

Chapitre 3

Identification de modèles de bioprocédés

3.1. Introduction

La complexité des modèles dynamiques de bioprocédés habituellement utilisés est le reflet de la complexité des procédés eux-mêmes ainsi que de la connaissance détaillée acquise à leur sujet au cours des dernières décennies. Ceux-ci comportent en réalité un grand nombre de paramètres auxquels il faut attribuer une valeur numérique soit sur base d'une connaissance *a priori* de leur valeur, soit sur base de données expérimentales. La qualité de l'estimation des paramètres dépendra intrinsèquement de la quantité et de la qualité de données qui sont mises à la disposition de la procédure de calibration (ou d'identification) des paramètres du modèle. A côté de ces limitations sur l'information disponible, un autre problème surgit dans l'identification des paramètres du modèle : les modèles dynamiques des bioprocédés sont intrinsèquement non linéaires et les paramètres du modèle peuvent être fortement corrélés entre eux. L'étude de l'*identifiabilité* des paramètres du modèle avant leur estimation s'avère être une tâche essentielle.

La question centrale de l'analyse d'identifiabilité peut se résumer comme suit : *supposons qu'un certain nombre de variables d'état du modèle sont disponibles à la mesure ; sur base de la structure du modèle (identifiabilité structurelle) ou sur base du type et de la qualité des données disponibles (identifiabilité pratique), peut-on s'attendre à pouvoir attribuer une valeur unique à chaque paramètre du modèle via l'estimation paramétrique ?* La réponse à l'analyse d'identifiable ne s'exprime souvent pas simplement sous forme d'un « oui » ou d'un « non », mais quand elle donne des résultats (ce qui n'est pas nécessairement évident avec des modèles non linéaires), ceux-ci

peuvent très bien dire que tel ou tel sous-ensemble ou combinaison de paramètres est identifiable.

Le chapitre est organisé comme suit. Nous abordons d'abord les questions d'identifiabilité structurelle et d'identifiabilité pratique, avant de voir comment concevoir et mettre en œuvre des expériences (plans d'expériences) qui permettent d'obtenir les meilleures données pour l'estimation des paramètres et de passer en revue différentes techniques d'identification paramétrique. Nous terminerons par un cas d'étude : l'identification des paramètres d'un modèle de digestion anaérobie.

3.2. Identifiabilité structurelle

Dans le cas de l'identifiabilité structurelle, on examine si, dans le cas (académique) de mesures idéales (mesures continues non bruitées qui s'ajustent parfaitement au modèle), tous les paramètres du modèle étudié sont identifiables. La conclusion sera peut-être que seulement une combinaison de paramètres est identifiable. Si le nombre de combinaisons de paramètres est plus faible que le nombre de paramètres du modèle, ou s'il n'y a pas de relation bijective entre ces combinaisons et les paramètres, alors seule une connaissance *a priori* de certains paramètres permettra de résoudre le problème de l'identifiabilité. Prenons un exemple simple : dans le modèle $y = ax_1 + bx_2 + c(x_1 + x_2)$, seuls les paramètres $a + c$ et $b + c$ sont structurellement identifiables (et non pas les trois paramètres a, b, c) ; deux paramètres (par exemple, a et b) seront identifiables si la valeur d'un troisième (ici c) est connue *a priori*.

Pour les systèmes linéaires, l'identifiabilité structurelle est bien comprise et il existe plusieurs tests disponibles (voir par exemple [GOD 85]). Par contre, le problème s'avère bien plus complexe pour les modèles non linéaires et seulement quelques méthodes permettent aujourd'hui de tester leur identifiabilité, avec souvent la possibilité de ne pouvoir mettre en évidence que des conditions nécessaires et/ou suffisantes d'identifiabilité locale (c'est-à-dire valide pour certains domaines du modèle seulement) [WAL 82]. La liste ci-dessous résume les méthodes disponibles, y compris les références à des applications centrées sur des modèles incluant des cinétiques de Monod :

- transformation du modèle non linéaire en un modèle linéaire [GOD 85]. Applications : [BOU 98, DOC 95, SPE 00] ;
- développement en séries :
 - séries de Taylor [GOD 85, POH 78]. Des applications : [BOU 98, DOC 95, HOL 82a, JEP 96, PET 00b, SPE 00] ;
 - séries génératrices [WAL 95]. Applications : [PET 00a] ;
- transformation de similarité ou isomorphisme d'état local [WAL 95]. Applications : [JUL 97, JUL 98] ;

– étude de l’observabilité du système non linéaire [CAS 85]. Des applications : [BOU 98].

De manière générale, à la fois pour les modèles linéaires et non linéaires, il est difficile de déterminer quelle approche demandera le moins de calculs pour un exemple spécifique. Dans beaucoup de cas, ce sont les méthodes de développement en série qui sont utilisées, car elles sont relativement simples à mettre en œuvre et surtout leur développement suit une démarche continue (on peut toujours calculer les termes d’ordre supérieur de la série). Pour la méthode des séries de Taylor, celles-ci sont générées dans le domaine temporel, tandis que les séries génératrices se développent dans le domaine des entrées. Celles-ci peuvent être vues comme une extension de la méthode des séries de Taylor dans le cas où l’on considère une classe d’entrées [RAK 85]. Dans le cas de modèle sans variable d’entrée, la méthode des séries génératrices devient équivalente à la méthode des séries de Taylor [VAJ 89].

3.2.1. Développement en séries de Taylor

L’approche par développement en série de Taylor a été à l’origine développée par Pohjanpalo [POH 78]. Considérons le modèle général suivant :

$$\frac{dx(t)}{dt} = f(x(t), u(t), t, \theta), \quad x(0) = x_0(\theta) \quad [3.1a]$$

$$y(t, \theta) = g(x(t), \theta) \quad [3.1b]$$

où x , u , y et θ représentent respectivement le vecteur d’état, le vecteur d’entrée, le vecteur de sortie (variables mesurées) et le vecteur des paramètres. Le développement en série de Taylor de $y(t)$ autour de $t = 0$ est égal à :

$$y(t, \theta) = y(0) + t \frac{dy}{dt}(0) + \frac{t^2}{2!} \frac{d^2y}{dt^2}(0) + \dots + \frac{t^k}{k!} \frac{d^k y}{dt^k}(0), \quad k = 0, 1, \dots, \infty \quad [3.2]$$

L’approche suppose que la variable $y(t)$ et ses dérivées successives sont connues (via la valeur de y). Ecrivons les expressions des dérivées successives de $y(t)$ [3.2] obtenues par calcul à partir du modèle [3.1] comme suit :

$$\frac{d^k y}{dt^k}(0) = a_k(\theta) \quad [3.3]$$

où $a_k(\theta)$ sont des fonctions (typiquement non linéaires) des paramètres θ . L’équation [3.3] est un système d’équations non linéaires où les inconnues sont les paramètres θ et où les termes $\frac{d^k y}{dt^k}(0)$ sont, par hypothèse, connus. La méthode consiste à résoudre ce système d’équations pour un nombre limité de dérivées de $y(t)$. Il est évident que ce nombre doit être au moins égal au nombre de paramètres inconnus. Une condition suffisante d’identifiabilité structurelle est qu’il existe une solution unique à ce système d’équations.

3.2.2. Séries génératrices

Dans le cas des séries génératrices, on considère le modèle général sous la forme suivante [WAL 95] :

$$\frac{dx(t)}{dt} = f^0(x(t), \theta) + \sum_{i=1}^m u_i(t) f^i(x(t), \theta), \quad x(0) = x_0(\theta) \quad [3.4a]$$

$$y(t, \theta) = g(x(t), \theta) \quad [3.4b]$$

où les entrées $u_i(t)$ apparaissent sous forme linéaire. La méthode des séries génératrices est basée sur la fonction de sortie $g(x(t), \theta)$ et ses dérivées de Lie successives, évaluées en $t = 0$:

$$L_{f_i} g(x(t), \theta)_{t=0} = \sum_{k=1}^n f^{i,k}(x, \theta) \frac{\partial}{\partial x_k} g(x(t), \theta)_{t=0} = s_i(\theta) \quad [3.5]$$

$f^{i,k}$ est la k^e composante de f^i . On retrouve ici également un système d'équations où les inconnues sont les paramètres θ (dans les fonctions $s_i(\theta)$). Une condition suffisante d'identifiabilité structurelle sera donc qu'il existe une solution unique pour θ au système d'équations ci-dessus, semblable à celui du développement en série de Taylor [3.3].

3.2.3. Exemples d'application des méthodes de développement en série

La mise en œuvre des deux méthodes de développement en série peut suivre les étapes suivantes :

- 1) calcul des dérivées successives de $g(x, \theta)$;
- 2) choix du jeu de paramètres θ à identifier ;
- 3) évaluation des dérivées successives en insérant les quantités connues et les dérivées d'ordre inférieur ;
- 4) expression des dérivées comme fonctions des combinaisons de paramètres choisis en 2 ;
- 5) résolution des équations de l'étape 4 ;
- 6) si la solution est unique, les paramètres θ sont structurellement identifiables ; sinon, retour à l'étape 2.

