 1.** *Pour toute fonction  $f \in L^1(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), \mu)$ , il existe une unique décomposition de Sobol de  $f$ .*

On considère donc la décomposition de Sobol de  $f$ , et on suppose désormais que  $f$  est de carré intégrable, *i.e.*  $f \in L^2(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), \mu)$ . On peut alors montrer que chaque  $f_{i_1, \dots, i_s}$  est aussi de carré intégrable : la remarque 2 se traduit alors par l'orthogonalité des  $f_{i_1, \dots, i_s}$  dans  $L^2(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), \mu)$  muni de son produit scalaire usuel. De plus, les quantités suivantes sont finies :

$$V = \int_{\mathbb{R}^p} f(x)^2 \mu(dx) - (f_0)^2,$$

$$V_{i_1, \dots, i_s} = \int_{\mathbb{R}^s} (f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}))^2 \mu_{i_1}(dx_{i_1}) \dots \mu_{i_s}(dx_{i_s}),$$

et sont appelées *variances* par Sobol. En mettant (3.1) au carré et en intégrant, on obtient :

$$V = \sum_{s=1}^p \sum_{i_1 < \dots < i_s} V_{i_1, \dots, i_s}. \quad (3.2)$$

**Définition 4.** *On définit les indices de sensibilité d'ordre  $s \in \{1, \dots, p\}$  par :*

$$S_{i_1, \dots, i_s} = \frac{V_{i_1, \dots, i_s}}{V}$$

pour tout  $\{i_1, \dots, i_s\} \subset \{1, \dots, p\}$ .

Il s'agit maintenant de comprendre le sens de ces indices de sensibilité. D'après l'égalité (3.2), les  $S_{i_1, \dots, i_s}$  sont positifs, et on a

$$\sum_{s=1}^p \sum_{i_1 < \dots < i_s} S_{i_1, \dots, i_s} = 1.$$

De plus, on voit facilement que  $S_{i_1, \dots, i_s} = 0$  si et seulement si  $f_{i_1, \dots, i_s} \equiv 0$ . Ainsi, la fonction  $f$  ne dépend pas d'un paramètre  $x_i$  si et seulement si tous les indices  $S_{i_1, \dots, i_s}$  tels que  $i \in \{i_1, \dots, i_s\}$  sont nuls. Ceci est déjà cohérent avec le sens intuitif de la sensibilité (si un indice  $S_{i_1, \dots, i_s}$  est non nul, alors  $f$  dépend de  $x_{i_1}, \dots, x_{i_s}$ ), mais on voudrait pouvoir utiliser des arguments plus quantitatifs comme : "plus les indices de sensibilité sont grands, plus la sensibilité est grande". Pour cela, on va se placer dans un cadre probabiliste et utiliser les notions de moyenne (ou espérance) et de variance (par exemple : une variance élevée pour la variable aléatoire  $Y$  signifie moralement que les valeurs prises par  $Y$  ont "tendance à beaucoup s'écartier de la moyenne").

### 3.3.2 Les indices de Sobol dans un cadre probabiliste

Pour appliquer les résultats précédents dans un cadre probabiliste, on utilise fortement la notion d'espérance conditionnelle, qui est assez abstraite dans sa définition générale. Les notations qui vont suivre s'appliquent à des v.a.r. quelconques, et on se concentrera plutôt sur le sens intuitif des espérances conditionnelles. Cependant, on donne dans l'annexe B des explications plus rigoureuses dans le cas particulier des variables à densité, pour lesquelles on a une formulation explicite de l'espérance conditionnelle.

On revient au cadre développé dans la section 3.2.1, et on considère  $X = (X_1, \dots, X_p)$  un vecteur aléatoire dans  $\mathbb{R}^p$  avec les variables  $X_i$  indépendantes. On va appliquer les résultats précédents aux mesures  $\mu_i = P_{X_i}$ . En effet, par indépendance des  $X_i$ , on a

$$\mu \stackrel{\text{déf}}{=} P_{X_1} \otimes \dots \otimes P_{X_p} = P_X.$$

On considère une fonction continue  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  dans  $L^2(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$ , et la v.a.r.  $Y = f(X)$ . Les quantités  $V$  et  $V_{i_1, \dots, i_s}$  sont alors de véritables variances (la deuxième égalité découlant de la définition 3) :

$$V = \int_{\mathbb{R}^p} f(x)^2 P_X(dx) - \left( \int_{\mathbb{R}^p} f(x) P_X(dx) \right)^2 = \text{Var}(f(X)) = \text{Var}(Y),$$

$$V_{i_1, \dots, i_s} = \int_{\mathbb{R}^s} (f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}))^2 P_{(X_{i_1}, \dots, X_{i_s})}(dx_{i_1}, \dots, dx_{i_s}) = \text{Var}(f_{i_1, \dots, i_s}(X_{i_1}, \dots, X_{i_s})),$$

et l'égalité (3.2) s'appelle *décomposition de la variance*.

Interprétons maintenant les quantités  $V_{i_1, \dots, i_s}$ . Pour simplifier les écritures, on note

$$\frac{P_X}{P_{X_i}} = P_{X_1} \otimes \dots \otimes P_{X_{i-1}} \otimes P_{X_{i+1}} \otimes \dots \otimes P_{X_p} = P_{(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p)},$$

la deuxième égalité étant due à l'indépendance des  $X_i$ . On écrit de la même façon  $\frac{P_X}{P_{X_i} \otimes P_{X_j}}$ ,  $\frac{P_X}{P_{X_i} \otimes P_{X_j} \otimes P_{X_k}}$ , etc. Ainsi, en intégrant l'égalité (3.1) successivement par rapport aux mesures  $P_X$ ,  $\frac{P_X}{P_{X_i}}$ ,  $\frac{P_X}{P_{X_i} \otimes P_{X_j}}$ ,  $\frac{P_X}{P_{X_i} \otimes P_{X_j} \otimes P_{X_k}}$ , ... , on montre que

$$\begin{aligned} f_0 &= \mathbb{E}(Y) \\ f_i(X_i) &= \mathbb{E}(Y|X_i) - f_0 \\ f_{i,j}(X_i, X_j) &= \mathbb{E}(Y|X_i, X_j) - f_i(X_i) - f_j(X_j) - f_0 \\ f_{i,j,k}(X_i, X_j, X_k) &= \mathbb{E}(Y|X_i, X_j, X_k) - f_{i,j}(X_i, X_j) - f_{i,k}(X_i, X_k) - f_{j,k}(X_j, X_k) \\ &\quad - f_i(X_i) - f_j(X_j) - f_k(X_k) - f_0 \\ &\dots \end{aligned}$$

et on obtient alors une expression explicite<sup>4</sup> des  $V_{i_1, \dots, i_s}$  :

$$\begin{aligned} V_i &= \text{Var}(\mathbb{E}(Y|X_i)) \\ V_{i,j} &= \text{Var}(\mathbb{E}(Y|X_i, X_j)) - V_i - V_j \\ V_{i,j,k} &= \text{Var}(\mathbb{E}(Y|X_i, X_j, X_k)) - V_{i,j} - V_{i,k} - V_{j,k} - V_i - V_j - V_k \\ &\dots \end{aligned}$$

Il est alors plus aisé d'interpréter les indices de Sobol : regarder  $V_i = \text{Var}(\mathbb{E}(Y|X_i))$ , c'est regarder "la tendance de la variable  $\mathbb{E}(Y|X_i)$  à s'écarter de sa moyenne". Or, comme  $\mathbb{E}(Y|X_i)$  peut s'écrire  $\phi(X_i)$  (où  $\phi$  est une fonction mesurable), calculer la variance de  $\mathbb{E}(Y|X_i)$  revient à quantifier sa tendance à varier lorsque  $X_i$  parcourt son ensemble de définition : c'est donc bien une sensibilité au sens intuitif que l'on observe. Pour visualiser ceci, on peut par exemple<sup>5</sup> fixer  $N$  valeurs de  $X_i$  (*i.e.* effectuer  $N$  tirages selon  $X_i$ ) et regarder, pour chaque valeur de  $X_i$ , la moyenne de  $Y$  vue comme fonction de  $(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_p)$ . Plus l'ensemble de ces  $N$  moyennes (vu simplement comme un ensemble de valeurs) présente une variance forte, plus la fonction  $f$  est sensible aux variations du paramètre  $x_i$  (en tout cas les variations associées à la loi de  $X_i$ ). L'indice de Sobol  $S_i = V_i/V$  exprime donc la sensibilité de  $f$  par rapport à  $x_i$  variant seul. De la même manière, les indices de Sobol d'ordre supérieur expriment

4. Preuve lourde en calculs (nombreux termes de covariance). Pour plus de précisions, voir [10] (p. 126-129).

5. C'est d'ailleurs la méthode que l'on va utiliser pour estimer les indices de Sobol.

la sensibilité par rapport à des groupes de paramètres variant simultanément, et uniquement simultanément : on enlève la sensibilité par rapport à chaque paramètre seul. Par exemple, on a  $S_{i,j} = \text{Var}(\mathbb{E}(Y|X_i, X_j))/V - S_i - S_j$ .

Il convient de noter que les indices de Sobol dépendent du choix de la loi de  $X$  (ou, ce qui revient au même, de la mesure  $\mu$ ). Un indice n'a donc de sens que comparé aux autres à  $P_X$  fixé.

### 3.3.3 Indices de sensibilité totaux

Nous disposons ainsi d'indices permettant de quantifier la sensibilité de  $Y$  par rapport à un paramètre ou un groupe de paramètres. Cependant, on aimerait bien avoir une notion globale de sensibilité par rapport à un paramètre (afin de pouvoir le fixer s'il est peu influent par exemple). On va donc définir un indice de sensibilité *total* par rapport à un paramètre, c'est-à-dire sous toutes ses formes, y compris lorsqu'il varie simultanément avec d'autres : l'idée est simplement de regrouper tous les indices liés à ce paramètre en un unique indice. On introduit donc l'*indice de sensibilité total selon  $x_i$* , noté  $S_{T_i}$  et défini par :

$$S_{T_i} = \sum_{\mathcal{I} \subset \{1, \dots, p\}, i \in \mathcal{I}} S_{\mathcal{I}}.$$

Par exemple, pour un modèle avec  $p = 3$  paramètres, on a  $S_{T_1} = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3}$ . Dans la suite, nous calculerons uniquement ces indices de sensibilité totaux.

### 3.3.4 Application au modèle algues-daphnies

Il s'agit maintenant d'appliquer la théorie des indices de Sobol à notre problème : pour étudier  $Y_1, Y_2, Y_3$  et  $Y_4$  nous allons bien sûr prendre à chaque fois comme fonction  $f$  une des deux composantes de  $\theta_t$ , avec  $p = 12$ .

Jusqu'à présent, nous n'avons pas eu à préciser la loi de  $X$ . Il apparaît quand même une condition nécessaire pour appliquer les résultats précédents : les composantes de  $\theta_t$  doivent être de carré intégrable par rapport à  $P_X$ . En prenant  $X$  à support compact  $B_1 \times \dots \times B_p$  inclus dans  $A_1 \times \dots \times A_p$ , la continuité de  $\theta_t$  nous assure que cette condition est vérifiée (on pourrait tenter de prendre un support égal à  $A_1 \times \dots \times A_p$ , mais l'intégrabilité semblerait alors fortement compromise).

En pratique, on prend tous les  $X_i$  de loi uniforme sur  $[0, 1]$  : cela revient à considérer tous les  $B_i$  égaux à  $[0, 1]$ , et à estimer la sensibilité avec la même importance en tout point de l'hypercube  $[0, 1]^p$ . Cela pose *a priori* un problème : pour garder une cohérence dans notre modèle, les  $B_i$  ne doivent pas être égaux à  $[0, 1]$  (cf. section 2.6.3). En fait cela n'a pas d'importance du



moment que l'on considère des intervalles fermés bornés  $B_i = [a_i, b_i] \subset A_i$ , quitte à étudier une version "normalisée" du modèle. On peut en effet contourner le problème en remplaçant dans notre modèle chaque paramètre  $x_i$  par  $a_i + (b_i - a_i)\hat{x}_i$  : tirer un échantillon  $x_i$  selon  $\mathcal{U}([a_i, b_i])$  équivaut alors à tirer  $\hat{x}_i$  selon  $\mathcal{U}([0, 1])$ . On étudie ainsi un nouveau modèle  $\hat{\theta}_t$  avec les paramètres  $\hat{x}_1, \dots, \hat{x}_p$ , mais l'influence de  $\hat{x}_i$  sur  $\hat{\theta}_t$  est la même que celle de  $x_i$  sur  $\theta_t$ .

Il reste encore un point important : le choix des intervalles  $B_i$ . Ils doivent être assez larges pour contenir toute valeur biologiquement cohérente des paramètres, sans être trop grands car les valeurs "extrêmes" ne donnent clairement pas un bon ajustement : inutile d'estimer la sensibilité du modèle là où l'on sait que les valeurs des paramètres sont aberrantes. En outre, cela nuirait à la précision des estimations des indices de sensibilité, la taille  $N$  des échantillons étant limitée. Nous utilisons dans ce rapport des valeurs déterminées à l'aide de la littérature et de simulations faites avec le logiciel Scilab, répertoriées dans le tableau 4. Il est à noter que l'on aurait pu prendre 1 comme borne inférieure de  $B_3$  plutôt que 0,99 : cela permet seulement de vérifier la condition  $B_3 \subset A_3 \subset ]0, 1[$  du premier chapitre, mais c'est en fait sans importance puisqu'en pratique Scilab simule des tirages selon  $\mathcal{U}([0, 1])$ .