Appliquons la procédure à un modèle de dégradation de substrat S dans un réacteur *batch* avec cinétique de Monod :

$$\frac{dS}{dt} = -\frac{\mu_{max} X}{Y} \frac{S}{K_S + S} \quad S(t=0) = S_0 \quad [3.6]$$

où le taux de respiration $OUR(t)$ est mesuré. Celui-ci est relié au substrat S par l'équation suivante :

$$OUR = -(1 - Y) \frac{\mu_{max} X}{Y} \frac{S}{K_S + S} \quad [3.7]$$

Calculons les dérivées de $OUR(t)$ en $t = 0$:

$$OUR(0) = \frac{\mu_{max} X (1 - Y)}{Y} \frac{S_0}{K_S + S_0} \quad [3.8]$$

$$\frac{dOUR}{dt}(0) = -\frac{\mu_{max}^2 X^2 (1 - Y)}{Y^2} \frac{K_S S_0}{(K_S + S_0)^3} \quad [3.9]$$

$$\frac{d^2 OUR}{dt^2}(0) = \frac{\mu_{max}^3 X^3 (1 - Y)}{Y^3} \frac{K_S S_0 (K_S - 2S_0)}{(K_S + S_0)^5} \quad [3.10]$$

Il y a cinq paramètres à identifier : Y , μ_{max} , X , K_S et S_0 . Si l'on écrit $z_i = \frac{d^i OUR}{dt^i}(0)$, $i = 0, 1, 2, \dots$, et que l'on considère les combinaisons de paramètres suivantes :

$$\theta_1 = \frac{\mu_{max} X (1 - Y)}{Y}, \quad \theta_2 = (1 - Y) S_0, \quad \theta_3 = (1 - Y) K_S \quad [3.11]$$

les équations [3.8], [3.9] et [3.10] peuvent être réécrites sous la forme (équivalente) suivante :

$$z_0 = \frac{\theta_1 \theta_2}{\theta_2 + \theta_3} \quad [3.12]$$

$$z_1 = -\frac{\theta_1^2 \theta_2 \theta_3}{(\theta_2 + \theta_3)^3} \quad [3.13]$$

$$z_2 = \frac{\theta_1^3 \theta_2 \theta_3 (\theta_3 - 2\theta_2)}{(\theta_2 + \theta_3)^5} \quad [3.14]$$

Dès lors, les « paramètres » θ_1 , θ_2 et θ_3 peuvent être (formellement) déterminés à partir des valeurs de z_i (ou autrement des dérivées successives de $OUR(t)$) en inversant les expressions ci-dessus :

$$\theta_1 = \frac{z_0 (z_0 z_2 - 3z_1^2)}{z_0 z_2 - z_1^2} \quad [3.15]$$

$$\theta_2 = -\frac{2z_0^2 z_1}{z_0 z_2 - 3z_1^2} \quad [3.16]$$

$$\theta_3 = -\frac{4z_0^2 z_1^3}{(z_0 z_2 - 3z_1^2)(z_0 z_2 - z_1^2)} \quad [3.17]$$

3.2.4. Quelques remarques sur les méthodes de test d'identifiabilité structurelle

De l'exemple ci-dessus, on peut déjà tirer quelques conclusions intéressantes sur la mise en œuvre des méthodes de test d'identifiabilité structurelle.

L'avantage des méthodes de développement en série est qu'elle permet de suivre une procédure clairement identifiée, ce qui n'est pas nécessairement le cas pour les autres méthodes. Ceci dit, le nombre maximum d'équations nécessaires pour obtenir un résultat est inconnu. Il n'y a en tout cas, de manière générale, aucune garantie qu'une information supplémentaire sera apportée en ajoutant des dérivées d'ordre supérieur. Cette absence de borne supérieure sur le nombre de termes à considérer fait que le résultat est une condition seulement suffisante (mais pas nécessaire) d'identifiabilité.

Il est par ailleurs difficile de savoir toujours *a priori* quelles seront les combinaisons de paramètres identifiables. Il n'y a en tout cas aucune règle systématique. Ceci dit, soit le bon sens, soit la structure des termes dans les développements successifs sont souvent une bonne source pour un choix initial. Par exemple, le choix de $\theta_1 (= -\frac{\mu_{max} X(1-Y)}{Y})$ apparaît évident à la lecture des équations [3.8], [3.9] et [3.10]. Récemment, Petersen [PET 00a] a proposé une méthode qui permet de suggérer des combinaisons potentiellement identifiables pour une classe de modèles de bioprocédés décrits par des cinétiques de Monod.

3.3. Identifiabilité pratique

L'identifiabilité pratique est complémentaire de l'identifiabilité structurelle, en ce sens que l'on s'intéresse maintenant à l'impact des données expérimentales disponibles (on ne considère plus des mesures idéales !) sur l'identifiabilité des paramètres du modèle étudié. Un exemple typique de la distinction entre les deux concepts d'identifiabilité est le modèle de Monod dans un procédé de croissance microbienne simple (un substrat, une biomasse). On peut montrer que ce modèle est structurellement identifiable [ABO 78], mais dans la pratique, celui-ci s'avère souvent non identifiable, étant donné la pauvre qualité des données expérimentales disponibles et leur nombre limité (voir [HOL 82a]). Cela se manifeste par exemple sous la forme suivante : toute variation d'un paramètre (par exemple, μ_{max} dans le cas du modèle de Monod) peut être presque parfaitement compensée par une variation proportionnelle d'un autre paramètre (K_S dans le cas du modèle de Monod) et pourtant toujours reproduire une bonne adéquation entre les données expérimentales et les prédictions numériques du modèle dynamique. De plus, il peut arriver que les algorithmes utilisés pour estimer la valeur des paramètres d'un modèle non linéaire aient de mauvaises propriétés de convergence ou s'avèrent mal conditionnés numériquement. Dans ce cas, les valeurs estimées peuvent être très sensibles aux conditions de l'estimation. On peut avoir alors

des valeurs de paramètres susceptibles de varier dans de larges gammes, ce qui rend difficile l'interprétation physique de leur valeur.

La question centrale est donc la suivante : étant donné certaines données expérimentales, quelle précision peut-on obtenir pour les valeurs estimées des paramètres, ou, en d'autres mots, quel est l'effet d'une petite variation de la valeur des paramètres sur l'adéquation entre modèle et données ?

3.3.1. Cadre théorique

Voyons comment formuler cette question mathématiquement [MUN 91]. L'estimation des paramètres peut, en réalité, être formulée comme la minimisation du critère quadratique suivant par un choix optimal de valeurs pour les paramètres θ :

$$J(\theta) = \sum_{i=1}^N (y_i(\hat{\theta}) - y_i)^T Q_i (y_i(\hat{\theta}) - y_i) \quad [3.18]$$

où y_i et $y_i(\hat{\theta})$ sont les vecteurs des N valeurs mesurées et prédictions du modèle aux instants t_i ($i = 1, \dots, N$) et Q_i est une matrice (carrée) de pondération. La valeur de J obtenue pour un jeu de paramètres légèrement différents des valeurs optimales est donnée par l'expression suivante [MUN 89] :

$$E[J(\theta + \delta\theta)] \approx \delta\theta^T \left[\sum_{i=1}^N \left[\frac{\partial y}{\partial \theta}(t_i) \right]^T Q_i \left[\frac{\partial y}{\partial \theta}(t_i) \right] \right] \delta\theta + \sum_{i=1}^N \text{tr}(C_i C_i^{-1}) \quad [3.19]$$

où C_i représente la matrice de covariance des erreurs de mesure (Q_i est souvent choisie égale à C_i^{-1}). Une conséquence importante de [3.19] est que pour optimiser l'identifiabilité pratique (c'est-à-dire maximiser la différence entre $J(\theta + \delta\theta)$ et $J_{opt}(\theta)$), il faut maximiser le terme entre parenthèses [.]. Ce terme est en réalité la matrice d'information de Fisher et exprime le contenu informatif des données expérimentales [LJU 87] :

$$F = \sum_{i=1}^N \left[\frac{\partial y}{\partial \theta}(t_i) \right]^T Q_i \left[\frac{\partial y}{\partial \theta}(t_i) \right] \quad [3.20]$$

Cette matrice est l'inverse de la matrice de covariance des erreurs d'estimation du meilleur estimateur linéaire non biaisé [GOD 85] :

$$V = F^{-1} = \left[\sum_{i=1}^N \left[\frac{\partial y}{\partial \theta}(t_i) \right]^T Q_i \left[\frac{\partial y}{\partial \theta}(t_i) \right] \right]^{-1} \quad [3.21]$$

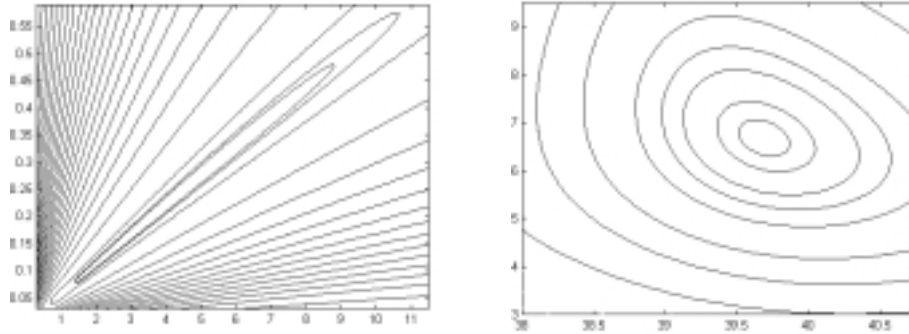


Figure 3.1. Tracé de contours d'une fonction objectif mal conditionnée (gauche) et bien conditionnée (droite) dans le cas de l'estimation de deux paramètres [PET 00a]

Les termes $\frac{\partial y}{\partial \theta}$ sont les fonctions de sensibilité des variables de sortie. Elles quantifient la dépendance des prédictions du modèle vis-à-vis des paramètres.

L'approximation [3.19] de la fonction objectif permet de tracer des courbes de la fonction pour des valeurs données dans l'espace des paramètres. Dans le cas d'un modèle à deux paramètres (μ_{max} et K_s dans le cas du modèle de Monod, par exemple), ces courbes sont des ellipses. Comme le souligne Munack [MUN 89], les axes des ellipses correspondent aux vecteurs propres de la matrice de Fisher et leur longueur est proportionnelle à la racine carrée de l'inverse de la valeur propre correspondante. Ainsi, le rapport de la plus grande valeur propre sur la plus petite (en valeur absolue) est une mesure de la forme de la fonction près des valeurs estimées optimales des paramètres. La figure 3.1 (gauche) montre une ellipse très aplatie correspondant à un rapport de valeurs propres très élevé, alors que dans la figure 3.1 (droite), les valeurs propres sont proches.

3.3.2. Intervalle de confiance des paramètres estimés

Un résultat important de l'identifiabilité pratique est la possibilité de calculer la variance des paramètres, qui permet de calculer un intervalle de confiance pour ceux-ci.

Si la matrice de covariance V [3.21] est connue et en considérant Q_i comme l'inverse de la matrice de covariance des erreurs de mesure, on peut calculer de manière approchée des écarts types pour les paramètres comme suit :

$$\sigma(\theta_i) = \sqrt{V_{ii}} \quad [3.22]$$

Les intervalles de confiance sont alors obtenus de la manière suivante :

$$\theta \pm t_{\alpha; N-p} \sigma(\theta_i) \quad [3.23]$$

pour un niveau de confiance de $100(1 - \alpha)$ % et des valeurs de t obtenues à partir des distributions de Student.

En appliquant l'équation [3.22], on commet une erreur, parce que V_{ii} ne contient que l'erreur de mesure. Dans le cas où une seule variable mesurée est utilisée pour l'identification, une estimation plus réaliste de l'intervalle de confiance peut être obtenue en évaluant :

$$s^2 = \frac{J_{opt}(\theta)}{N - p} \quad [3.24]$$

avec p le nombre de paramètres du modèle et $J_{opt}(\theta)$ défini par l'équation [3.18] avec Q_i la matrice identité. En effet, s^2 calculé avec [3.24] contient aussi l'erreur de modèle. Les déviations standard peuvent alors être calculées de manière approchée comme suit :

$$\sigma(\theta_i) = s \sqrt{V_{ii}} \quad [3.25]$$

Une manière d'obtenir les intervalles de confiance exacts pour les systèmes non linéaires est d'explorer systématiquement toutes les valeurs de la fonction objectif pour un très grand nombre de combinaisons de paramètres. Cela se fait au prix d'un grand nombre de calculs informatiques, car le nombre d'évaluations croît exponentiellement avec le nombre de paramètres [LOB 91].

De manière générale, les intervalles de confiance sont décrits par des hypervolumes bornés par des hypersurfaces dans l'espace des paramètres. Dans le cas particulier d'un modèle à 2 paramètres, cela correspond à un domaine de dimension 2 dans l'espace de dimension 3 généré par $(\theta_1, \theta_2, J(\theta))$. Cela permet de visualiser certains problèmes d'estimation paramétrique (figure 3.1). Dans cet espace, la fonction objectif est une surface et l'intervalle de confiance est délimité par une courbe. L'intervalle de confiance $(1 - \alpha)$ correspond à l'ensemble des valeurs de paramètres pour lesquels la fonction objectif est plus petite que la valeur limite [BEA 60] :

$$J_{opt} * \left(1 + \frac{p}{N - p} F_{\alpha; p, N-p}\right) \quad [3.26]$$

où N et p sont respectivement le nombre de mesures et de paramètres et où F est la valeur de la F -distribution avec p et $N - p$ degrés de liberté et un niveau de confiance α . Dans la figure 3.1, les courbes de contour sont calculées selon cette approche pour des valeurs de α croissantes.

3.3.3. Fonctions de sensibilité

Comme suggéré ci-dessus, les fonctions de sensibilité jouent un rôle-clé dans l'évaluation de l'identifiabilité pratique de paramètres. Une propriété importante est que si les fonctions de sensibilité sont proportionnelles, la matrice de covariance [3.21] est singulière : le modèle est dès lors non identifiable [ROB 85].

Il s'avère que dans de nombreux cas de modèles de procédés biologiques, les fonctions de sensibilité sont souvent quasi linéairement dépendantes. Cela donne souvent lieu à des valeurs estimées de paramètres fortement corrélées. C'est ce que l'on observe dans la figure 3.1 (gauche), où la fonctionnelle de l'erreur est très étalée (comme une « vallée »), exprimant par là que plusieurs combinaisons de paramètres peuvent décrire les mêmes données de manière équivalente.

C'est pourquoi une manière d'étudier l'identifiabilité pratique d'un modèle est de tracer les fonctions de sensibilité. De nombreux exemples d'une telle approche sont disponibles dans la littérature scientifique, en particulier dans le cas du modèle de Monod avec mesures de biomasse et de substrat [HOL 82a, HOL 82b, MAR 89, PET 00c, POS 90, ROB 85, VIA 85].

Différentes approches sont possibles pour la détermination des fonctions de sensibilité $\frac{\partial y_j}{\partial \theta_i}$. La méthode la plus précise est la dérivation analytique. Pour des modèles complexes, il s'avère cependant rapidement nécessaire d'utiliser des logiciels de manipulation symbolique. On peut également avoir recours aux calculs numériques sur base de valeurs numériques des paramètres qui s'écartent légèrement des valeurs nominales. Pratiquement, si l'on considère une déviation $\Delta\theta_i$ pour le paramètre θ_i , la sensibilité de la variable de sortie y_j à θ_i se calcule comme suit :

$$\frac{\partial y_j}{\partial \theta_i} = \frac{y_j(\theta_i) - y_j(\theta_i + \Delta\theta_i)}{\Delta\theta_i} \quad [3.27]$$

A titre d'illustration, faisons le calcul dans le cas du modèle de Monod [3.6] avec mesures de OUR . La sensibilité de OUR par rapport à μ_{max} est égale à :

$$\frac{\partial OUR}{\partial \mu_{max}} = \frac{\partial}{\partial \mu_{max}} \left(-(1 - Y) \frac{dS}{dt} \right) = -(1 - Y) \frac{d}{dt} \left(\frac{\partial S}{\partial \mu_{max}} \right) \quad [3.28]$$

Dans cette équation, la sensibilité $\frac{\partial S}{\partial \mu_{max}}$ est obtenue par intégration (avec zéro comme condition initiale) de l'équation différentielle suivante :

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial S}{\partial \mu_{max}} \right) &= \frac{\partial}{\partial \mu_{max}} \left(-\frac{\mu_{max} X}{Y} \frac{S}{K_S + S} \right) \\ &= -\frac{X}{Y} \left(\frac{S}{K_S + S} + \frac{\mu_{max} K_S}{(K_S + S)^2} \frac{\partial S}{\partial \mu_{max}} \right) \end{aligned} \quad [3.29]$$

où la concentration en substrat S est calculée en intégrant l'équation dynamique du substrat :

$$\frac{dS}{dt} = -\frac{\mu_{max}X}{Y} \frac{S}{K_S + S} \quad [3.30]$$

La résolution simultanée des équations différentielles [3.29] et [3.30] permet d'obtenir les fonctions de sensibilité [3.28]. On peut pratiquer de manière similaire pour la sensibilité de OUR par rapport à K_S . On obtient alors les relations suivantes :

$$\frac{\partial OUR}{\partial K_S} = -(1-Y) \frac{d}{dt} \left(\frac{\partial S}{\partial K_S} \right) \quad [3.31]$$

$$\frac{d}{dt} \left(\frac{\partial S}{\partial K_S} \right) = -\frac{\mu_{max}X}{Y} \left(\frac{K_S}{(K_S + S)^2} \frac{\partial S}{\partial K_S} - \frac{S}{(K_S + S)^2} \right) \quad [3.32]$$

On peut observer, en examinant les relations ci-dessus, que les fonctions de sensibilité dépendent de la valeur des paramètres du modèle. Ceci est une caractéristique des modèles non linéaires [DRA 81]. La matrice de Fisher sera donc aussi fonction de la valeur des paramètres du modèle : cela aura des conséquences sur la conception optimale d'expériences (voir ci-dessous).

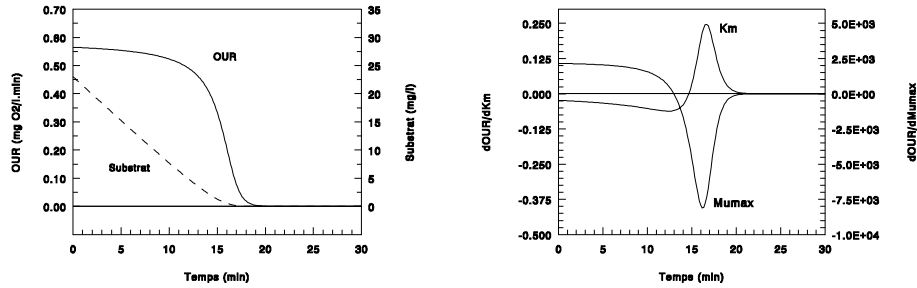


Figure 3.2. Fonctions de sensibilité (droite) pour un profil d' OUR avec un modèle de Monod (gauche)

Un exemple de profil de OUR avec les évolutions des fonctions de sensibilité est représenté à la figure 3.2. On peut observer que les fonctions de sensibilité pour K_S et μ_{max} sont pratiquement linéairement dépendantes. Intuitivement, les fonctions de sensibilité expriment la dépendance d'une variable à un changement de valeur d'un paramètre. Dans l'exemple de la figure 3.2, ces conditions de quasi-linéarité prédominent lorsque la concentration en substrat est tombée à une valeur proche de la constante d'affinité K_S . On comprend facilement qu'une première approche possible pour améliorer le contenu informatif des données est de choisir des instants d'échantillonnage correspondant aux zones où les paramètres ont de l'influence, c'est-à-dire quand la sensibilité est élevée.

3.4. Planification optimale d'expériences en vue de l'estimation des paramètres (POE/EP)

3.4.1. Introduction

Une manière d'améliorer la qualité des données pour l'identification en vue d'obtenir des valeurs de paramètres fiables est de planifier les expériences qui servent à la collecte des données de manière à augmenter le contenu informatif de ces données pour l'identification. Cette tâche est rendue particulièrement difficile dans le contexte des bioprocédés, vu les ressources expérimentales limitées, tant du point de vue du temps d'expériences (celles-ci peuvent s'avérer très longues et s'étendre sur plusieurs semaines, voire plusieurs mois), que des coûts (les mesures peuvent s'avérer rapidement très onéreuses).