Indice	Paramètre	Borne inférieure	Borne supérieure
1	$a_1$	0,01	1
2	$a_2$	0,01	1
3	$a_3$	0,01	0,99
4	$a_4$	0,01	2
5	$a_5$	0,01	0,2
6	$ILL_\infty$	$0,01 \times 10^5$	$2 \times 10^5$
7	$TI_\infty$	$0,01 \times 10^7$	$2 \times 10^7$
8	$V_2$	0,01	100
9	$r_1$	0,01	2
10	$r_2$	0,01	2
11	$K_1^0$	$0,1 \times 10^9$	$7 \times 10^9$
12	$K_2$	$0,1 \times 10^9$	$7 \times 10^9$

TABEAU 4 – Bornes des intervalles de variation de chaque paramètre pour l'analyse de sensibilité

## 3.4 Estimation des indices de Sobol

### 3.4.1 Estimation par Monte Carlo

Il s'agit maintenant d'estimer numériquement ces indices, c'est-à-dire estimer les intégrales  $\int_{\mathbb{R}^s} (f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}))^2 P_{X_{i_1}}(dx_{i_1}) \cdots P_{X_{i_s}}(dx_{i_s})$ . Nous utilisons pour cela la méthode dite de Monte Carlo.

Soit  $g \in L^1(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p), P_X)$ , on cherche à estimer l'intégrale

$$I = \mathbb{E}(g(X)) = \int g(x) P_X(dx).$$

On considère  $(\tilde{X}_k)_{k \in \mathbb{N}}$  une suite de v.a.r. indépendantes et identiquement distribuées (i.i.d), de même loi que  $X$ . La méthode de Monte Carlo consiste à utiliser la convergence presque-sûre de  $I_N = \frac{1}{N} \sum_{k=1}^N g(\tilde{X}_k)$  vers  $I$ , qui est une conséquence directe de la loi forte des grands nombres. Ainsi, si  $\tilde{x}_1, \dots, \tilde{x}_N$  est une réalisation de  $\tilde{X}_1, \dots, \tilde{X}_N$ , on obtient une approximation de  $I$  par

$$\tilde{\mathbb{E}}(g(X)) = \frac{1}{N} \sum_{k=1}^N g(\tilde{x}_k)$$

avec une convergence en  $\mathcal{O}(N^{-\frac{1}{2}})$ . Cette méthode permet, moyennant quelques astuces de calcul, d'obtenir une estimation des indices de sensibilité de tout ordre, ainsi que des indices de sensibilité totaux.

Dans la pratique, pour un modèle comportant une dizaine de paramètres, une taille d'échantillon de  $N = 10000$  permet d'avoir une estimation correcte des indices de sensibilité. Il est possible de quantifier l'erreur d'estimation commise grâce au théorème central limite, mais nous n'abordons pas cet aspect ici.

### 3.4.2 Résultats

On considère les indices de sensibilité totaux pour quatre instants où le modèle est fortement sensible aux variations des paramètres :  $t = 1, 2, 5, 10$  jours. On ne s'intéresse ici qu'aux variables  $Y_3$  et  $Y_4$ , car  $Y_1$  et  $Y_2$  ne dépendent clairement pas de certains paramètres (cas  $D = 0$ ). On applique donc la théorie de Sobol aux première et deuxième composantes de  $\theta_t$  dans le cas  $D = 10$ . Pour chaque valeur de  $t$ , on a ainsi 24 indices :

- $S(t, N_1, x_1), \dots, S(t, N_1, x_{12})$  où  $S(t, N_1, x_i) = S_{T_i}$  (notations précédentes) ;
- $S(t, N_2, x_1), \dots, S(t, N_2, x_{12})$  de la même façon pour la deuxième composante de  $\theta_t$ .

Nous avons donc au total  $4 \times 24 = 96$  indices à estimer. On prend comme taille d'échantillon  $N = 10000$  : les résultats sont répertoriés dans le tableau 5. Pour avoir une idée plus générale de la sensibilité de  $N_1$  et  $N_2$  sur la totalité des instants considérés, on considère pour chaque paramètre  $x_i$  la somme des 8 indices associés :

$$\tilde{S}_i = \sum_{k \in \{1,2\}} \sum_{t \in \{1,2,5,10\}} S(t, N_k, x_i).$$

et une représentation graphique est donnée sur la figure 13. On décide de négliger l'influence des paramètres  $x_i$  tels que  $\tilde{S}_i < 0,15$  :  $a_3$ ,  $a_4$  et  $V_2$ . On les fixe à une valeur cohérente (cf. tableau 6) : ainsi, on ne considère plus que 9 paramètres dans notre modèle.

**Remarque 3.** *Le choix et le nombre des instants d'étude est complètement arbitraire. Pour affiner l'analyse de sensibilité, on pourrait tout à fait considérer plus d'instant, cela soulevant seulement une question de temps de calcul.*

$t$	$N_1$				$N_2$			
	1	2	5	10	1	2	5	10
$a_1$	0,021	0,027	0,044	0,071	0,005	0,050	0,070	0,096
$a_2$	0,124	0,141	0,118	0,099	0,018	0,110	0,120	0,173
$a_3$	<b>0,000</b>	<b>0,000</b>	<b>0,001</b>	<b>0,010</b>	<b>0,000</b>	<b>0,009</b>	<b>0,015</b>	<b>0,032</b>
$a_4$	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,005</b>	<b>0,000</b>	<b>0,003</b>	<b>0,035</b>	<b>0,078</b>
$a_5$	0,031	0,013	0,147	0,429	0,852	0,291	0,054	0,041
$ILL_\infty$	0,031	0,043	0,075	0,098	0,003	0,060	0,096	0,143
$TI_\infty$	0,307	0,234	0,172	0,187	0,014	0,133	0,238	0,315
$V_2$	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>	<b>0,000</b>
$r_1$	0,568	0,737	0,748	0,517	0,059	0,311	0,373	0,414
$r_2$	0,000	0,000	0,002	0,020	0,082	0,367	0,293	0,205
$K_1^0$	0,002	0,006	0,219	0,238	0,000	0,005	0,020	0,029
$K_2$	0,000	0,000	0,001	0,006	0,000	0,026	0,244	0,313

TABLEAU 5 – Estimation des indices de Sobol totaux pour chaque paramètre à  $t = 1, 2, 5, 10$  : les lignes en gras ont la somme de leurs termes  $\tilde{S}_i$  inférieure à 0,15.

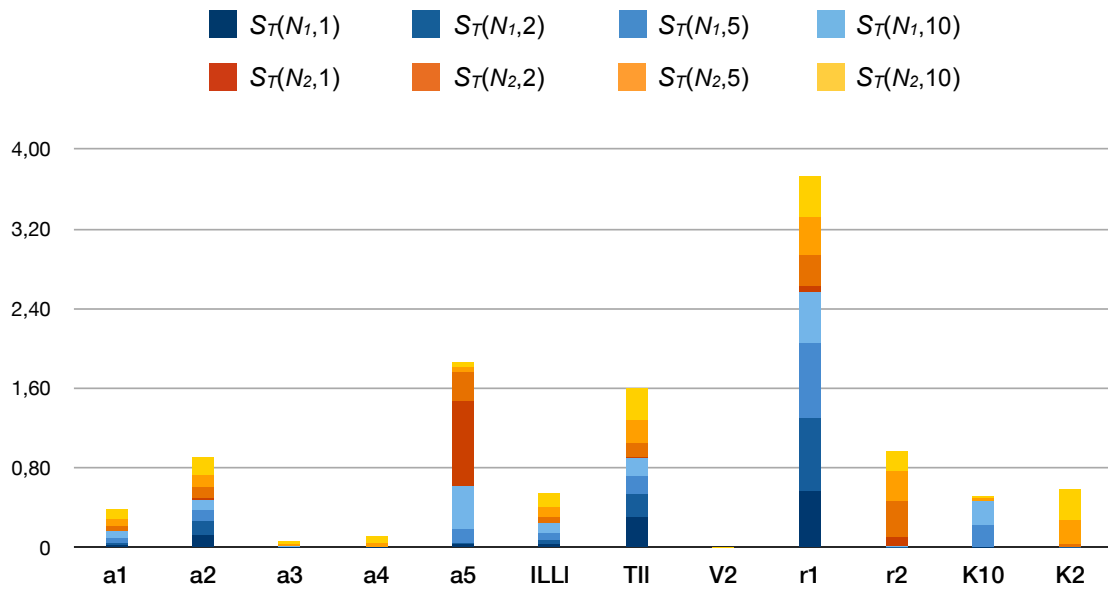


FIGURE 13 – Représentation des sommes des indices de Sobol totaux pour chaque paramètre : les variations de bleu (resp. orange) correspondent aux indices associés à  $N_1$  (resp.  $N_2$ ) pour chaque instant  $t = 1, 2, 5, 10$ .

Indice	Paramètre	Valeur fixée
3	$a_3$	0,05
4	$a_4$	1
8	$V_2$	1

TABLEAU 6 – Valeurs des paramètres considérés comme peu influents

## Chapitre 4

# Estimation des paramètres

### 4.1 Estimation ou optimisation ?

Bien que le titre de ce rapport comporte le mot *optimisation*, nous parlerons également d'*estimation* des paramètres. En fait cela semble être surtout une affaire de point de vue : l'optimisation fait référence au fait de déterminer des valeurs *idéales* des paramètres selon un certain critère : elle est principalement liée à une approche déterministe (cf. section 4.3). L'estimation correspond plus à une approche probabiliste : on estime la probabilité qu'un paramètre prenne telle ou telle valeur au vu des données récoltées (cf. section 4.4).

Ce stage nous a conduit à privilégier la seconde approche pour deux raisons : d'une part elle semble plus adaptée à notre problème (qui est non convexe), d'autre part elle semble donner des résultats plus fins. Elle fournit en effet un moyen de connaître les *densités* de probabilité de chaque paramètre, et pas seulement les valeurs les plus probables : à l'avenir cela peut certainement permettre une analyse plus fine de la qualité de notre modèle.

### 4.2 Notations en vue de l'utilisation des données

Nous avons à présent un modèle du système algues-daphnies avec  $p = 9$  paramètres. Les données dont nous disposons concernent à la fois le système sans daphnies, modélisé par  $(E_1)$ , et le système avec daphnies, modélisé par  $(E_2)$ . Afin prendre en compte l'ensemble de ces données, nous devons introduire certaines notations.

Soient  $x = (x_1, \dots, x_p)$  le vecteur des paramètres et  $\mathcal{D} = A_1 \times \dots \times A_p$ . Désormais, la fonction du modèle est  $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)$  où  $(\phi_1(t, x), \phi_2(t, x)) = \phi(t, x)$  pour  $D = 0$  et  $(\phi_3(t, x), \phi_4(t, x)) = \phi(t, x)$  pour  $D = 10$ . En d'autres termes :

- $\phi_1$  est le nombre d’algues **en suspension** dans le système **sans daphnies**,
- $\phi_2$  est le nombre d’algues **en sédimentées** dans le système **sans daphnies**,
- $\phi_3$  est le nombre d’algues **en suspension** dans le système **avec daphnies**,
- $\phi_4$  est le nombre d’algues **en sédimentées** dans le système **avec daphnies**.

Le problème est que l’on ne dispose pas du même nombre de données pour ces 4 quantités, ni pour les 3 essais. On doit donc considérer  $3 \times 4 = 12$  tableaux de données, que l’on définit sous forme de matrices : la matrice  $M_{ij} \in \mathcal{M}_{2, n_{ij}}(\mathbb{R})$  représente les  $n_{ij}$  données disponibles sur la quantité  $j \in \{1, 2, 3, 4\}$  pour l’essai  $i \in \{1, 2, 3\}$ . Ainsi, on note

$$M_{ij} = \begin{pmatrix} t_1^{ij} & y_1^{ij} \\ \vdots & \vdots \\ t_{n_{ij}}^{ij} & y_{n_{ij}}^{ij} \end{pmatrix}$$

où  $y_k^{ij}$  est la mesure au temps  $t_k^{ij}$  de la quantité  $j$  lors de l’essai  $i$ .

Par exemple, les données de l’essai 3 sur les algues en suspension dans le système sans daphnies sont représentées par  $M_{3,1}$ , avec  $n_{3,1} = 9$ .

### 4.3 Approche déterministe : méthode des moindres carrés

Dans un premier temps, on définit la notion d’ajustement de façon empirique : une courbe est dite *ajustée* si elle minimise la somme des carrés des différences entre les mesures et les prédictions de  $\Phi(t, x)$ . Cette méthode, appelée **ajustement par la méthode des moindres carrés**, correspond à une façon intuitive (car graphique) de rechercher la courbe la plus “proche” des données. Nous verrons qu’elle n’est pas sans fondement mathématique (cf section 4.4.1).

Cette approche se traduit par la définition d’une *fonction coût*, positive et continue sur  $\mathcal{D}$ , que l’on cherche à minimiser. Sachant que les 3 essais n’ont probablement pas été réalisés dans des conditions *strictement* identiques, il semble peu cohérent de tenter d’ajuster le modèle en même temps aux 3 jeux de données correspondants. Nous sommes donc amenés à définir une fonction coût pour chaque essai, notée  $h_i$ , et définie par :

$$\forall x \in \mathcal{D}, h_i(x) = \sum_{j=1}^4 \sum_{k=1}^{n_{ij}} \left( \phi_j(t_k^{ij}, x) - y_k^{ij} \right)^2.$$

Ces fonctions sont continues d’après la propriété 2, et elles admettent un minimum global sur  $\bar{\mathcal{D}}$  : en effet, soit  $\Phi$  n’est pas continue sur  $\bar{\mathcal{D}}$  et alors  $h_i$  tend vers  $+\infty$  quand un des  $x_i$  tend vers les bornes de  $A_i$  (on dit que  $h_i$  est *coercive*), soit  $\Phi$  est continue sur  $\bar{\mathcal{D}}$  et le résultat est le même (si  $\bar{\mathcal{D}}$  est borné on applique la continuité sur un compact, sinon on conclut par l’argument

de coercivité). En revanche, il se peut très bien que les fonctions  $h_i$  présentent plusieurs minima locaux (cf. 4.3.2).

### 4.3.1 Algorithmes de minimisation

Soit  $h = h_i$  où  $i$  est fixé, on s'intéresse maintenant à la recherche des minima de  $h$ , globaux ou locaux. Comme on l'a vu précédemment, on ne connaît pas la forme analytique de la fonction  $\Phi$  : il en est donc de même pour  $h$ . En revanche, on est capable de résoudre numériquement  $(E_1)$  et  $(E_2)$  avec des algorithmes informatiques (méthode de Runge-Kutta par exemple). Si l'on suppose  $h$  de classe  $\mathcal{C}^1$ , on peut estimer ses dérivées partielles.

On note  $\nabla h = \mathbf{grad} h$ , et pour  $\alpha \in \mathcal{D}$ ,  $L_h(\alpha) = \{x \in \mathcal{D} \mid h(x) \leq \alpha\}$ . On utilise une *méthode de descente* : l'idée générale est de construire, à partir d'un point  $x^{(0)} \in \mathcal{D}$  donné, une suite  $(x^{(k)})_k$  qui vérifie :

$$\begin{aligned} (i) \quad & \forall k \geq 1, x^{(k)} \in L_h(h(x^{(0)})); \\ (ii) \quad & \nabla h(x^{(k)}) \xrightarrow[k \rightarrow +\infty]{} 0; \\ (iii) \quad & \|x^{(k+1)} - x^{(k)}\| \xrightarrow[k \rightarrow +\infty]{} 0. \end{aligned}$$

Si  $L_h(h(x^{(0)}))$  est compact, on a existence d'une sous-suite de  $(x^{(k)})_k$  qui converge vers un point  $\bar{x}$  tel que  $\nabla h(\bar{x}) = 0$ . La forme générale d'un algorithme de descente est la suivante [13] :

*choisir  $x^{(0)}$ , poser  $k = 0$*

***tant que***  $\|\nabla h(x^{(k)})\| > \text{seuil de tolérance}$

*déterminer une direction  $d^{(k)}$  telle que  $\exists \sigma > 0$  tel que  $h(x^{(k)} + \sigma d^{(k)}) < h(x^{(k)})$*

*déterminer un pas convenable  $\sigma_k > 0$*

*poser  $x^{(k+1)} = x^{(k)} + \sigma d^{(k)}$  et  $k = k + 1$*

Ainsi, en fixant un seuil de tolérance approprié, l'algorithme converge vers une valeur  $\bar{x}$  qui est arbitrairement proche d'un **minimum local** de  $h$ . On parle de méthode déterministe car si l'on fixe  $x^{(0)}$  et un nombre d'itérations  $k$ , l'algorithme renverra toujours vers la même valeur.

Il existe plusieurs algorithmes effectuant une telle descente : on peut citer par exemple les méthodes de descente du gradient, de relaxation, du gradient conjugué et de Newton. Nous utilisons ici l'algorithme de Levenberg-Marquardt, qui est spécifiquement conçu pour la méthode des moindres carrés, et qui est implémenté dans Scilab. La figure 14 montre le résultat avec deux valeurs différentes de  $x^{(0)}$ .

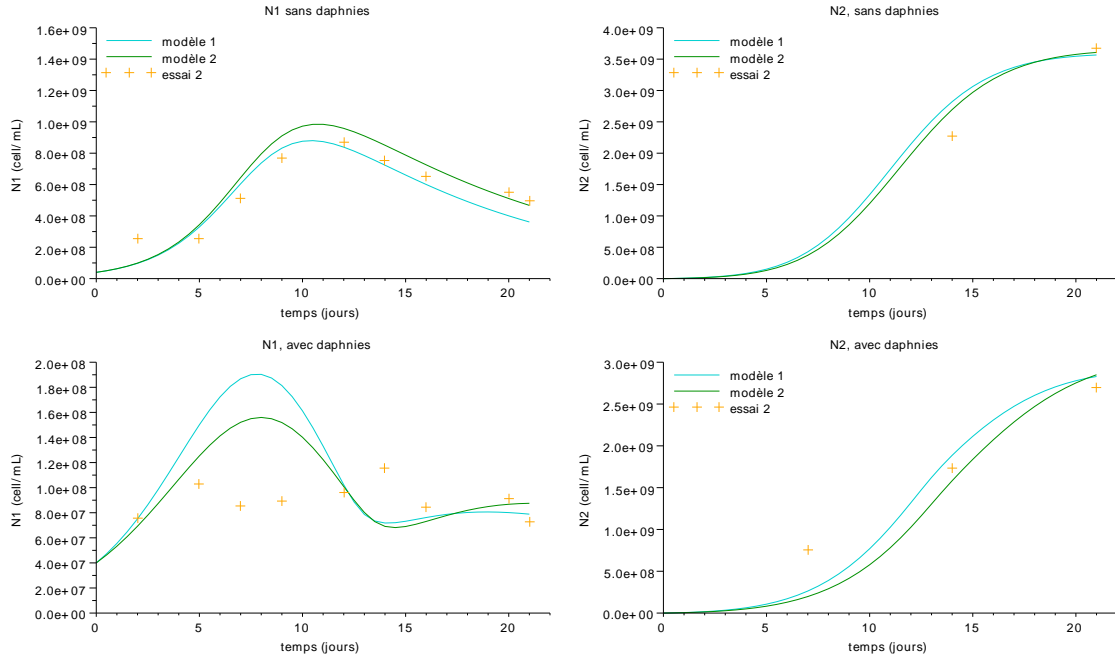


FIGURE 14 – Courbes ajustées par méthode de descente : “modèle 1” et “modèle 2” sont les courbes associées à deux valeurs  $\bar{x}_1$  et  $\bar{x}_2$  pour lesquelles  $h_2$  atteint un minimum local.

### 4.3.2 Problème fondamental : minima locaux

La figure 14 illustre un problème fondamental : la présence de minima locaux. En effet, malgré la continuité de la fonction coût qui assure l’existence d’un minimum global, on n’a *a priori* aucune connaissance sur l’existence ou non de minima locaux. En fait, la fonction coût n’est pas convexe (cf. fig 15), et on peut montrer numériquement l’existence d’un grand nombre de minima locaux. Il s’agit du principal point faible de la méthode de descente : la valeur  $\bar{x}$  retournée dépend fortement du vecteur initial  $x^{(0)}$ . Ainsi, à moins de connaître assez précisément la valeur idéale de  $x$ , l’algorithme ne retourne pas le minimum global. En outre, l’approche par les moindres carrés associés à une méthode de descente possède d’autres inconvénients :

- La méthode suppose que  $\Phi$  est  $\mathcal{C}^1$  sur  $\mathbb{R}_+ \times \mathcal{D}$ , or on ne sait pas le prouver dans notre cas.
- Elle renvoie une seule valeur pour chaque paramètre, ce qui suppose implicitement que le système algues-daphnies peut-être décrit avec uniquement ces valeurs. On omet alors la notion fondamentale de **variabilité**, c’est-à-dire l’aspect aléatoire du système réel.
- L’**ajustement** est pour le moment défini de façon assez empirique. Or, le fait d’avoir à ajuster le modèle suggère que l’on a de l’*incertitude* sur les paramètres, notamment à cause d’erreurs de mesure : il serait judicieux de chercher à traiter ces erreurs de mesure autrement que par un ajustement “graphique”.



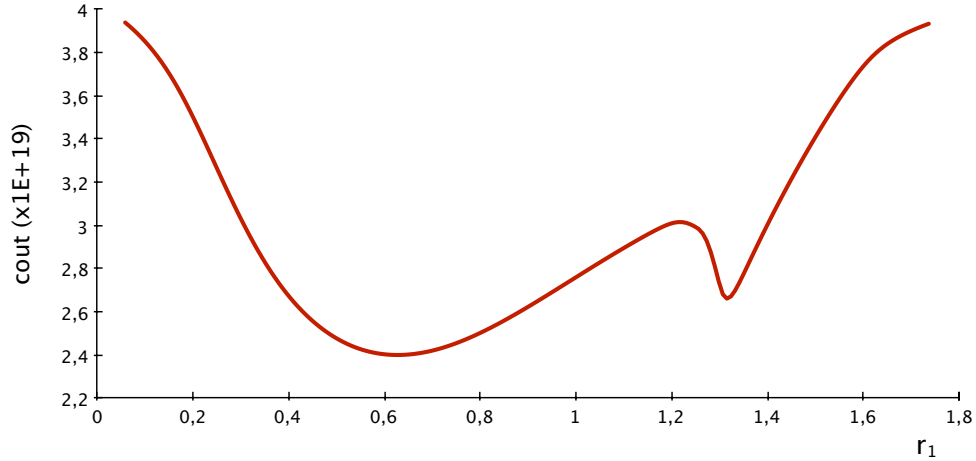


FIGURE 15 – Projection de  $h_1$  sur l’axe  $r_1$ , les autres paramètres étant fixés

Cette approche semble donc inappropriée à notre problème. Outre la présence de minima locaux qui la rend difficile à appliquer, il semble plus cohérent de rechercher, plutôt qu’une valeur unique, une densité de probabilité sur chaque paramètre.

#### 4.4 Approche stochastique : méthodes de Monte Carlo

On qualifie de Monte Carlo les méthodes de résolution numérique qui font intervenir, contrairement aux méthodes déterministes, des procédés aléatoires (ou *stochastiques*). C’est ce type de méthode qui permet de faire de l’*inférence statistique* sur les paramètres, c’est-à-dire obtenir une information sur ceux-ci, à partir de données et d’éventuelles connaissances statistiques sur les erreurs de mesure. La nature de cette information diffère selon deux approches historiques (ou *cadres d’inférence statistique*) :

- Le cadre **fréquentiste**, dans lequel on assimile les probabilités à des fréquences d’observation à long terme (beaucoup d’observations), à paramètres fixés. L’inférence statistique consiste alors à fournir les valeurs des paramètres *les plus probables*, selon une loi de probabilité à définir (cf. section 4.4.1).
- Le cadre **bayésien**, où l’on considère des probabilités à la fois sur les observations et les paramètres. On détermine alors la distribution *a posteriori* des paramètres, c’est-à-dire la loi conditionnelle des paramètres sachant les observations (*i.e.* les données), en utilisant le théorème de Bayes (cf. section 4.4.2).

#### 4.4.1 Lien entre l'approche fréquentiste et la méthode des moindres carrés

Si l'on considère que l'erreur de mesure est gaussienne, centrée et de variance constante  $\sigma^2$ , alors l'inférence au sens fréquentiste est équivalente à l'optimisation par la méthode des moindres carrés. En effet, on cherche dans ce cadre à maximiser une fonction appelée *vraisemblance*, qui est en fait la probabilité d'obtenir une certaine mesure de la valeur de sortie (erronée) lorsque l'on connaît les paramètres (donc la vraie valeur de sortie). Soit  $i \in \{1, 2, 3, 4\}$  fixé, considérons uniquement  $\phi_i$  pour fixer les idées. Supposons que l'on dispose de  $m$  mesures  $y_1, \dots, y_m$  aux instants  $t_1, \dots, t_m$ , telles que pour tout  $j \in \{1, \dots, m\}$ , on ait  $y_j = \phi_i(x, t_j) + \epsilon_j$  où les  $\phi_i(x, t_j)$  sont les valeurs exactes du modèle et les  $\epsilon_j$  sont des erreurs aléatoires i.i.d de densité  $g_\sigma$  et plus précisément de loi  $\mathcal{N}(0, \sigma^2)$ . Le vecteur des données  $y = (y_1, \dots, y_m)$  est donc vu un  $m$ -échantillon i.i.d. d'une v.a.r. à densité notée  $Y$ , et on définit la vraisemblance  $L$  par :

$$L(y, x, \sigma) = \prod_{j=1}^m g_\sigma(\epsilon_j) = \frac{1}{(2\pi\sigma^2)^{\frac{m}{2}}} \exp\left(-\frac{1}{2} \sum_{j=1}^m \frac{(y_j - \phi_i(x, t_j))^2}{\sigma^2}\right).$$

On remarque que lorsque  $x$  et  $\sigma$  sont fixés, la densité de  $Y$  est  $f_{x,\sigma}(y) = L(y, x, \sigma)$ . Puisqu'en fait  $y$  est connu, on cherche à obtenir le *maximum de vraisemblance* de  $y$ , c'est-à-dire trouver la (ou les) valeur(s) de  $x$  qui maximise(nt)  $f_{x,\sigma}(y)$ . Comme  $\sigma$  est fixé, cela revient à minimiser la quantité  $\sum_{j=1}^m (y_j - \phi_i(x, t_j))^2$  : on retrouve donc la méthode des moindres carrés (et aussi le problème des minima locaux).