Les buts poursuivis par la planification d'expériences dans le contexte de l'identification peuvent être triple [VAN 98] :

- la sélection fiable de structures de modèles mathématiques adéquats pour décrire la dynamique du procédé ;
- l'estimation précise des paramètres du modèle ;
- la combinaison des deux objectifs ci-dessus.

Chacun de ces objectifs est associé à l'une ou l'autre fonction quantitative qui déterminera sur quels aspects devront se concentrer les expériences. Dans les paragraphes suivants, nous nous concentrerons sur la planification d'expériences pour l'estimation des paramètres (deuxième objectif). Le lecteur intéressé pourra utilement consulter le chapitre [VAN 98] pour les aspects planification expérimentale pour les deux autres objectifs.

Un plan d'expérience est typiquement constitué de trois étapes :

- 1) définition d'une fonction objectif de la POE, qui est la traduction mathématique du but poursuivi par la planification expérimentale ;
- 2) inventaire des degrés de liberté disponibles et des contraintes expérimentales ;
- 3) détermination de l'optimum de la fonction objectif en jouant de manière appropriée sur les degrés de liberté, tout en s'assurant de respecter les contraintes.

Fondamentalement, l'algorithme qui réalise la planification d'expériences le fait sous forme de simulation numérique d'expériences, de manière à quantifier les effets potentiels des conditions expérimentales proposées sur la fonction objectif. Sur la base des résultats donnés par un algorithme d'optimisation qui est capable de tenir compte des contraintes expérimentales, on peut proposer une expérience optimale, qui sera ultérieurement mise en œuvre.

Puisque la planification expérimentale est largement dépendante du modèle du procédé, il est évident qu'un prérequis à une bonne planification est l'existence d'un modèle suffisamment fiable du procédé.

Les questions suivantes peuvent servir de guide avant la mise en œuvre de l'étude de planification expérimentale.

Quel mesurer ? Cette question met en évidence l'importance d'une bonne définition du système (choix des variables d'entrée, de sortie et d'état...).

Où mesurer ? C'est le problème du placement des capteurs.

Quand mesurer ? On s'intéresse alors au choix de la (des) période(s) d'échantillonnage.

Quel type de manipulations ? C'est souvent la question la plus importante. Une fois les degrés de liberté déterminés, il faut encore choisir quel type de signal à mettre en œuvre. Cela requiert souvent de l'imagination et de la créativité pour concevoir les signaux qui agiront sur le procédé de manière à récupérer une information de grande qualité. Comme exemples de signaux possibles, on trouvera par exemple les signaux périodiques (sinus...), les impulsions, les signaux binaires pseudo-aléatoires (PRBS *pseudo random binary signals* en anglais) [LUB 91]...

Quel traitement des données ? Bien que, la plupart du temps, non considéré comme étant partie intégrante de la planification expérimentale, le prétraitement des données est souvent appliqué avant que l'identification du modèle ne commence. Typiquement, celui-ci inclut l'élimination du bruit de mesure, la détection et l'élimination de données aberrantes et l'élimination des dynamiques non intéressantes.

La qualité d'un jeu de valeurs estimées de paramètres peut être évaluée de diverses manières, par exemple sur base de la qualité des prédictions données par le modèle. Cependant, dans la plupart des cas, on caractérise la qualité de l'estimation des paramètres sur base des erreurs d'estimations (intervalles de confiance comme données ci-dessus [3.23]) ou, plus généralement, sur base de la matrice de covariance [3.21]. La qualité apparaît dès lors clairement reliée à l'identifiabilité pratique.

Si l'objectif de la planification expérimentale est d'améliorer l'estimation des paramètres, il est évident que la matrice de covariance [3.21], ses éléments, ou son inverse (la matrice de Fisher) en constitue un élément central.

Notons que d'autres méthodes peuvent s'avérer également appropriées pour traiter la planification expérimentale. Par exemple, il se peut que la transformation du modèle en une forme équivalente numériquement plus robuste permette d'obtenir une estimation plus fiable [RAT 86] (cf. également le cas d'étude à la fin du chapitre). On

peut également penser à utiliser une information *a priori* (comme la connaissance du taux de croissance maximum) pour imposer des bornes aux valeurs des paramètres et, de cette manière, améliorer l'identifiabilité des autres paramètres [MUN 89].

3.4.2. Base théorique du POE/EP

Rappelons-nous d'abord que la variable qui décrit la fiabilité de la valeur estimée d'un paramètre, c'est-à-dire la déviation standard de l'erreur d'estimation du paramètre, est donnée par la relation suivante [3.22] :

$$\sigma(\theta_i) = s \sqrt{V_{ii}} \quad [3.33]$$

dans le cas d'une seule variable de sortie mesurée. On observe alors que 2 termes peuvent être manipulés en vue d'accroître la précision de l'estimation du paramètre. La première (le carré du résidu moyen s^2) est calculée à partir de la somme des erreurs au carré entre les N prédictions du modèle et les N données expérimentales, J_{opt} [3.24] :

$$s^2 = \frac{J_{opt}(\theta)}{N - p} \quad [3.34]$$

Il est possible de diminuer cette valeur en augmentant le nombre de données expérimentales N , par exemple en répétant les expériences. Cela est particulièrement utile lorsque p n'est pas négligeable devant N . Quand N est très grand, l'augmentation de la valeur du dénominateur sera proportionnelle à l'augmentation de la valeur de la fonction objectif J_{opt} qui est typiquement une somme de N erreurs au carré.

On peut, de manière alternative, réduire la covariance de l'erreur d'estimation V [3.21]. Nous allons maintenant nous concentrer sur les méthodes qui permettent de diminuer V . Différentes approches ont été développées pour planifier des expériences sur base de la matrice de covariance V , ou de manière équivalente, sur base de la matrice d'information de Fisher F , via l'optimisation de certaines mesures scalaires de ces matrices [MUN 91] :

- critère de planification A-optimale : $\min[\text{tr}(F^{-1})]$,
- critère de planification A-optimale modifié : $\max[\text{tr}(F)]$,
- critère de planification D-optimale : $\max[\det(F)]$,
- critère de planification E-optimale : $\max[\lambda_{\min}(F)]$,
- critère de planification E-optimale modifié : $\min[\frac{\lambda_{\max}(F)}{\lambda_{\min}(F)}]$.

où $\lambda_{\min}(F)$ et $\lambda_{\max}(F)$ sont respectivement la plus petite et la plus grande valeur propre de la matrice d'information de Fisher. Ces différents critères peuvent être interprétés comme suit [MUN 91] : les planifications A et D minimisent respectivement les

moyennes arithmétique et géométrique des erreurs d'identification, le critère E permet alors la minimisation de l'erreur la plus large. Pour les deux critères E, l'optimisation se fait sur les valeurs propres de la matrice de Fisher : cela garantit la maximisation de la distance vis-à-vis du cas singulier (non informatif). Le critère E modifié devrait être interprété sur la base de la forme de la fonction objectif. Le rapport de la plus grande valeur propre à la plus petite est une indication de cette forme. Comme l'objectif est d'avoir des valeurs propres aussi proches que possible les unes des autres, cela revient à avoir un cercle (il faut malgré tout noter que le critère ne vise pas la taille de la fonction objectif, c'est-à-dire le *volume* de l'intervalle de confiance ; en d'autres mots, optimiser le critère E modifié peut très bien aboutir à un intervalle de confiance plus rond mais plus large). Quand $\lambda_{min}(F)$ est égale à zéro, ce rapport est égal à l'infini, c'est-à-dire qu'il existe une infinité de combinaisons de paramètres qui permettent de décrire les données expérimentales, et dès lors l'expérience est non informative. La matrice de Fisher est alors singulière. Dans ce cas, les critères D et E sont égaux à zéro, alors qu'il est impossible de déterminer le critère A, vu que l'inversion de F est impossible. Ce cas permet également de mettre en évidence des problèmes qui peuvent surgir avec le critère A modifié : même dans le cas d'une expérience non informative et donc non identifiable, le critère A modifié peut toujours être maximisé, puisqu'une des autres valeurs propres peut être très élevée [GOO 87].

Il n'est par ailleurs pas inutile de rappeler que les éléments de la matrice d'information de Fisher dépendent des unités des paramètres. En conséquence, les valeurs propres de la matrice de Fisher dépendent des unités : elle peut donc faire l'objet d'une remise à l'échelle. Et il se peut qu'une expérience optimale pour un certain jeu d'unités ne l'est pas pour un autre. Ceci est vrai pour chaque critère ci-dessus, sauf pour le critère D [PET 00a] : bien que la valeur absolue de ce critère soit différente pour différentes unités des paramètres, l'expérience optimale demeure la même. Et seule l'expérience optimale générée par le critère D est indépendante des unités choisies. D'autre part, cette dépendance de la matrice d'information de Fisher vis-à-vis des unités peut être utilisée de manière avantageuse dans le cas où des problèmes numériques surgissent lors de son inversion. Dans ce cas, on peut modifier le nombre de conditionnement (qui est équivalent au critère E modifié) par un simple changement d'unités et ainsi récupérer la fiabilité de l'inversion matricielle. Cela peut s'avérer particulièrement intéressant pour un certain nombre d'algorithmes d'optimisation numérique utilisés pour l'estimation paramétrique (voir ci-dessous). Mentionnons finalement que d'autres critères peuvent être proposés, par exemple la réduction de l'erreur d'estimation d'un paramètre peut être obtenue en utilisant comme critère la variance de ce paramètre.