#### 4.4.2 L'approche bayésienne

Bien que l'approche fréquentiste considère  $y$  comme une variable aléatoire, elle fournit seulement une ou plusieurs valeurs optimales de  $x$ , et non une distribution sur  $x$ . Pour obtenir une telle distribution, il faut assimiler à la fois  $y$  et  $x$  à des variables aléatoires notées  $Y$  et  $X$ .

On se place désormais dans le cas où  $Y = (Y^{ij})_{i \in \{1,2,3\}, j \in \{1,2,3,4\}}$  où les  $Y^{ij}$  sont des variables aléatoires à valeurs dans  $\mathbb{R}^{n^{ij}}$  (données relatives à  $\phi_1, \phi_2, \phi_3, \phi_4$  pour les essais 1, 2, 3), **on suppose que  $(X, Y)$  admet une densité** notée  $f_{(X,Y)}(x, y)$  et on considère les densités conditionnelles associées (cf. annexe B). Ainsi, à  $\sigma$  fixé,  $L(y, x, \sigma) = f_{x,\sigma}(y)$  peut-être vue comme la densité conditionnelle de  $Y$  sachant  $X = x$ , notée  $f_{Y|X=x}(y)$ . Enfin, on suppose que les variables  $Y^1 := (Y^{1j})_{j \in \{1,2,3,4\}}$ ,  $Y^2 := (Y^{2j})_{j \in \{1,2,3,4\}}$  et  $Y^3 := (Y^{3j})_{j \in \{1,2,3,4\}}$  sont indépendantes (indépendance des trois essais).

Etant donné qu'on a ici peu d'information sur l'erreur de mesure, on décide de considérer  $\sigma$  comme un autre paramètre aléatoire, et on note  $\sigma^2 = x_{p+1}$ . On se ramène donc à un modèle à

10 paramètres, et dans la suite on note  $p = 10$ .

**Remarque 4.** *La vraisemblance  $L$  est dans notre cas définie à partir d'un modèle d'erreur gaussien, mais on pourrait la définir à partir de n'importe quelle erreur aléatoire (avec ou sans paramètres du type  $\sigma$ ) en écrivant  $L(y, x) = f_{Y|X=x}(y)$ .*

On suppose connue la loi de  $X$  (de densité  $f_X$ ), appelée “probabilité *a priori* des paramètres”. La densité que l'on cherche à calculer est  $f_{X|Y=y}$  : c'est la densité conditionnelle de  $X$  sachant  $Y = y$ , ou encore la “probabilité des paramètres sachant les données  $y$ ”, définie par

$$f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}.$$

Or, on ne connaît ni  $f_{(X,Y)}$ , ni  $f_Y$ . Pour calculer  $f_{X|Y=y}$ , l'idée consiste à appliquer le théorème de Bayes pour utiliser ce que l'on connaît. On a alors

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y) f_X(x)}{f_Y(y)} = \frac{f_{Y|X=x}(y) f_X(x)}{\int_{\mathbb{R}^p} f_{(X,Y)}(\tilde{x}, y) d\tilde{x}},$$

ce qui peut s'écrire

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y) f_X(x)}{\int_{\mathbb{R}^p} f_{Y|X=\tilde{x}}(y) f_X(\tilde{x}) d\tilde{x}}. \quad (4.1)$$

On remarque que le dénominateur du terme de droite dans (4.1) n'est rien d'autre qu'une constante de normalisation. Bien que l'on connaisse  $f_{Y|X=x}$  et  $f_X$ , le calcul analytique de l'intégrale est souvent impossible, et une estimation numérique semble ardue, surtout en grande dimension comme dans notre cas ( $p = 10$ ). La force de certains algorithmes de Monte Carlo est de ne pas avoir besoin du calcul de cette constante : c'est le cas des deux algorithmes que nous allons détailler dans ce rapport : l'algorithme de Metropolis-Hastings (cf. 4.5) et une méthode de filtrage particulière (cf 4.6).

**Remarque 5.** *La définition de la densité a priori  $f_X$  est fondamentale : elle reflète l'information que l'on a avant les mesures. Par exemple, on peut décider que les paramètres ne peuvent pas dépasser certaines bornes sans être aberrants. Lorsque l'on a peu d'information, on cherche la loi la moins “informative” possible : on est alors tenté de prendre une loi uniforme, ce qui permet par exemple de résoudre le problème le maximum de vraisemblance vu plus haut, en remarquant que dans ce cas  $\max_x f_{Y|X=x}(y) = \max_x f_{X|Y=y}(x)$ . Cependant, la “non informativité” de ce type de loi n'est pas claire et le choix de  $f_X$  en général est controversé.*

**Remarque 6.** *Si certains algorithmes requièrent la connaissance de la forme analytique de la fonction  $f_{Y|X=x}$ , d'autres nécessitent seulement de savoir la simuler, c'est-à-dire de pouvoir tirer des échantillons  $y$  selon cette densité.*

### 4.4.3 Principe des méthodes de simulation de Monte Carlo

D'après la section précédente, on a  $f_{X|Y=y}(x) \propto f_{Y|X=x}(y)f_X(x)$ . On note  $\pi$  la loi associée à  $f_{X|Y=y}$ , et on considère  $Z$  une v.a. de loi  $\pi$ . Au lieu de calculer  $f_{X|Y=y}$ , les algorithmes de Monte Carlo dits "de simulation" renvoient un échantillon ayant les mêmes propriétés vis-à-vis de la loi des grands nombres qu'un échantillon tiré selon  $\pi$ . En d'autres termes, pour  $N \in \mathbb{N}$ , on obtient un  $N$ -échantillon tiré selon une loi  $Z_N$  telle que  $(Z_N)_{N \in \mathbb{N}}$  converge en loi vers  $\pi$ . Ainsi, pour  $N$  assez grand on peut considérer que ces algorithmes renvoient un  $N$ -échantillon approximativement tiré selon  $\pi$ . On peut ensuite estimer  $f_{X|Y=y}$ , par exemple avec la méthode d'estimation par noyau.

Le principe général des méthodes de simulation de Monte Carlo est d'effectuer un tel tirage sans avoir recours à la constante  $\int f_{Y|X=x}(y)f_X(x) dx$ . Cependant, le principe qui assure la bonne propriété de l'échantillon diffère selon les algorithmes : nous présentons respectivement dans les sections 4.5 et 4.6 l'algorithme de Metropolis-Hastings, basé sur une chaîne de Markov, et un algorithme de filtrage particulière, basé sur l'échantillonnage pondéré.

#### Application au système algues-daphnies

Puisque le but est de fournir des distributions de probabilité, on va considérer globalement les 3 essais : si certains paramètres diffèrent selon les essais, cette différence sera prise en compte dans la distribution *a posteriori* des paramètres.

**Remarque 7.** Grâce à l'hypothèse d'indépendance des paramètres  $X_i$ , on a :

$$f_{X|Y=y} = \prod_{i=1}^p f_{X_i|Y=y}$$

et donc pour avoir la forme de la fonction  $f_{X|Y=y}$  dans  $\mathbb{R}^p$  il suffit de multiplier chaque  $f_{X_i|Y=y}$ .

## 4.5 L'algorithme de Metropolis-Hastings

Le principe est de simuler la loi de probabilité  $\pi$  de densité  $f$  à partir d'une chaîne de Markov  $(X_t)_{t \in \mathbb{N}}$  ergodique de loi invariante  $\pi$  : la propriété d'ergodicité implique que pour un nombre d'itérations suffisamment grand, on obtient un échantillon approximativement tiré selon  $\pi$ .

La chaîne de Markov  $(X_t)_{t \in \mathbb{N}}$  est construite de la façon suivante. On effectue une marche aléatoire sur  $\Omega = \mathbb{R}^{10}$  : on utilise une loi  $Q$  pour choisir, étant en position à l'instant  $t$ , une nouvelle position potentielle (appelée proposition) avec la probabilité  $Q(x, x_t)$ . La probabilité

$\mathbb{P}(X_{t+1} = x|X_t = x_t)$  marche se déplace en  $x$  à l'instant  $t + 1$  est alors définie par

$$\mathbb{P}(X_{t+1} = x|X_t = x_t) = \min\left(\frac{f(x)Q(x_t, x)}{f(x_t)Q(x, x_t)}, 1\right). \quad (4.2)$$

La probabilité restante est la probabilité que la marche reste en  $x_t$ , c'est-à-dire

$$\mathbb{P}(X_{t+1} = x_t|X_t = x_t) = 1 - \mathbb{P}(X_{t+1} = x|X_t = x_t).$$

En pratique, on va prendre  $Q(\cdot, x_t) \sim \mathcal{N}(x_t, \lambda)$  et on a ainsi  $Q(x, x_t) = Q(x_t, x)$ , d'où :

$$\mathbb{P}(X_{t+1} = x|X_t = x_t) = \min\left(\frac{f(x)}{f(x_t)}, 1\right). \quad (4.3)$$

Le principal intérêt de l'algorithme est dans la formule (4.2). En effet on ne connaît  $f(x)$  qu'à une constante près : le quotient  $f(x)/f(x_t)$  permet précisément d'ignorer cette constante.

**Remarque 8.** *La chaîne de Markov ainsi construite est à espace d'états non dénombrable ( $\mathbb{R}^{10}$ ) et les résultats de convergence ne sont pas simples à montrer. Pour plus de précisions sur cet aspect, on renvoie à [14].*

Ainsi, en effectuant  $m$  pas sur cette chaîne avec  $m$  assez grand, on obtient en prenant le  $m$ -ième pas un tirage selon une approximation de  $\pi$ . La figure 16 montre un exemple de tirage effectué grâce à cet algorithme.

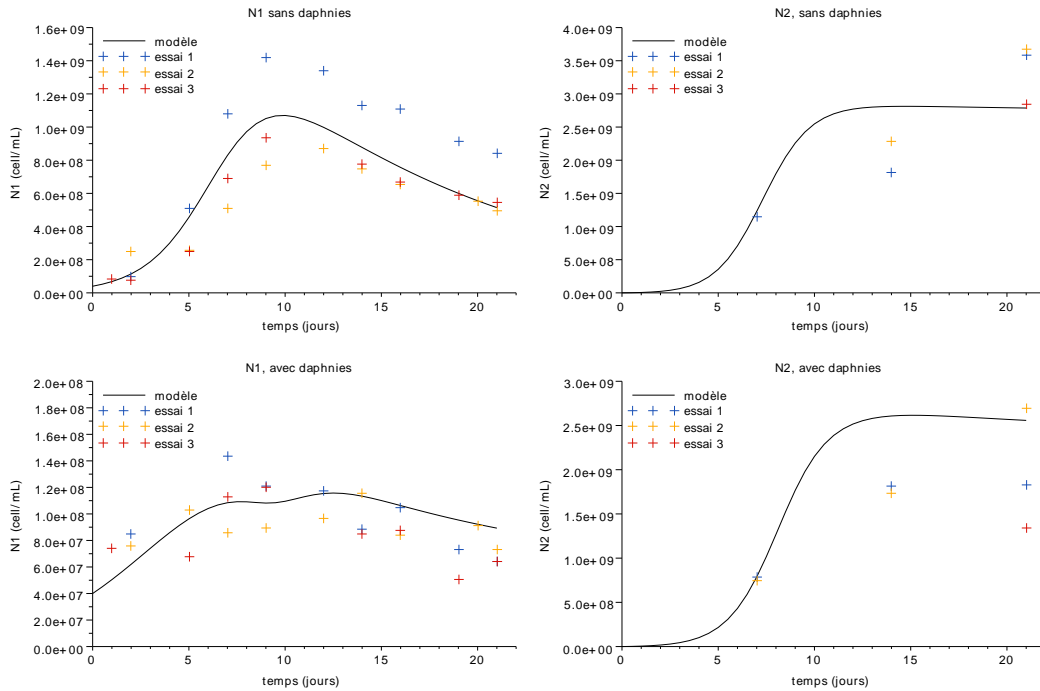


FIGURE 16 – Courbes correspondant à un tirage aléatoire selon une approximation de  $\pi$

## 4.6 Un algorithme de filtrage particulaire

On pourrait utiliser l'algorithme précédent pour obtenir un  $N$ -échantillon approximativement distribué selon  $\pi$ . Cependant cela pose un problème d'efficacité :

- soit on effectue  $N$  marches aléatoires de longueur  $m$  sur la chaîne de Markov (avec  $m$  assez grand), ce qui est extrêmement coûteux en temps ;
- soit on lance une seule chaîne, et à partir du rang  $m$  on enregistre l'état visité à chaque pas pendant  $N$  pas. On obtient ainsi un  $N$ -échantillon, mais les tirages ne sont pas indépendants puisque chaque état dépend de l'état précédent : on perd donc le résultat théorique de convergence. On peut s'arranger pour avoir une indépendance approximative (par exemple en enregistrant tous les  $\tilde{m}$  pas avec  $\tilde{m}$  bien choisi), mais on est alors obligé de faire tourner la chaîne plus longtemps et donc d'augmenter le temps de calcul.