3.4.3. Exemples

Passons maintenant en revue quelques exemples de planification expérimentale pour l'estimation paramétrique. Vialas *et al.* [VIA 85] ont proposé d'échantillonner

plus fréquemment pour des périodes expérimentales définies. Holmberg [HOL 82a] a montré que l'identifiabilité des paramètres du modèle de Monod dépend fortement, pour les expériences *batch*, de la valeur initiale de la concentration en substrat ; elle écrit également que la concentration initiale optimale en substrat dépend du niveau de bruit et des instants d'échantillonnage. Il est également évident sur base de ses résultats que la planification expérimentale dépend de la valeur des paramètres, ce qui peut induire une modification de la planification des expériences si l'on veut tenir compte du caractère changeant du procédé dans le temps. Munack [MUN 89] a proposé différentes modifications des expériences *batch* et a montré que d'importantes améliorations des intervalles de confiance peuvent être obtenues par le biais des techniques de planification expérimentale.

Vanrolleghem *et al.* (1995) ont mis en évidence l'amélioration de la précision de l'estimation pour les deux paramètres d'un modèle de biodégradation avec cinétique de Monod. L'ajout d'un pulse de substrat (figure 3.3) a permis de réduire l'intervalle de confiance de 25 %. Cette diminution de la variance était répartie de manière équivalente entre les deux paramètres μ_{max} et K_S , ce qui est un résultat typique d'un PEO basé sur le critère D.

Une étude similaire a été réalisée par [PET 00a] pour une application réelle où un modèle de traitement d'eaux usées devait être identifié. Le modèle contenait deux sous-modèles, un pour l'oxydation de l'azote (nitrification), l'autre pour l'oxydation du carbone. Ici également, une amélioration de la précision a été obtenue, en ajoutant à l'échantillon d'eaux usées une quantité déterminée d'ammoniac. L'intervalle de confiance a été réduit de 20 % pour le sous-modèle d'oxydation du carbone et de 50 % pour le sous-modèle de nitrification.

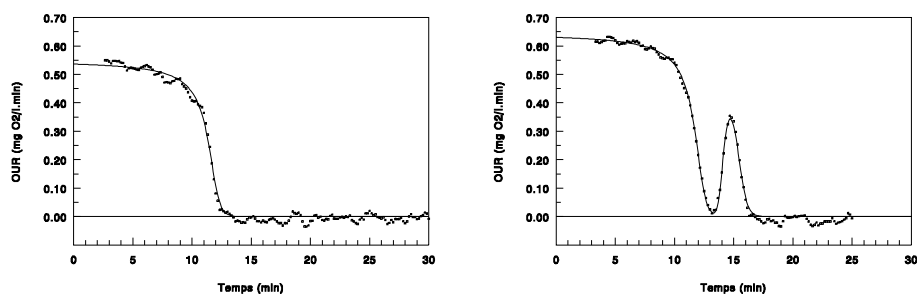


Figure 3.3. Respirogrammes expérimentaux obtenus pour une expérience *batch* et pour une expérience *fed-batch* avec un pulse additionnel après 14,6 minutes. Les variances des paramètres sont de 50 % inférieures pour l'expérience *fed-batch*

Baetens *et al.* (2000) ont identifié six paramètres dans un procédé de déphosphatation biologique. La planification expérimentale a été évaluée en considérant plusieurs

degrés de liberté. Le critère D pouvait être amélioré d'un facteur au moins égal à 4. Cela correspond à une amélioration des intervalles de confiance de l'estimation des paramètres d'un facteur 2. L'évolution du critère E modifié pour des conditions expérimentales changeantes ont indiqué que l'axe le plus long de l'ellipsoïde de confiance pouvait être réduit d'un facteur 2, car l'accroissement du critère E modifié d'un facteur 4 pouvait être obtenu en augmentant le dosage en acétate.

Une autre approche a été développée par Petersen *et al.* [PET 00c] en relation avec les degrés de liberté pour la planification expérimentale. Les différences de précision d'estimation paramétrique ont été évaluées lorsque différents ensembles de variables mesurées sont utilisées pour identifier les paramètres d'un modèle de nitrification. Les possibilités suivantes ont été évaluées :

- une seule mesure d'oxygène dissous S_O dans un réacteur *batch* aéré ;
- deux mesures d'oxygène dissous S_O à l'entrée et à la sortie de la chambre de respiration ;
- le taux de respiration OUR_{ex} calculé à partir de ces deux mesures d'oxygène ;
- le taux de production de protons H_p obtenu à partir du régulateur de pH ;
- deux mesures d'oxygène dissous S_O et le taux de production de protons H_p .

En dépit du fait que dans ce cas les paramètres de transfert de masse K_La et $S_{O,sat}$ doivent être estimés (en plus des trois paramètres de nitrification μ_{max} , K_{NH} et $S_{NH}(0)$), les estimations des paramètres de nitrification peuvent être plus précises avec les mesures d'oxygène dissous. Les intervalles de confiance sont alors dix fois plus petits (les variances sont cent fois plus petites). Le bruit (plus) élevé des mesures de taux de respiration est sans aucun doute une explication partielle de ce phénomène. Ce résultat met en évidence l'importance d'un choix approprié des variables mesurées utilisées pour l'estimation des paramètres.

Munack [MUN 91] avait également montré que l'effet de différentes configurations de mesures pouvait mener à des contenus informatifs significativement différents des jeux de données. Dans son analyse intervient non seulement le type de données, mais également leur nombre et leur position au sein du réacteur.

Versyck *et al.* [VER 97] mettent en évidence un résultat remarquable en ce qui concerne la planification expérimentale, en liaison avec la commande optimale : le profil de commande optimale pour un modèle de croissance avec cinétique de Haldane constitue un excellent point de départ pour trouver une expérience pour lequel le critère E modifié est optimal (égal à 1). Par ailleurs, il est montré dans [VER 00] que, dans le cas de la modélisation de l'effet de la température sur la cinétique, l'unicité du critère E modifié peut être atteint par un profil approprié de température. Les auteurs montrent également que plus d'un plan expérimental permet d'atteindre cet optimum et qu'il est donc possible de tenir compte d'autres critères de sélection, comme la faisabilité pratique ou la validité du modèle (voir aussi [BAL 94]).

3.5. Algorithmes d'estimation

Jusqu'à présent, nous nous sommes concentrés sur l'analyse d'identifiabilité des paramètres et sur la conception d'expériences pour l'estimation paramétrique. Il est temps de passer à la question suivante. Nous avons construit et/ou sélectionné un modèle et nous avons un jeu de données appropriées pour l'estimation (obtenues idéalement après planification expérimentale) : comment allons-nous procéder pour estimer les valeurs des paramètres de ce modèle ?

Comme ce chapitre a comme vocation de donner un aperçu sur l'identification des paramètres, nous avons décidé de ne pas entrer trop dans les détails des algorithmes utilisés pour l'estimation paramétrique. De toute façon, il y a peu de chances que le lecteur soit amené à développer lui-même l'algorithme d'estimation de ces paramètres, il aura plus que certainement recours à des logiciels disponibles sur le marché (dont Matlab et ses boîtes à outils sont des exemples types). Nous suggérons de toute façon au lecteur intéressé par le sujet de se référer à des ouvrages spécialisés, par exemple [WAL 94].

3.5.1. *Choix de deux jeux de données*

La première étape essentielle et fondamentale, une fois les données expérimentales disponibles et avant de passer à l'estimation des paramètres du modèle, est de séparer les données en deux jeux de données distincts :

- 1) un premier jeu de données utilisées pour la calibration des paramètres ;
- 2) le reste des données pour la validation de l'estimation.

Le premier jeu de données va être utilisé pour calculer les estimations des paramètres en exploitant un algorithme d'estimation approprié. Le deuxième va servir à vérifier que ces paramètres permettent de bien reproduire la dynamique du procédé.

L'indépendance entre ces deux jeux de données est fondamentale pour valider la procédure d'identification (on peut espérer que le modèle obtenu sur la base de certaines données les reproduise bien, il est beaucoup plus important de savoir si ce modèle reproduit bien d'autres données qui n'ont pas servi à calibrer ses paramètres). La séparation des données doit se faire de manière à ce que le premier jeu de données (calibration) soit suffisamment informatif et couvre donc un spectre suffisant de conditions expérimentales (suivant par exemple un plan d'expérience). Le deuxième jeu de données devra contenir suffisamment de données, de manière à rendre la validation aussi crédible que possible. On évitera d'avoir des jeux de données trop déséquilibrés en termes de nombre de données, même si souvent le premier jeu de données (calibration) en comprendra souvent plus que le deuxième (validation).

3.5.2. Eléments d'estimation paramétrique : l'estimation par moindres carrés dans le cas linéaire

Nous avons écrit dans la section 3.3 que l'estimation des paramètres consistait à la base à optimiser un certain critère d'erreur, typiquement de prédiction, comme la fonction objectif [3.18]. La forme concrète que prendra l'algorithme d'estimation dépendra largement du modèle, en particulier suivant que le modèle est ou non linéaire en les paramètres. Par exemple, dans le cas linéaire à une équation, le modèle s'écrira de manière générique sous la forme suivante :

$$y_i = \phi_i^T \theta \quad [3.35]$$

où θ est le vecteur des paramètres, ϕ_i est appelé le « régresseur », y_i contient les termes indépendants des paramètres et l'indice i correspond à l'indice temporel. Le choix d'une représentation du modèle en temps discret est naturel, dans la mesure où l'estimation se base sur des données expérimentales elles-mêmes disponibles sous forme échantillonnée dans le temps.