Nous décrivons ici une méthode de génération d'un tel  $N$ -échantillon par filtrage particulaire (voir [15]) qui utilise le principe de l'échantillonnage avec rééchantillonnage par importance (ou SIR pour *Sampling Importance Resampling*). Plutôt que de considérer une marche aléatoire sur une chaîne de Markov, on travaille directement avec un  $N$ -échantillon appelé *population de  $N$  particules* et tiré initialement selon une densité arbitraire  $f_0$ . Ici, c'est l'interaction entre les particules (sous la forme d'un poids affecté à chaque particule) qui fait la puissance de l'algorithme : plus  $N$  est grand, plus l'approximation est précise.

On le fait ensuite évoluer séquentiellement en utilisant les données par paquets : le  $N$ -échantillon est ainsi tiré approximativement selon  $f_{X|Y^1=y^1}$ , puis selon  $f_{X|Y^1=y^1, Y^2=y^2}$ , etc. Ainsi, les méthodes de ce type sont parfois appelées *Méthodes de Monte Carlo séquentielles*. La convergence en  $N$  de chaque étape SIR est assurée par le résultat suivant.

### 4.6.1 Convergence de l'algorithme SIR

On se place dans le cas  $p = 1$  pour simplifier les notations. Le problème est le suivant : on est capable d'échantillonner (*i.e.* tirer des valeurs du paramètre au hasard) selon une certaine loi de densité  $f_0$  dont le support est  $\mathbb{R}$ , et on voudrait obtenir, à partir d'un  $N$ -échantillon selon  $f_0$ , un  $N$ -échantillon approximativement distribué selon une loi cible  $\pi$  de densité  $f$  que l'on ne connaît qu'à une constante multiplicative près .

Soit  $X$  une v.a.r de densité  $f_0$ , et soit  $(X_i)_{i \in \mathbb{N}^*}$  une suite de variables i.i.d. de même loi que  $X$ . On considère la suite  $(x_i)_{i \in \mathbb{N}^*}$  où pour tout  $i \in \mathbb{N}^*$ ,  $x_i$  est un échantillon tiré selon  $X_i$ .

Pour  $i \in \mathbb{N}^*$  et  $N \in \mathbb{N}^*$  on définit le *poids d'ordre  $N$*  de  $x_i$  par

$$w_N(x_i) = \frac{\frac{f(x_i)}{f_0(x_i)}}{\sum_{j=1}^N \frac{f(x_j)}{f_0(x_j)}}.$$

Notons que cette quantité est bien définie car  $f_0$  ne s'annule pas sur  $\mathbb{R}$  et  $f(x_j) \neq 0 \forall j \in \mathbb{N}^*$ . De plus, il apparaît que le calcul de  $w_N$  ne nécessite pas la connaissance la constante multiplicative de  $f$ . Soit  $(Z_N)_{N \in \mathbb{N}^*}$  une suite de v.a. discrètes de loi respectives

$$P_{Z_N} = \sum_{i=1}^N \delta_{x_i} w_N(x_i).$$

L'algorithme SIR est justifié par le théorème suivant :

**Théorème 2.** *On suppose que pour toute fonction  $g$  définie sur  $\mathbb{R}$ , de classe  $\mathcal{C}^0$ , bornée,  $\int_{\mathbb{R}} g(x)f(x) dx < +\infty$ . Alors la suite  $(Z_N)_{N \in \mathbb{N}^*}$  converge en loi vers  $\pi$ .*

*Démonstration.* Soit  $\widehat{X}$  une v.a.r. de loi  $\pi$ , c'est-à-dire de densité  $f$ . Soit  $g$  une fonction définie sur  $\mathbb{R}$ , continue et bornée, on pose  $h(x) = \frac{g(x)f(x)}{f_0(x)}$ . On a par la loi forte des grands nombres :

$$\frac{1}{N} \sum_{i=1}^N \frac{g(x_i)f(x_i)}{f_0(x_i)} \xrightarrow{N \rightarrow +\infty} \mathbb{E}(h(X)) \quad (4.4)$$

et

$$\frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{f_0(x_i)} \xrightarrow{N \rightarrow +\infty} \int_{\mathbb{R}} \frac{f(x)}{f_0(x)} f_0(x) dx = 1. \quad (4.5)$$

Or on constate que

$$\mathbb{E}(h(X)) = \int_{\mathbb{R}} \frac{g(x)f(x)}{f_0(x)} f_0(x) dx = \int_{\mathbb{R}} g(x)f(x) dx = \mathbb{E}(g(\widehat{X}))$$

et

$$\mathbb{E}(g(Z_N)) = \sum_{i=1}^N w_N(x_i)g(x_i) = \frac{\sum_{i=1}^N \frac{g(x_i)f(x_i)}{f_0(x_i)}}{\sum_{j=1}^N \frac{f(x_j)}{f_0(x_j)}}.$$

Ainsi, en divisant (4.4) par (4.5), on obtient

$$\mathbb{E}(g(Z_N)) \xrightarrow{N \rightarrow +\infty} \mathbb{E}(g(\widehat{X}))$$

et il s'ensuit que la suite  $(Z_N)$  converge en loi vers  $\widehat{X}$ . □

Ainsi, l'échantillonnage selon  $f_0$  pondéré par  $w_N$  suivi du rééchantillonnage selon  $Z_N$  permet d'obtenir un échantillon asymptotiquement distribué selon la densité cible  $f$ . En pratique, on s'assure que l'hypothèse d'intégrabilité du théorème est vérifiée en prenant  $f_X$  à support compact, ce qui implique que  $f$  l'est aussi.

## 4.6.2 En pratique : étapes de l'algorithme

On se place à  $N \in \mathbb{N}$  fixé. L'algorithme combine trois étapes, dont les deux premières correspondent à l'algorithme SIR :

- **échantillonnage pondéré** (ou *préférentiel*) : échantillonnage selon  $f_0$  puis génération des poids  $w_N(x_i)$  ;
- **bootstrap multinomial** : tirage de  $N$  particules avec remise dans  $(x_i)_{i \in \{1, \dots, N\}}$  avec les probabilités  $w_N(x_i)$  ;
- **dispersion markovienne** : étant donné la manière dont on tire les  $N$  particules, il est fort probable qu'un même  $x_i$  soit tiré plusieurs fois (*i.e.* plusieurs particules sont confondues). Pour éviter ces doublons, on rééchantillonne chaque particule au moyen d'une chaîne de Markov de loi invariante  $\pi$ , en utilisant par exemple l'algorithme de Metropolis-Hastings.

## 4.6.3 Résultats

On rappelle qu'à présent  $p = 10$ . Le but est de fournir un échantillon approximativement distribué selon la loi  $\pi$  de densité  $f$ , où  $f(x) \stackrel{\text{déf}}{=} f_{X|Y=y}(x) \propto f_{Y|X=x}(y)f_X(x)$ . La valeur de  $y$  est connue (c'est l'ensemble des données disponibles), on a donc besoin de la densité *a priori*  $f_X$  et de la vraisemblance  $x \mapsto f_{Y|X=x}(y)$ . On choisit ici  $f_X$  telle que les variables  $X_1, \dots, X_{10}$  soient indépendantes et de loi uniforme, mais de supports différents :  $f_X$  est donc un produit (normé) d'indicatrices d'intervalles, dont les bornes sont données dans le tableau 7.

La vraisemblance est définie comme dans la section 4.4.1 : étant donné que l'on n'a aucune information sur les erreurs de mesure, on considère une erreur gaussienne centrée de variance  $\sigma^2$ , où  $\sigma^2$  est aussi un paramètre aléatoire correspondant à  $X_{10}$  (mais constant sur toute une expérience). On applique ensuite trois fois la séquence (de 3 étapes) décrite ci-dessus sur une population de  $N = 10000$  particules : une première fois avec les données de l'essai 1, puis avec l'essai 2 et enfin l'essai 3. Les résultats sont représentés graphiquement sur les figures 17 et 18.

Si pour certains paramètres ( $K_2$  par ex.) on observe clairement un mode pour la densité, le résultat est plus flou pour d'autres paramètres ( $a_5$  par ex.). En outre, on ne sait pas comment interpréter la présence de pics très marqués sur les histogrammes : cela correspond-il à un vrai pic de densité ou alors est-ce un artefact lié à l'étape de rééchantillonnage ? On a tendance à pencher pour la deuxième hypothèse : il se pourrait que l'algorithme de dispersion (3<sup>ème</sup> étape) ne soit pas assez efficace et laisse de nombreux doublons. Dans tous les cas, il serait intéressant de lancer l'algorithme de filtrage avec un plus grand nombre de particules.



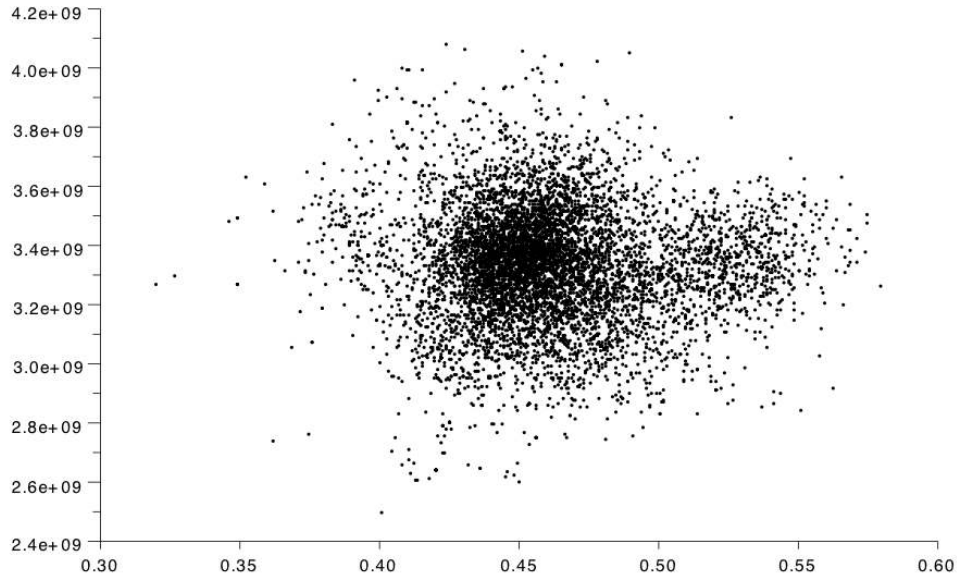


FIGURE 17 – Exemple de projection sur deux axes ( $r_2$  en abscisse,  $K_1^0$  en ordonnée) du nuage de particules final issu de l’algorithme de filtrage particulaire (10000 particules).

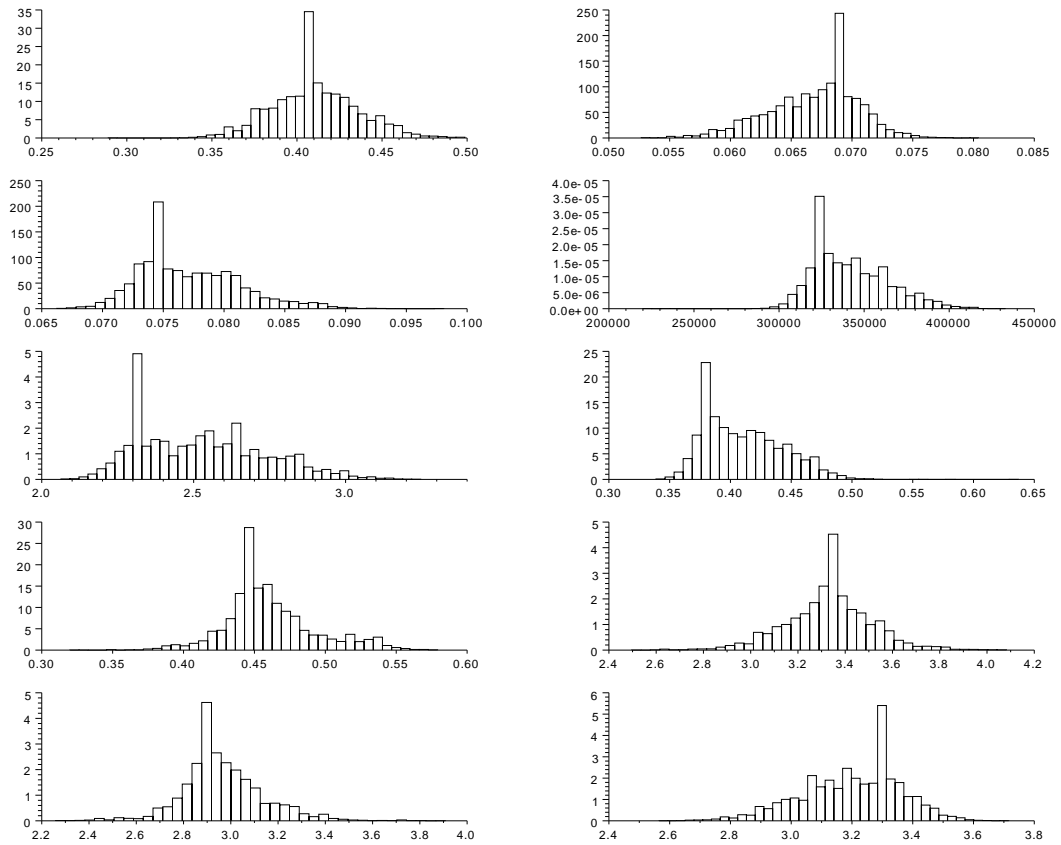


FIGURE 18 – Histogrammes obtenus par l’algorithme de filtrage particulaire (échantillon de taille 10000). Chaque histogramme correspond à un paramètre (de gauche à droite et de haut en bas :  $a_1$ ,  $a_2$ ,  $a_5$ ,  $TI_\infty$ ,  $ILL_\infty$ ,  $r_1$ ,  $r_2$ ,  $K_1^0$ ,  $K_2$ ,  $\sigma^2$ ).

# Chapitre 5

## Conclusion

Dans ce stage, nous avons eu l'occasion d'adapter deux outils mathématiques très généraux à au cas très concret du système algues-daphnies : l'analyse de sensibilité par décomposition de la variance, et l'inférence bayésienne par des méthodes de Monte Carlo.

L'analyse de sensibilité permet de répondre aux questions :

- Le modèle est-il cohérent par rapport au phénomène modélisé ?
- Quelles sont les variables qui contribuent le plus à la variance de la réponse du modèle ?
- Quelles sont les variables les moins influentes ?
- Quelles interactions y a-t-il dans l'influence des variables ?

Quant à l'inférence bayésienne, outre son aspect pratique (fournir par exemple les valeurs des paramètres les plus probables, fournir des marges d'erreur), elle permet de comparer plusieurs modèles entre eux. En effet, bien que difficile à interpréter, la forme de la densité a posteriori des paramètres donne des informations sur la qualité du modèle : on peut imaginer, par exemple, qu'une densité quasiment plate ou avec de nombreux modes implique que le modèle n'est pas assez fidèle au phénomène modélisé.

### Remarque sur l'hypothèse de densité

On a fait l'hypothèse que  $(X, Y)$  admettait une densité, mais ceci n'est pas évident en pratique : il n'est pas certain que si  $X$  admet une densité alors  $Y$  aussi, et encore moins  $(X, Y)$ . En outre, on a fait cette hypothèse de densité pour simplifier les notations, mais il est peut-être possible de généraliser ces résultats à des variables aléatoires quelconques.

## Améliorations possibles

Plusieurs améliorations sont envisageables pour l'étude du système algues-daphnies. Tout d'abord il serait intéressant de comparer les résultats obtenus avec plusieurs formes de densité a priori. Ensuite, il serait utile de réaliser des expériences permettant d'avoir des informations sur l'erreur de mesure, c'est-à-dire l'erreur commise lors de la prise d'échantillon dans le bécher, ainsi que l'erreur de l'appareil de comptage des algues.

De plus, on constate que l'on a ici utilisé le filtrage particulaire au minimum de ses capacités. En effet, on a fourni à l'algorithme la forme analytique de la vraisemblance, mais il serait tout à fait possible de se passer d'une telle expression analytique si l'on savait seulement *simuler* l'erreur de mesure (*i.e.* tirer des échantillons selon la loi d'erreur), par exemple en connaissant les lois des différentes sources d'erreur successives, prises séparément (voir [16]).

Enfin, si des études plus poussées du modèle mathématique défini dans ce rapport révèlent un manque de cohérence trop important, il pourrait être intéressant de se tourner vers un modèle basé sur des équations aux dérivées partielles. En effet, il est clair que le phénomène de sédimentation des algues ne peut se décrire finement qu'avec de telles équations : en considérant un gradient de concentration en algues plutôt que deux compartiments, on pourrait probablement obtenir des résultats plus proches de la réalité. En revanche, un tel modèle nécessiterait des données plus précises, c'est-à-dire sur plusieurs hauteurs de la colonne d'eau par exemple.

## Le mot de la fin

On a pu constater au cours de ce stage que les outils mathématiques employés pour étudier le système algues-daphnies sont en réalité extrêmement généraux : ils interviennent dans de nombreux domaines, de la biologie à la technologie wifi, en passant par les systèmes de visée embarqués dans les hélicoptères. On a ainsi une illustration supplémentaire de la formidable capacité de généralisation liée à l'abstraction mathématique.

# Références

- [1] B. Clément and C. Cadier, “Development of a new laboratory freshwater sediment microcosm test,” *Ecotoxicology*, vol. 7, pp. 279–290, 1998.
- [2] C. Velluet, “Modélisation de la dynamique algale dans un microcosme aquatique de laboratoire avec renouvellement d’eau,” travail de fin d’études, ENTPE, 2009.
- [3] F.-X. Gauci, “Contribution à la modélisation du fonctionnement d’un microcosme de laboratoire,” rapport de master, ENTPE, septembre 2011.
- [4] F. Orias, “Contribution expérimentale à la modélisation d’un microcosme algues-daphnies en conditions normale et contaminée (cadmium),” rapport de master, Université Paris Sud XI / ENTPE, 2012.
- [5] J. J. O’Connor and E. F. Robertson, “The MacTutor History of Mathematics archive,” 1996.
- [6] T. G. Preuss, M. Hammers-Wirtz, U. Hommen, M. N. Rubach, and H. T. Ratte, “Development and validation of an individual based *Daphnia magna* population model : the influence of crowding on population dynamics,” *Ecological modelling*, vol. 220, pp. 310–329, 2009.
- [7] F. Martínez-Jerónimo, “Description of the individual growth of *Daphnia magna* (crustacea : Cladocera) through the von Bertalanffy growth equation. Effect of photoperiod and temperature,” *Limnology*, vol. 13, pp. 65–71, 2012.
- [8] L. Bertalanffy, “Untersuchungen über die gesetzmäßigkeit des wachstums,” *Allgemeine Grundlagen der Theorie ; mathematische und physiologische Gesetzmäßigkeiten des Wachstums bei Wassertieren. Arch. Entwicklungsmech.*, vol. 1, no. 13, pp. 613–652, 1934.
- [9] J. Jacques, “Pratique de l’analyse de sensibilité : comment évaluer l’impact des entrées aléatoires sur la sortie d’un modèle mathématique,” *Pub. IRMA Lille*, vol. 71-III, 2011.
- [10] J. Jacques, *Contributions à l’analyse de sensibilité et à l’analyse discriminante*. PhD thesis, Université Joseph Fourier, 2005.
- [11] I. Sobol’, “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates,” *Mathematics and Computers in Simulation*, vol. 55, pp. 271–280, Février 2001.
- [12] I. Sobol’, “Sensitivity estimates for nonlinear mathematical models,” *MMCE (Engl. Transl.)*, vol. 1, pp. 407–414, 1993.
- [13] G. Koepfler, *Optimisation et algorithmique*. Université Paris Descartes, UFR de Mathématiques et Informatique, 2010.
- [14] J. Bérard, “Chaînes de Markov.” Notes de cours, Université Claude Bernard Lyon 1.
- [15] E. Parent, B. Amzal, and P. Girard, “Investigations particulières pour l’inférence statistique et l’optimisation de plan d’expériences,” *Journal de la Société Française de Statistique*, vol. 149, no. 1, 2008.
- [16] J.-P. Gauchi and J.-P. Vila, “Nonparametric particle filtering approaches for identification and inference in nonlinear state-space dynamic systems,” *Statistics and Computing*, 2012.
- [17] J. Vovelle, “Résultats généraux sur les équations différentielles.” Cours de L3, Université Claude Bernard Lyon 1.

# Annexe A

## Compléments du chapitre 2

### A.1 Valeurs des paramètres associés à la figure 12

$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$ILL_\infty$	$TI_\infty$	$V_2$	$r_1$	$r_2$	$K_1^0$	$K_2$	$D$	$N_1^0$	$V_1^0$
0,3	0,3	0,05	1	0,1	$4,6 \cdot 10^5$	$2,6 \cdot 10^7$	1	0,6	0,7	$3,1 \cdot 10^9$	$3,0 \cdot 10^9$	10	$4 \cdot 10^7$	2000

### A.2 Régularité de la fonction $F$

Définissons en toute rigueur la fonction  $F$  de l'équation différentielle (2.9). Soient  $I = ]-1, +\infty[$ ,  $\Omega = ]-\frac{1}{2}, +\infty[ \times ]-\frac{1}{2}, +\infty[$  et  $E = \mathbb{R}^2$ . On identifie  $E$  avec  $\mathcal{M}_{2,1}(\mathbb{R})$ , et on suppose que  $(x_1, \dots, x_{12})$  appartient au produit d'intervalles ouverts  $A_1 \times \dots \times A_{12}$  avec  $A_i \subset ]0, +\infty[$  pour tout  $i \neq 3$  et  $A_3 \subset ]0, 1[$ .

On définit  $F$ , pour  $t \in I$  et  $(x, y) \in \Omega$ , par

$$F \left( t, \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} r_1 x \left( 1 - \frac{x \exp(a_5 t)}{K_1^0} \right) - a_5 x - D \cdot g \left( \frac{x \exp(a_5 t)}{V_1^0}, t \right) \cdot \tilde{h}(x, a_3 y) \\ r_2 y \left( 1 - \frac{y}{K_2} \right) + a_5 x - D \cdot a_4 \cdot g \left( \frac{y}{V_2}, t \right) \cdot (1 - \tilde{h}(x, a_3 y)) \end{pmatrix}.$$

La fonction  $\tilde{h}$  est  $\mathcal{C}^1$  sur  $\Omega$ , la fonction  $g$  est  $\mathcal{C}^0$  sur  $\mathbb{R} \times I$ , et  $\mathcal{C}^1$  par morceaux sur tout compact de  $\mathbb{R} \times I$ . Or, pour tout  $t \in I$  et pour tout  $(x, y) \in \Omega$ , on a  $(\frac{x \exp(a_5 t)}{V_1^0}, t) \in \mathbb{R} \times I$ ,  $(\frac{y}{V_2}, t) \in \mathbb{R} \times I$  et  $(x, a_3 y) \in \Omega$  car  $a_3 \in ]0, 1[$ .

Ainsi,  $F$  est continue sur  $I \times \Omega$  et  $\mathcal{C}^1$  par morceaux par rapport à sa deuxième variable sur tout compact de  $I \times \Omega$ . On en déduit que  $F$  localement lipschitzienne par rapport à sa deuxième variable sur  $I \times \Omega$ .

$I$  est un intervalle ouvert de  $\mathbb{R}$ , et  $\Omega$  est un ouvert connexe de l'espace de Banach  $E = \mathbb{R}^2$ . On est donc dans les conditions d'application du théorème de Cauchy-Lipschitz (version locale) : il existe une unique solution maximale  $\mathbf{X}$  au problème de Cauchy, définie sur  $J$ , intervalle ouvert inclus dans  $I$ .

### A.3 Preuve de la proposition 1

**Proposition 1.** *On suppose que  $N_1^0 > 0$ . Alors :*

- (i)  $\forall t \in J, N_1(t) > 0$ ;
- (ii)  $\forall t \in J \cap \mathbb{R}_+^*, N_2(t) > 0$ ;
- (iii)  $\mathbb{R}_+ \subset J$ .