A titre d'illustration, considérons l'équation de bilan d'un substrat S soumis à deux réactions : une réaction de croissance (où, de plus, on suppose que la dépendance de la cinétique vis-à-vis de la concentration en substrat est du premier ordre) et une réaction de maintenance. Cette équation s'écrit comme suit :

$$\frac{dS}{dt} = -k_1 \alpha S X - k_m X + D S_{in} - D S \quad [3.36]$$

Faisons l'hypothèse que ni α (constante cinétique de croissance), ni k_m (coefficient de maintenance) ne sont connus et que nous disposons de données de concentration en substrat S et en biomasse X , ainsi que du taux de dilution D et de la concentration en substrat dans l'alimentation S_{in} à intervalles de temps T (autrement dit, suivant une période d'échantillonnage T). Supposons également que le coefficient de rendement k_1 est connu. Le vecteur θ est alors égal à :

$$\theta = \begin{bmatrix} \alpha \\ k_m \end{bmatrix} \quad [3.37]$$

En approximant la dérivée temporelle $\frac{dS}{dt}$ par une différence finie $\frac{S_{i+1} - S_i}{T}$ (où l'indice i indique la valeur à l'instant $t = iT$), on peut réécrire l'équation de bilan [3.36] comme suit :

$$S_{i+1} - S_i - T D_i S_{in,i} + T D_i S_i = -k_1 S_i X_i T \alpha + X_i T k_m \quad [3.38]$$

Autrement dit, dans le formalisme de l'équation [3.35], y_i et ϕ_i sont égaux à :

$$y_i = S_{i+1} - S_i - T D_i S_{in,i} + T D_i S_i, \quad \phi_i = \begin{bmatrix} -k_1 S_i X_i T \\ X_i T \end{bmatrix} \quad [3.39]$$

Dans le cas simple de l'équation linéaire en les paramètres θ , des techniques de régression linéaire, par exemple un estimateur par moindres carrés, peuvent être utilisées. Dans ce cas, on considère la fonction objectif [3.18] sous la forme suivante :

$$J(\theta) = \sum_{i=1}^N \beta_i (y_i(\hat{\theta}) - y_i)^2 \quad [3.40]$$

L'estimateur par moindres carrés qui minimise ce critère s'écrit simplement :

$$\hat{\theta} = \left[\sum_{i=1}^N \beta_i \phi_i \phi_i^T \right]^{-1} \sum_{i=1}^N \beta_i \phi_i y_i \quad [3.41]$$

Lorsque β_i dans la fonction objectif est différent de 1, on parle d'estimateur par moindres carrés pondérés. Un choix possible pour β_i est de prendre l'inverse de la variance de la mesure y_i .

3.5.3. *Survol des algorithmes d'estimation de paramètres*

De manière générale, l'estimation des paramètres suit la logique décrite ci-dessus pour le cas simple du modèle linéaire monovarié. Elle est basée sur la minimisation d'un critère, dont le modèle type est celui déjà vu plus haut [3.18] ¹ :

$$J(\theta) = \sum_{i=1}^N (y_i(\hat{\theta}) - y_i)^T Q_i (y_i(\hat{\theta}) - y_i) \quad [3.42]$$

Les poids Q_i peuvent être différents à chaque instant d'échantillonnage, ce qui peut s'avérer intéressant dans le cas d'une qualité des mesures qui varie dans le temps. Un choix type pour la matrice Q_i consiste à prendre l'inverse de la matrice de covariance des erreurs de mesure (comme suggéré ci-dessus dans le cas monovarié).

Avant de passer en revue les méthodes de minimisation qui permettent de déterminer les valeurs de paramètres de la fonction objectif choisie, on peut dire que, de manière générale, les problèmes rencontrés lors de l'estimation de paramètres par des méthodes de type moindres carrés sont dus à trois sortes de phénomènes. D'abord, il se peut que les résidus ne soient pas distribués de manière normale, en contradiction avec les hypothèses à la base de la conception de ce type d'algorithmes. Ensuite, pour pondérer correctement les résidus, il faut connaître la valeur de la variance de l'erreur en chaque point de mesure. Celle-ci n'est disponible de manière précise que

1. Mais d'autres fonctions objectives sont parfois utilisées, basées par exemple sur la minimisation de la valeur absolue de l'erreur de prédiction, sur la minimisation de l'erreur absolue maximum, ou sur la maximisation du nombre de changements de signe dans la séquence des résidus.

lorsqu'il y a un nombre suffisant de mesures répétées en un point, ce qui est rarement le cas. Finalement, les valeurs aberrantes de mesures peuvent mener à une estimation imprécise des paramètres, ainsi qu'à des intervalles de confiance biaisés.

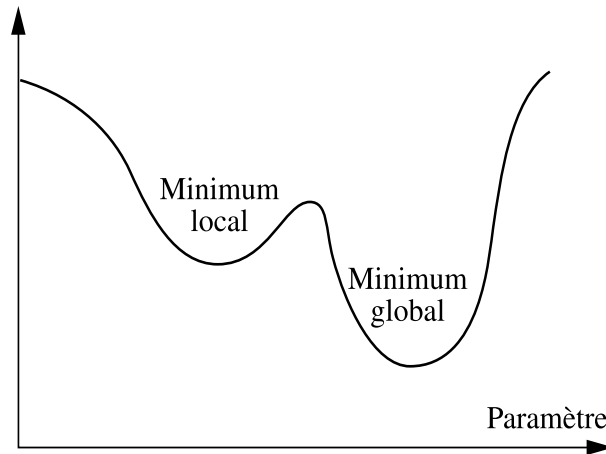


Figure 3.4. Minima local et global pour une fonction objectif

La recherche d'un optimum d'une fonction multivariée est un problème commun dans de nombreux domaines de recherche et il existe une expertise substantielle à ce sujet. Il faut cependant attirer l'attention sur le fait que le minimum que l'on cherche (le jeu de paramètres qui minimise le critère) peut être soit local, soit global (voir figure 3.4) : dans le premier cas, il correspond à un minimum dans une région limitée de l'espace, alors que dans le second cas, il correspond au minimum sur la totalité de l'espace des paramètres. En dépit d'efforts importants, il n'existe à ce jour aucun algorithme de minimisation qui puisse garantir de trouver le minimum global pour n'importe quel problème d'optimisation non linéaire. On constate en particulier que la convergence vers un minimum (local ou global) dépend fortement des valeurs initiales choisies pour les paramètres $\hat{\theta}(0)$. On peut aisément comprendre ce phénomène sur base de la courbe 3.4 : en prenant des valeurs initiales à gauche du minimum local, il est probable que l'algorithme de minimisation ira se fixer sur le minimum local et n'atteindra pas le minimum global, alors qu'il l'atteindra en partant à droite de celui-ci. Pour éviter de ne converger que vers un minimum local, il peut donc s'avérer important de choisir un nombre suffisant d'itérations de l'algorithme de minimisation en prenant des jeux de valeurs initiales de paramètres suffisamment différents pour couvrir un spectre de valeurs aussi large que possible.

Schuetze [SCH 98] a tenté de classer les différentes méthodes de minimisation en deux groupes : les méthodes de minimisation locale et celles de minimisation glo-

bale, même si cela a quelque chose d'artificiel, dans la mesure où les méthodes dites « globales » font toutes intervenir à un certain niveau une procédure locale.

3.5.3.1. *Minimisation locale*

Il existe un grand nombre de procédures de minimisation locale. Nombre d'entre elles se basent sur les dérivées de la fonction objectif à minimiser, information qui est alors soit considérée comme directement disponible, soit calculée numériquement. L'idée de base de ces méthodes dites « du gradient » est qu'une direction de recherche du minimum est recherchée dans l'espace des paramètres. Une fois cette direction choisie, l'algorithme choisit l'amplitude avec laquelle les valeurs des paramètres sont modifiées. Si l'on peut montrer la convergence de la méthode, la convergence peut s'avérer très lente [DRA 81].

Une alternative consiste à utiliser la méthode de Newton. Si elle permet d'avoir éventuellement une convergence plus rapide, un désavantage de la méthode est qu'elle peut diverger ou osciller autour de la solution.

La modification qui est probablement la plus connue des méthodes de base est l'algorithme de Levenberg-Marquardt [MAR 63], un compromis entre la méthode du gradient et celle de Newton. Un choix approprié des paramètres de l'algorithme permet de combiner les avantages du gradient (propriétés de convergence) avec ceux de la méthode de Newton (convergence rapide).

Pour améliorer la convergence, on peut utiliser le fait que la plupart des fonctions sont raisonnablement bien approximées par une fonction quadratique près de l'optimum. La méthode de BFGS (Broyden, Fletcher, Goldfarb et Shanno) [FLE 87] est probablement la plus connue de cette classe de méthodes. [VAN 96] a constaté cependant que ces méthodes, si elles convergent très rapidement, peuvent être sensibles aux minima locaux.

Une méthode de minimisation locale qui n'est pas basée sur le calcul ou l'approximation des dérivées est la méthode de Powell avec les améliorations proposées par Brent [GEG 92, PRE 86]. Elle est basée sur une combinaison répétée d'une recherche unidimensionnelle le long d'un ensemble de plusieurs directions. Il s'agit probablement d'une des meilleures méthodes en termes de compromis entre vitesse de convergence et insensibilité aux minima locaux.

Une autre méthode bien connue de minimisation locale qui ne requiert pas d'information sur les dérivées a été proposée par Nelder et Mead, mieux connue sous le nom de méthode du Simplex (à ne pas confondre avec la méthode du simplexe en programmation linéaire !). Un simplexe est un ensemble de $(p + 1)$ vecteurs de paramètres dans l'espace des paramètres. La méthode consiste en une succession de comparaisons de la fonction objectif à ses arêtes, suivie d'un remplacement des arêtes ayant la valeur

la plus élevée par un autre point dans l'espace des paramètres de dimension p . La méthode est bien appréciée pour sa robustesse aux minima locaux, sa facilité de mise en œuvre et sa vitesse de convergence raisonnable [PRE 86, SCH 98, VAN 96].

3.5.3.2. Minimisation globale

Les méthodes de minimisation globale peuvent *grosso modo* être classées en deux groupes [SCH 98]. Le premier groupe comprend des méthodes purement déterministes, comme la méthode des grilles (*gridding*). Celle-ci consiste à évaluer la fonction objectif pour un nombre élevé de points prédéfinis sur une grille couvrant l'espace des paramètres. Si le nombre d'évaluations de fonctions est suffisant, il y a des chances d'atteindre les minima. Cette méthode n'est pas très efficace, sauf à l'améliorer en raffinant la grille après une série d'évaluations.

Le second groupe de méthodes globales peut être appelé méthodes aléatoires, car des décisions aléatoires sont incluses dans la procédure de recherche de l'optimum. Parmi celles-ci, les méthodes adaptatives tiennent compte de l'information obtenue lors des évaluations précédentes. Pour l'une de celles-ci (*Simulated Annealing* [PRE 86]), l'idée est que la recherche ne va pas toujours vers une solution possible (qui pourrait n'être qu'un minimum local), mais aussi, de temps en temps, dans une direction différente. Cette méthode peut être vue comme précédant des méthodes populaires tels que les algorithmes génétiques (GA) [GOL 89]. Ces algorithmes commencent avec une population initiale de solutions candidates (un peu similaires aux arêtes dans la méthode du Simplex) échantillonnées de manière aléatoire dans l'espace des paramètres. Dans les algorithmes génétiques, de nouvelles solutions candidates sont obtenues par imitation des processus d'évolution biologique de croisement, mutation et sélection parmi les « populations » de paramètres. La définition des paramètres de l'algorithme lui-même est cruciale pour une mise en œuvre correcte.

3.6. Un cas d'étude : l'identification des paramètres d'un modèle de procédé de digestion anaérobie

Le modèle de digestion anaérobie présenté chapitre 2 (formules [2.16] et [2.29]) a fait l'objet d'une étude systématique d'identification des paramètres sur base d'un réacteur à lit fixe du LBE de l'INRA à Narbonne (voir [BER 00, BER 01] pour une étude plus fouillée). Ce cas d'étude est remarquable sous plusieurs aspects. D'abord, l'identification des paramètres est truffée d'embûches typiques pour les systèmes biologiques : le procédé est très complexe (de nombreuses populations bactériennes participent au procédé, elles peuvent avoir des comportements différents selon les conditions opératoires) ; il n'existe aucune mesure directe de chacune des populations bactériennes acidogènes et méthanigènes et, de manière générale, un nombre limité de variables du procédé est accessible ; le processus est lent et peut se déstabiliser facilement par accumulation d'acides gras. Ces caractéristiques ont des conséquences importantes sur le modèle choisi ; le modèle ne peut être de complexité trop élevée

de peur d'être non identifiable, mais les hypothèses simplificatrices de modélisation peuvent avoir un impact sur sa capacité de prédiction de la dynamique du procédé. Le choix de la structure est donc une étape critique, en ce sens que le modèle se doit de contenir les éléments essentiels de la dynamique du procédé.

Par ailleurs, l'identification elle-même contient des éléments originaux par rapport au reste du chapitre. Les contraintes sur le procédé ont amené à concevoir un plan expérimental, non pas sur la base des techniques ci-dessus, mais motivée par le souci de couvrir une gamme de conditions opératoires aussi large que possible, tout en limitant, autant que possible, la durée (forcément longue ici) des expériences.

De plus, la structure des modèles de systèmes réactionnels permet d'opérer une séparation des paramètres en trois classes (coefficients de rendement, paramètres cinétiques et paramètres de transfert) et d'effectuer l'identification de chaque classe de paramètres de manière distincte.

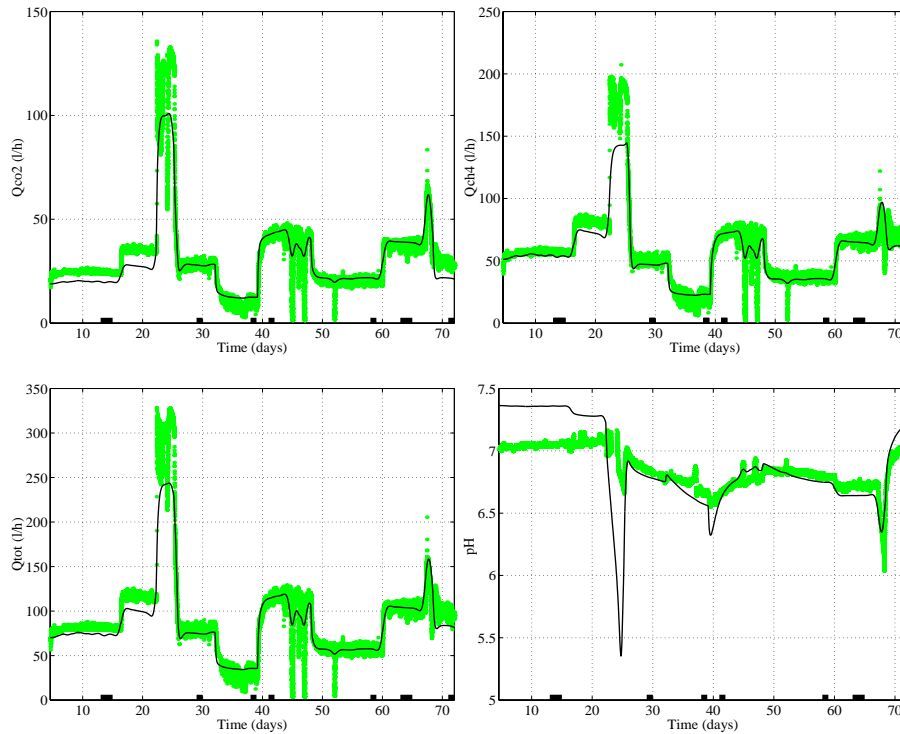


Figure 3.5. Valeurs simulées (trait fin) et données (trait épais) des débits gazeux et du pH. Les périodes utilisées pour la calibration sont représentées en gras sur l'axe des abscisses

Dans le présent cas d'étude, l'accent est mis sur l'estimation paramétrique par régression linéaire. Ceci dit, il est utile d'attirer l'attention sur le fait que cette approche peut ne pas s'avérer optimale d'un point de vue statistique, dans la mesure où les conditions statistiques sur les variables utilisées dans la régression linéaire pourraient ne pas être remplies complètement pour permettre une estimation qui soit statistiquement suffisamment précise et fiable. De manière générale, il faut être prudent dans l'interprétation des valeurs estimées des paramètres et de leur écart type. Il est d'ailleurs souvent conseillé d'utiliser ces valeurs de paramètres données par régression linéaire comme valeurs initiales d'une estimation sur base du modèle non linéaire. Une telle estimation ne s'est cependant pas avérée utile dans le cas présent.

3.6.1. Le modèle

Repartons des équations développées au chapitre 2. Une variante de celles-ci a été considérée, car le réacteur est un réacteur à lit fixe où les bactéries sont fixées sur des supports. Donc formellement, il n'y a pas de terme de dilution dans les équations de bilan. Par contre, il a été constaté qu'une partie de la biomasse se détachait : l'hypothèse que le taux de détachement était proportionnel au taux de dilution, moyennant un coefficient α , a alors été faite. Dans ces conditions, le modèle se récrit comme suit :

$$\frac{dX_1}{dt} = \mu_1 X_1 - \alpha D X_1 \quad [3.43]$$

$$\frac{dX_2}{dt} = \mu_2 X_2 - \alpha D X_2 \quad [3.44]$$

$$\frac{dS_1}{dt} = D(S_{1in} - S_1) - k_1 \mu_1 X_1 \quad [3.45]$$

$$\frac{dS_2}{dt} = D(S_{2in} - S_2) + k_2 \mu_1 X_1 - k_3 \mu_2 X_2 \quad [3.46]$$

$$\frac{dZ}{dt} = D(Z_{in} - Z) \quad [3.47]$$

$$\frac{dC}{dt} = D(C_{in} - C) - q_C + k_4 \mu_1 X_1 + k_5 \mu_2 X_2 \quad [3.48]$$

$$q_C = k_L a (C - S_2 - Z - K_H P_C) \quad [3.49]$$

$$P_C = \frac{\phi - \sqrt{\phi^2 - 4K_H P_T (C + S_2 - Z)}}{2K_H} \quad [3.50]$$

$$\phi = C + S_2 - Z + K_H P_T + \frac{k_6}{k_L a} \mu_2 X_2 \quad [3.51]$$

$$q_M = k_6 \mu_2 X_2 \quad [3.52]$$

$$pH = -\log_{10} \left(K_b \frac{C - Z + S_2}{Z - S_2} \right) \quad [3.53]$$

De plus, des modèles de croissance ont été choisis, un modèle de Monod pour l'acidogénèse, un modèle de Haldane pour la méthanisation :

$$\mu_1 = \frac{\mu_{max1} S_1}{K_{S1} + S_1}, \quad \mu_2 = \frac{\mu_0 S_2}{K_{S2} + S_2 + S_2^2/K_{I2}} \quad [3.54]$$

En l'absence de règle systématique, ce choix a été dicté par une volonté d'avoir des modèles cinétiques suffisamment simples, cohérents avec ceux utilisés habituellement pour la digestion anaérobie, mais aussi capables de mettre en évidence l'instabilité potentielle du procédé en présence d'accumulation d'acides gras volatils (d'où le choix du modèle de Haldane pour μ_2).