*Démonstration.* (i). On va se ramener à un système autonome de dimension 3 en considérant le temps comme une coordonnée. Soit  $\tilde{F} : \Omega \mapsto \mathbb{R}^3$  définie par :

$$\forall (x, y, z) \in \Omega, \tilde{F} \left( \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = \begin{pmatrix} r_1 x \left( 1 - \frac{x \exp(a_5 z)}{K_1^0} \right) - a_5 x - D \cdot g \left( \frac{x \exp(a_5 z)}{V_1^0}, z \right) \cdot \tilde{h}(x, a_3 y) \\ r_2 y \left( 1 - \frac{y}{K_2} \right) + a_5 x - D a_4 \cdot g \left( \frac{y}{V_2}, z \right) \cdot (1 - \tilde{h}(x, a_3 y)) \\ 1 \end{pmatrix},$$

avec  $\Omega$  ouvert connexe de  $\mathbb{R}^3$ . On a équivalence entre les problèmes de Cauchy suivants :

$$(E) \begin{cases} \mathbf{X}'(t) = F(t, \mathbf{X}(t)) \\ \mathbf{X}(0) = \begin{pmatrix} N_1^0 \\ 0 \end{pmatrix} \end{cases} \iff (\tilde{E}) \begin{cases} \tilde{\mathbf{X}}'(t) = \tilde{F}(\tilde{\mathbf{X}}(t)) \\ \tilde{\mathbf{X}}(0) = \begin{pmatrix} N_1^0 \\ 0 \\ 0 \end{pmatrix} \end{cases}.$$

D'après la définition de  $\tilde{F}$ , toute solution de  $(\tilde{E})$  avec  $N_1^0 = 0$  a sa première coordonnée constante égale à 0. Ainsi, d'après le théorème de Cauchy-Lipschitz, l'ensemble des solutions maximales de  $(\tilde{E})$  avec  $N_1^0 = 0$  forme une partition de  $\Omega \cap (\{0\} \times \mathbb{R} \times \mathbb{R})$ .

Or le système  $(\tilde{E})$  est autonome : par le principe de non-intersection des trajectoires, toute solution de  $(\tilde{E})$  avec  $N_1^0 > 0$  reste dans  $\Omega \cap (\mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R})$ . Il s'ensuit que pour tout  $t \in J, N_1(t) > 0$ .

(ii). D'après (i) et la définition de  $\tilde{F}$ , on remarque que pour tout  $t \in J, (N_2(t) = 0) \Rightarrow (N_2'(t) = a_5 N_1(t) > 0)$ . En particulier,  $N_2(0) = 0$  et  $N_2'(0) > 0$ , donc l'ensemble

$$M = \{\alpha > 0 \mid \forall t \in J \cap ]0, \alpha], N_2(t) > 0\}$$

est non vide. On note  $\beta = \sup M \leq +\infty$ .

Si  $\beta < +\infty$ , comme  $N_2$  est une fonction  $\mathcal{C}^1$  de  $t$ , on a  $N_2(\beta) = 0$  et  $N_2'(\beta) \leq 0$  : contradiction. D'où  $\beta = +\infty$  et  $\forall t \in J \cap \mathbb{R}_+^*, N_2(t) > 0$ .

(iii). On montre par le lemme de Gronwall que  $N_1$  et  $N_2$  sont majorées (preuve laissée au lecteur courageux). Donc d'après (i) et (ii) la solution  $\mathbf{X}$  est bornée, et par le théorème de sortie de tout compact elle est globale (car  $E$  est de dimension finie), d'où  $J = I$  et  $\mathbb{R}_+ \subset J$ .

□

## A.4 Preuve de la proposition 2

**Proposition 2.** *La fonction  $\phi$  est continue sur  $\mathbb{R}_+ \times A_1 \times \cdots \times A_{12}$ .*

*Démonstration.* On note  $x = \mathbf{X}$ ,  $x_0 = \mathbf{X}_0$ ,  $\mathcal{D} = A_1 \times \cdots \times A_p$  et  $\lambda = (x_1, \dots, x_p)$  pour simplifier l'écriture (on est ici dans le cas  $p = 12$ ). On rappelle que  $I = ]-1, +\infty[$  et  $\Omega = ]-\frac{1}{2}, +\infty[ \times ]-\frac{1}{2}, +\infty[$  et  $E = \mathbb{R}^2$ . On considère la fonction  $F$  comme une fonction de trois arguments  $(t, x, \lambda) \in I \times \Omega \times \mathcal{D}$  dans le problème de Cauchy :

$$(P) \begin{cases} x'(t) = F(t, x(t), \lambda) \\ x(0) = x_0 \end{cases}$$

Dans un premier temps, on se ramène au problème de la dépendance en la donnée initiale. On note  $\widehat{\Omega} = \Omega \times \mathcal{D}$  et  $\widehat{E} = E \times \mathbb{R}^p = \mathbb{R}^{2+p}$  (on a donc  $\widehat{\Omega} \subset \widehat{E}$ ). Comme les  $A_i$  sont des intervalles ouverts (par définition, cf. section 2.6.2),  $\widehat{\Omega}$  est bien un ouvert connexe de  $\widehat{E}$ . On identifie  $\widehat{E}$  avec  $\mathcal{M}_{2+p,1}(\mathbb{R})$ , et on définit la fonction  $\widehat{F}$  par :

$$\forall t \in I, \forall (\omega, a) \in \Omega \times \mathcal{D}, \widehat{F} \left( t, \begin{pmatrix} \omega \\ a \end{pmatrix} \right) = \begin{pmatrix} F(t, \omega, a) \\ 0 \end{pmatrix}$$

Le problème de Cauchy  $(P)$  est équivalent au nouveau problème de Cauchy :

$$(\widehat{P}) \begin{cases} y'(t) = \widehat{F}(t, y(t)) \\ y(0) = y_0 \end{cases}$$

avec  $y(t) = \begin{pmatrix} x(t) \\ \rho(t) \end{pmatrix} \in \widehat{\Omega}$ , et  $y_0 = \begin{pmatrix} x_0 \\ \lambda \end{pmatrix} \in \widehat{\Omega}$  (i.e.  $(x(t), \rho(t)) \in \Omega \times \mathcal{D}$  et  $(x_0, \lambda) \in \Omega \times \mathcal{D}$ ).

La fonction  $\widehat{F}$  est bien définie de  $I \times \widehat{\Omega}$  dans  $\widehat{E}$ , et la fonction  $\rho$  vérifie l'équation  $\rho' = 0$ , dont la solution est :  $\rho(t) = cste = \lambda$ . Ainsi, la dépendance de la solution de  $(P)$  par rapport à  $(x_0, \lambda)$  (donc en particulier  $\lambda$ ) se ramène à la dépendance de la solution de  $(\widehat{P})$  par rapport à  $y_0 \in \widehat{\Omega}$ . Pour conclure on a besoin du lemme suivant, démontré en fin de section.

**Lemme.** *La fonction  $\widehat{F}$  est une application  $I \times \widehat{\Omega} \rightarrow \widehat{E}$  continue et localement lipschitzienne par rapport à sa deuxième variable.*

Le problème de Cauchy  $(\widehat{P})$  admet donc une unique solution maximale. Pour  $y_0 \in \widehat{\Omega}$ , on note  $J_{y_0}$  l'intervalle de définition de la solution maximale de  $(\widehat{P})$  avec la condition initiale  $y(0) = y_0$ . On pose :

$$\mathcal{D}_\Phi = \bigcup_{y_0 \in \widehat{\Omega}} (J_{y_0} \times \{y_0\})$$

et on définit le flot  $\Phi$  de l'équation différentielle  $y'(t) = \widehat{F}(t, y(t))$  comme l'application  $\Phi : \mathcal{D}_\Phi \rightarrow \widehat{\Omega}$  qui à  $(t, y_0)$  associe la valeur au temps  $t$  de la solution maximale de  $(\widehat{P})$  de condition initiale  $y(0) = y_0$ . Les conditions suivantes sont vérifiées :  $I$  intervalle ouvert de  $\mathbb{R}$ ,  $\widehat{\Omega}$  ouvert connexe de l'espace de Banach  $\widehat{E}$

et  $\widehat{F}$  application  $I \times \widehat{\Omega} \rightarrow \widehat{E}$  continue et localement lipschitzienne par rapport à sa deuxième variable, et par conséquent le théorème de *continuité du flot*<sup>1</sup> s'applique :

**Théorème.** *L'ensemble  $\mathcal{D}_\Phi$  est ouvert et  $\Phi : \mathcal{D}_\Phi \rightarrow \widehat{\Omega}$  est une application localement lipschitzienne.*

La fonction  $\Phi$  est donc continue sur  $\mathcal{D}_\Phi$ , et si on note  $\Phi = (\Phi_1, \Phi_2)$  où  $\Phi_1$  est à valeurs dans  $\Omega$  et  $\Phi_2$  à valeurs dans  $\mathcal{D}$ , alors la fonction  $\Phi_1$  continue sur  $\mathcal{D}_\Phi$ . Examinons cet ensemble : on remarque que  $\mathcal{D}_\Phi \subset I \times \Omega \times \mathcal{D}$ , et puisque  $\mathbb{R}_+ \subset J_{y_0}$  pour tout  $y_0 \in \widehat{\Omega}$  d'après la propriété 1, on a  $\mathbb{R}_+ \times \Omega \times \mathcal{D} \subset \mathcal{D}_\Phi$ . Le résultat final vient du fait que la fonction  $\phi : \mathbb{R}_+ \times \mathcal{D} \rightarrow \mathbb{R}^2$  peut se voir comme la restriction de  $\Phi_1$  au sous ensemble  $\mathbb{R}_+ \times \{x_0\} \times \mathcal{D}$  (où  $x_0 = \mathbf{X}_0$ ). En effet, on a :

$$\forall t \in \mathbb{R}_+, \forall \lambda \in \mathcal{D}, \phi(t, \lambda) = \Phi_1(t, x_0, \lambda)$$

et *a fortiori* la fonction  $\phi$  est continue sur  $\mathbb{R}_+ \times \mathcal{D}$  (et est même à valeurs dans  $\Omega$ ).  $\square$

*Preuve du lemme.* Montrons que la fonction  $\widehat{F}$  est une application  $I \times \widehat{\Omega} \rightarrow \widehat{E}$  continue et localement lipschitzienne par rapport à sa deuxième variable. Par définition de  $\widehat{F}$ , il suffit de montrer cela pour  $F$ , vue comme application  $I \times \widehat{\Omega} \rightarrow E$ .

On rappelle que la fonction  $\tilde{h}$  est  $\mathcal{C}^1$  sur  $\Omega$  et ne dépend d'aucun des paramètres  $x_i$ . Pour la fonction  $g$ , il est nécessaire de préciser les choses. Définissons d'abord la fonction  $H$  par

$$\forall (x, y, z) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^*, H(x, y, z) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{y}{z}x & \text{si } 0 < x \leq z \\ y & \text{si } x > z \end{cases}$$

puis les fonctions  $\widehat{g}_1$  et  $\widehat{g}_2$  sur  $I \times \widehat{\Omega}$  par

$$\widehat{g}_1(t, \widehat{\omega}) = H\left(\frac{x \exp(x_5 t)}{V_1^0}, x_7 (1 - \exp(-x_2(t+1))), x_6 (1 - \exp(-x_1(t+1)))\right)$$

et

$$\widehat{g}_2(t, \widehat{\omega}) = H\left(\frac{y}{x_8}, x_7 (1 - \exp(-x_2(t+1))), x_6 (1 - \exp(-x_1(t+1)))\right)$$

où  $t \in I$  et  $\widehat{\omega} = (x, y, x_1, \dots, x_p) \in \widehat{\Omega}$ . D'après les définitions (2.3) et (2.4), on a les égalités :

$$\widehat{g}_1(t, \widehat{\omega}) = g\left(\frac{x \exp(x_5 t)}{V_1^0}, t\right) \quad \text{et} \quad \widehat{g}_2(t, \widehat{\omega}) = g\left(\frac{y}{x_8}, t\right)$$

pour tout  $t \in I$  et pour tout  $\widehat{\omega} = (x, y, x_1, \dots, x_p) \in \widehat{\Omega}$ . La définition de  $F$  en tant qu'application  $I \times \widehat{\Omega} \rightarrow E$  est donc (toujours pour  $t \in I$  et  $\widehat{\omega} = (x, y, x_1, \dots, x_p) \in \widehat{\Omega}$ ) :

$$F(t, \widehat{\omega}) = \begin{pmatrix} x_9 x \left(1 - \frac{x \exp(x_5 t)}{x_{11}}\right) - x_5 x - D \cdot \widehat{g}_1(t, \widehat{\omega}) \cdot \tilde{h}(x, x_3 y) \\ x_{10} y \left(1 - \frac{y}{x_{12}}\right) + x_5 x - D \cdot x_4 \cdot \widehat{g}_2(t, \widehat{\omega}) \cdot (1 - \tilde{h}(x, x_3 y)) \end{pmatrix}$$

---

1. On trouve une preuve de ce théorème dans un cours de J. Vovelle [17].



La fonction  $F : I \times \widehat{\Omega} \rightarrow E$  est clairement continue. Pour montrer que  $F$  est localement lipschitzienne (loc. lip.) par rapport à sa deuxième variable, il suffit de montrer que les fonctions  $\widehat{g}_1$  et  $\widehat{g}_2$  le sont. Il ne semble pas simple de formuler un théorème général pour de telles fonctions composées, aussi nous allons montrer le résultat “à la main”. Considérons  $\widehat{g}_1$ . On est en dimension finie et on a donc équivalence des normes, mais pour fixer les idées on munit  $I \times \widehat{\Omega}$  de la norme 1 sur  $\mathbb{R}^{p+3}$ , notée  $\|\cdot\|_1$  (somme des valeurs absolues). Il s’agit de montrer que pour tout  $(t_0, z_0) \in I \times \widehat{\Omega}$ , il existe  $V_0 \subset I \times \widehat{\Omega}$  un voisinage de  $(t_0, z_0)$  et  $L_0 \in \mathbb{R}_+$  tels que :

$$\forall (t, z_1) \in V_0, \forall (t, z_2) \in V_0, |\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq L_0 \|z_1 - z_2\|_1.$$

On adopte dans la suite les notations suivantes (il n’y aura pas d’ambiguïté avec d’éventuelles puissances) :  $z_1 = (x^1, y^1, x_1^1, \dots, x_p^1)$  et  $z_2 = (x^2, y^2, x_1^2, \dots, x_p^2)$ . On définit les ensembles :

$$\begin{aligned} A &:= \left\{ (x, y, x_1, \dots, x_p) \in \widehat{\Omega} ; x < 0 \right\}, \\ B &:= \left\{ (x, y, x_1, \dots, x_p) \in \widehat{\Omega} ; 0 < \frac{x \exp(x_5 t)}{V_1^0} < x_6 (1 - \exp(-x_1(t+1))) \right\}, \\ C &:= \left\{ (x, y, x_1, \dots, x_p) \in \widehat{\Omega} ; \frac{x \exp(x_5 t)}{V_1^0} > x_6 (1 - \exp(-x_1(t+1))) \right\}. \end{aligned}$$

Ce sont des ouverts disjoints et on a  $\widehat{\Omega} \subset \bar{A} \cup \bar{B} \cup \bar{C}$  et  $\widehat{\Omega} \cap \bar{A} \cap \bar{C} = \emptyset$ .

Soit maintenant  $(t_0, z_0) \in I \times \widehat{\Omega}$ . Il y a 3 cas possibles :

- Cas  $z_0 \in A$  (resp.  $z_0 \in B, z_0 \in C$ ) : il existe  $V_0$  voisinage de  $(t_0, z_0)$  inclus dans  $I \times A$  (resp.  $I \times B, I \times C$ ), et le caractère loc. lip. découle du fait que  $\widehat{g}_1$  est  $\mathcal{C}^1$  sur  $I \times A$  (resp.  $I \times B, I \times C$ ).
- Cas  $z_0 \in (\bar{A} \cap \bar{B})$  : il existe  $V_0$  voisinage de  $(t_0, z_0)$  inclus dans  $I \times (\bar{A} \cup \bar{B})$  et inclus dans un compact de  $I \times \widehat{\Omega}$ . Soient  $(t, z_1) \in V_0$  et  $(t, z_2) \in V_0$ . Si  $z_1 \in \bar{A}$  et  $z_2 \in \bar{A}$  le résultat est trivial par définition de  $H$ . Si  $z_1 \in B$  et  $z_2 \in B$ , d’après le cas précédent il existe  $k_1 \in \mathbb{R}_+$  tel que  $|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq k_1 \|z_1 - z_2\|_1$ . Sinon, par symétrie on peut supposer que  $z_1 \in \bar{A}$  et  $z_2 \in \bar{B} \setminus A$ . On a alors :

$$\begin{aligned} |\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| &= \frac{x^2 \exp(x_5^2 t) x_7^2 (1 - \exp(-x_2^2(t+1)))}{V_1^0 x_6^2 (1 - \exp(-x_1^2(t+1)))} \\ &\leq k_2 x^2 \text{ où } k_2 \in \mathbb{R}_+ \text{ (car } V_0 \text{ est inclus dans un compact de } I \times \widehat{\Omega}) \\ &\leq k_2 |x^2 - x^1| \text{ puisque } x^1 = 0 \text{ et } x^2 \geq 0. \\ &\leq k_2 \|z_1 - z_2\|_1 \end{aligned}$$

d’où  $|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq \max(k_1, k_2) \|z_1 - z_2\|_1 \quad \forall (t, z_1) \in V_0, \forall (t, z_2) \in V_0$ .

- Cas  $z_0 \in (\bar{B} \cap \bar{C})$  : il existe  $V_0$  voisinage de  $(t_0, z_0)$  inclus dans  $I \times (\bar{B} \cup \bar{C})$  et inclus dans un compact de  $I \times \widehat{\Omega}$ . Soient  $(t, z_1) \in V_0$  et  $(t, z_2) \in V_0$ . De la même manière que précédemment, on peut se restreindre au cas  $z_1 \in C$  et  $z_2 \in \bar{B} \setminus A$ . On a :

$$|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| = \left| x_1^1 (1 - \exp^{(1)}(-x_2^1(t+1))) - \frac{x^2 \exp(x_5^2 t) x_7^2 (1 - \exp^{(2)}(-x_2^2(t+1)))}{V_1^0 x_6^2 (1 - \exp(-x_1^2(t+1)))} \right|.$$

– Si (1)  $\geq$  (2) :  $z_1 \in C$  donc

$$|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq \left| \frac{x^1 \exp(x_5^1 t) x_7^1 (1 - \exp(-x_2^1(t+1)))}{V_1^0 x_6^1 (1 - \exp(-x_1^1(t+1)))} - \frac{x^2 \exp(x_5^2 t) x_7^2 (1 - \exp(-x_2^2(t+1)))}{V_1^0 x_6^2 (1 - \exp(-x_1^2(t+1)))} \right|$$

$$\leq k_3 \|z_1 - z_2\|_1 \text{ (fonction } C^1 \text{ sur } V_0 \text{ inclus dans un compact)}$$

– Si (1)  $<$  (2) :  $z_2 \in \bar{B}$  donc

$$|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq |x_7^1 (1 - \exp(-x_2^1(t+1))) - x_7^2 (1 - \exp(-x_2^2(t+1)))|$$

$$\leq k_4 \|z_1 - z_2\|_1 \text{ (fonction } C^1 \text{ sur } V_0 \text{ inclus dans un compact)}$$

d'où  $|\widehat{g}_1(t, z_1) - \widehat{g}_1(t, z_2)| \leq \max(k_3, k_4) \|z_1 - z_2\|_1 \quad \forall (t, z_1) \in V_0, \forall (t, z_2) \in V_0$ .

Ainsi  $\widehat{g}_1$  est bien loc. lip. par rapport à sa deuxième variable. On montre de la même manière que  $\widehat{g}_2$  l'est aussi, et on a donc le résultat sur  $F$ . □

## A.5 Distribution *a priori* pour les méthodes de Monte Carlo

Indice	Paramètre	Borne inférieure	Borne supérieure
1	$a_1$	0,01	0.5
2	$a_2$	0,01	0.5
3	$a_5$	0,01	0.5
4	$ILL_\infty$	$0,1 \times 10^5$	$10 \times 10^5$
5	$TI_\infty$	$0,1 \times 10^7$	$10 \times 10^7$
6	$r_1$	0	2
7	$r_2$	0	2
8	$K_1^0$	$0,1 \times 10^9$	$10 \times 10^9$
9	$K_2$	$0,1 \times 10^9$	$10 \times 10^9$
10	$\sigma^2$	$0,1 \times 10^{17}$	$10 \times 10^{17}$

TABLEAU 7 – Bornes de la distribution *a priori* pour les méthodes de Monte Carlo

## Annexe B

# Conditionnement dans le cas des variables à densité

Soit  $Y$  une variable aléatoire réelle. On note  $p$  la dimension de  $X$ , *i.e.* le nombre de paramètres (dans notre cas  $p = 12$ ). On suppose que le vecteur aléatoire  $(X, Y) = (X_1, \dots, X_p, Y)$  admet une densité notée  $f_{(X,Y)}$ . Les variables  $X_1, \dots, X_p$  et  $Y$  admettent donc également des densités, appelées *densités marginales* et notées respectivement  $f_{X_1}, \dots, f_{X_p}$  et  $f_Y$ . Il en est de même pour chaque  $n$ -uplet de ces variables pour  $n \leq p$  (on note les densités associées  $f_{(X_1, X_2)}, f_{(X_1, X_2, X_3)}, \dots$ ).

On souhaite utiliser la notion de loi conditionnelle dans le cas de variables à densité. Ce concept semble paradoxal : en effet, soit  $B \in \mathcal{B}(\mathbb{R})$  et soit  $i \in \{1, \dots, p\}$ , l'écriture  $\mathbb{P}(Y \in B | X_i = x_i)$  n'a *a priori* pas de sens, puisque  $\mathbb{P}(X_i = x_i) = 0$  pour tout  $x_i \in A_i$ . Cependant, il est possible de lui en donner un, par analogie avec le cas discret, en définissant la notion de densité conditionnelle (qui peut en fait se justifier rigoureusement dans le cadre plus général de l'espérance conditionnelle).

**Proposition 3.** Soit  $i \in \{1, \dots, p\}$ . Pour tout  $x_i \in A_i$  tel que  $f_{X_i}(x_i) > 0$ , la formule suivante définit une densité sur  $\mathbb{R}^2$  :

$$f_{Y|X_i=x_i}(y) = \frac{f_{(X_i, Y)}(x_i, y)}{f_{X_i}(x_i)}.$$

Cette fonction s'appelle **densité conditionnelle** de  $Y$  sachant  $X_i = x_i$ .

La preuve est immédiate puisque  $f_{Y|X_i=x_i}$  est une fonction positive d'intégrale 1. On peut alors définir  $\mathbb{P}(Y \in B | X_i = x_i) = \int_B f_{Y|X_i=x_i}(y) dy$ , ce qui nous permet, en supposant  $Y$  intégrable, de généraliser la notion d'espérance conditionnelle.

**Définition 5.** On suppose que  $Y$  est intégrable.

1) L'espérance conditionnelle de  $Y$  sachant  $X_i = x_i$  est définie par :

$$\mathbb{E}(Y | X_i = x_i) = \int_{\mathbb{R}^2} y f_{Y|X_i=x_i}(y) dy.$$

2) **L'espérance conditionnelle de  $Y$  sachant  $X_i$**  est la variable aléatoire définie par :

$$\mathbb{E}(Y|X_i) = \phi_i(X_i), \text{ avec } \forall x_i \in A_i, \phi_i(x_i) = \mathbb{E}(Y|X_i = x_i).$$

On peut définir de la même manière (en supposant cette fois que  $\text{Var}(Y) < +\infty$ ) :

- la **variance conditionnelle de  $Y$  sachant  $X_i = x_i$** , notée  $\text{Var}(Y|X_i = x_i)$ ;
- la **variance conditionnelle de  $Y$  sachant  $X_i$** , notée  $\text{Var}(Y|X_i)$ .

Pour  $i \in \{1, \dots, p\}$  fixé, on obtient les trois résultats suivants, importants pour la suite.

**Théorème 3** (Bayes). *Pour tout  $x_i \in A_i$  et pour tout  $y \in \mathbb{R}^2$ , on a :*

$$f_{Y|X_i=x_i}(y) = \frac{f_Y(y)}{f_{X_i}(x_i)} f_{X_i|Y=y}(x_i).$$

*Démonstration.* Immédiat d'après la proposition 3, en écrivant que  $f_{(X_i, Y)}(x_i, y) = f_{(Y, X_i)}(y, x_i)$ .  $\square$

**Théorème 4** (Loi de la probabilité totale). *On a l'égalité suivante :*

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X_i)).$$

*Démonstration.* D'après la définition 5 et le théorème de transport, on a

$$\mathbb{E}(\mathbb{E}(Y|X_i)) = \mathbb{E}(\phi_i(X_i)) = \int_{A_i} \phi_i(x_i) f_{X_i}(x_i) dx_i,$$

et par conséquent d'après le théorème de Bayes et le théorème de Fubini :

$$\begin{aligned} \mathbb{E}(\mathbb{E}(Y|X_i)) &= \int_{A_i} \mathbb{E}(Y|X_i = x_i) f_{X_i}(x_i) dx_i \\ &= \int_{A_i} \left( \int_{\mathbb{R}^2} y f_{Y|X_i=x_i}(y) dy \right) f_{X_i}(x_i) dx_i \\ &= \int_{A_i} \left( \int_{\mathbb{R}^2} y \frac{f_Y(y)}{f_{X_i}(x_i)} f_{X_i|Y=y}(x_i) f_{X_i}(x_i) dy \right) dx_i \\ &= \int_{\mathbb{R}^2} \left( \int_{A_i} y f_Y(y) f_{X_i|Y=y}(x_i) dx_i \right) dy \\ &= \int_{\mathbb{R}^2} y f_Y(y) \left( \int_{A_i} f_{X_i|Y=y}(x_i) dx_i \right) dy \\ &= \int_{\mathbb{R}^2} y f_Y(y) dy \\ &= \mathbb{E}(Y) \end{aligned}$$

car  $\int_{A_i} f_{X_i|Y=y}(x_i) dx_i = 1$  puisque  $f_{X_i|Y=y}$  est une fonction de densité sur  $A_i$ .  $\square$

**Remarque 9.** *Ce résultat est en fait une propriété immédiate de l'espérance conditionnelle en général.*

**Remarque 10.** *On peut clairement généraliser tous ces résultats en remplaçant  $X_i$  par n'importe quel  $n$ -uplet  $(X_{i_1}, \dots, X_{i_n})$  avec  $n \leq p$  (ex :  $\mathbb{E}(Y|X_1, X_2)$ ,  $\text{Var}(Y|X_3, X_7, X_8)$ , ...), et même pour  $Y$  variable aléatoire à densité quelconque.*