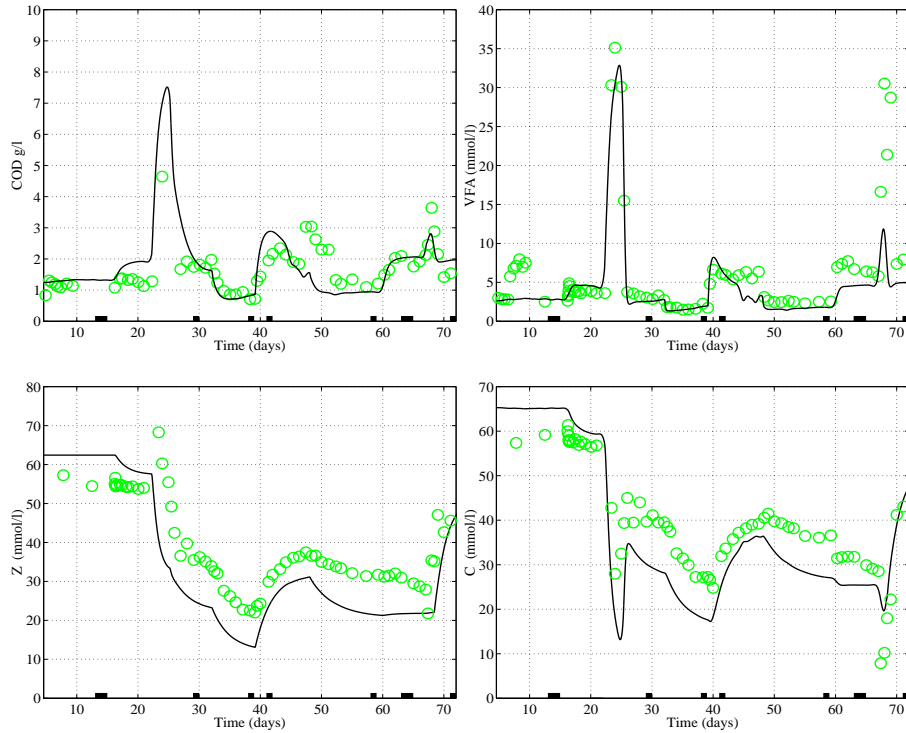


Figure 3.6. Valeurs simulées (trait continu) et données (o) de S_1 , S_2 , Z et C . Les périodes utilisées pour la calibration sont représentées en gras sur l'axe des abscisses

Sachant que K_b et K_H sont des constantes chimique et physique connues ($K_b = 6,5 \cdot 10^{-7}$ mol/l, $K_H = 16$ mmol/l/atm), on constate que le modèle contient treize paramètres à identifier. Par ailleurs, les variables disponibles à la mesure sont le taux de dilution D , les concentrations d'entrées, S_{1in} , S_{2in} , Z_{in} , C_{in} , les débits gazeux q_C , q_M et les concentrations S_1 , S_2 , Z , C et le pH.

3.6.2. Le plan d'expérience

Etant donné la complexité du modèle et du nombre élevé de paramètres, ainsi que les contraintes expérimentales (constantes de temps longues et instabilité potentielle du procédé), la stratégie suivie ici pour la planification expérimentale a été de parcourir un nombre de points d'opérations suffisamment représentatifs du fonctionnement du procédé. Le plan expérimental est résumé dans le tableau 3.1.

D (jour⁻¹)	S_{1in} (g/l)	S_{2in} (mmole/l)	pH
0,34	9,5	93,60	5,12
0,35	10,0	73,68	4,46
0,35	4,8	38,06	4,49
0,36	15,6	112,70	4,42
0,26	10,6	72,98	4,42
0,51	10,7	71,60	4,47
0,53	9,1	68,78	5,30

Tableau 3.1. *Caractéristiques des conditions moyennes d'alimentation*

3.6.3. Le choix des données pour la calibration et pour la validation

Dans le cas présent, une des priorités était d'obtenir un modèle qui soit capable de reproduire prioritairement les états d'équilibre correctement. C'est pourquoi les données recueillies ont été séparées en deux lots : les données correspondant aux états stationnaires pour la calibration des paramètres, les données dynamiques pour la validation. Les périodes correspondant aux valeurs utilisées pour la calibration sont représentées sur les figures 3.5 et 3.6 par un trait gras sur l'axe des abscisses.

3.6.4. L'identification des paramètres

Avec un modèle aussi complexe, on doit s'attendre à rencontrer des problèmes d'identifiabilité structurelle et pratique. Nous espérons pouvoir gommer, au moins partiellement, les problèmes d'identifiabilité pratique grâce à la planification expérimentale. Mais qu'en est-il de l'identifiabilité structurelle ? Sans préjuger des résultats d'une analyse qui pourrait s'avérer complexe, il paraît intéressant de profiter de la structure du modèle et de ses propriétés pour formuler le problème de l'identification des paramètres sous une forme différente. Nous verrons au chapitre suivant (Estimation d'état) qu'il existe une transformation d'état [4.42]² qui permet de récrire

2. Cette transformation est présentée en détail et sous une forme plus générale dans [BAS 90].

une partie du modèle sous une forme indépendante de la cinétique. Cette transformation est dans notre cas doublement intéressante. D'une part, elle permet de procéder à l'identification des paramètres cinétiques et des autres paramètres du modèle de manière distincte. D'autre part, cette séparation est particulièrement intéressante dans la mesure où la modélisation de la cinétique est le maillon le plus faible de celle-ci, étant donné que l'on manque de base physique pour le choix d'un modèle approprié et que de plus, celui-ci est la source principale de non-linéarités dans le modèle.

C'est fondamentalement cette approche qui a été suivie ici. Comme l'identification se fait sur la base des données de régime, des équations de bilan statique des biomasses X_1 et X_2 et des expressions des taux spécifiques de croissance, on en déduit les expressions suivantes :

$$\frac{1}{D} = \frac{\alpha}{\mu_{1 \max}} + K_{S1} \frac{\alpha}{\mu_{1 \max}} \frac{1}{\bar{S}_1} \quad [3.55]$$

$$\frac{1}{D} = \frac{\alpha}{\mu_0} + K_{S2} \frac{\alpha}{\mu_0} \frac{1}{\bar{S}_2} + \frac{1}{K_{I2}} \frac{\alpha}{\mu_0} \bar{S}_2 \quad [3.56]$$

Ces équations sont linéaires en les « paramètres » $\frac{\alpha}{\mu_{1 \max}}$, $K_{S1} \frac{\alpha}{\mu_{1 \max}}$, $\frac{\alpha}{\mu_0}$, $K_{S2} \frac{\alpha}{\mu_0}$ et $\frac{1}{K_{I2}} \frac{\alpha}{\mu_0}$. Ceux-ci peuvent être déterminés par régression linéaire. Le seul problème est qu'il n'est pas possible de distinguer entre $\mu_{1 \max}$ et α : c'est pourquoi une valeur de la littérature [GHO 74] a été considérée pour $\mu_{1 \max}$. Ce choix s'avère d'autant plus acceptable qu'une analyse de sensibilité montre une faible sensibilité de $\mu_{1 \max}$.

On peut ensuite utiliser l'équation de débit de CO_2 pour déterminer $k_L a$, en se rappelant que $q_C = k_L a (CO_2 - K_H P_C)$ et que la réaction de dissociation du bicarbonate permet de relier le CO_2 à C et au pH suivant la relation :

$$CO_2 = C \frac{1}{1 + K_b 10^{pH}} \quad [3.57]$$

Cela nous permet de déterminer les paramètres cinétiques et de transfert indépendamment des coefficients de rendement. Les résultats sont résumés dans le tableau 3.2.

Il nous reste maintenant à calculer les valeurs des coefficients de rendement. La difficulté principale provient de l'absence de mesures des populations X_1 et X_2 . Sans ces mesures, il s'avère que les coefficients de rendement ne sont pas structurellement identifiables. Il suffit pour s'en convaincre de faire une transformation d'échelle pour X_1 et X_2 : $X'_1 = \lambda_1 X_1$, $X'_2 = \lambda_2 X_2$. Cette transformation d'échelle peut être compensée par la remise à l'échelle suivante pour les coefficients de rendement :

$$k'_1 = \frac{k_1}{\lambda_1}, k'_2 = \frac{k_2}{\lambda_1}, k'_4 = \frac{k_4}{\lambda_1}, k'_3 = \frac{k_3}{\lambda_2}, k'_5 = \frac{k_5}{\lambda_2}, k'_6 = \frac{k_6}{\lambda_2} \quad [3.58]$$

Pour contourner la difficulté, l'identification s'est faite en deux étapes : d'abord estimation des rapports de coefficients de rendement structurellement identifiables

Paramètre	Unités	Valeur	Ecart type
$\mu_{1 \max}$	jour ⁻¹	1,20	[GHO 74]
K_{S1}	g/l	7,10	5,0
μ_0	jour ⁻¹	0,74	0,9
K_{S2}	mmol	9,28	13,7
K_{I2}	mmol	16,00	17,9
α	/	0,50	0,4
$k_L a$	jour ⁻¹	19,80	3,5

Tableau 3.2. Valeurs estimées des paramètres cinétiques et de transfert

(k_2/k_1 , k_6/k_3 , k_5/k_3 , k_4/k_1), ensuite, utilisation de données de solides volatils en suspension (SVS) pour déterminer les valeurs de chaque coefficient de rendement. L'estimation des rapports de coefficient de rendement s'est faite sur la base des équations de bilan en régime permanent, qui peuvent se récrire comme suit après élimination de X_1 et X_2 :

$$q_C = D(C_{in} - C) + \left(\frac{k_4}{k_1} + \frac{k_2 k_5}{k_1 k_3} \right) D(S_{1in} - S_1) + \frac{k_5}{k_3} D(S_{2in} - S_2)$$

$$q_M = \frac{k_2 k_6}{k_1 k_3} D(S_{1in} - S_1) + \frac{k_6}{k_3} D(S_{2in} - S_2)$$

Les rapports $\frac{k_4}{k_1} + \frac{k_2 k_5}{k_1 k_3}$, $\frac{k_5}{k_3}$, $\frac{k_2 k_6}{k_1 k_3}$, $\frac{k_6}{k_3}$ peuvent alors être déterminés par régression linéaire. Le résultat de la calibration est présenté au tableau 3.3.

Rapport	Unités	Valeur	Ecart type
k_2/k_1	mmol/g	2,72	2,16
k_6/k_3	/	1,62	0,12
k_5/k_3	/	1,28	0,13
k_4/k_1	mmol/g	1,18	3,02

Tableau 3.3. Valeurs estimées des rapports de coefficients de rendement

La détermination des coefficients s'est alors faite sur la base des mesures de SVS. En effet, les SVS sont un indicateur approximatif de $X_1 + X_2$. La détermination de la répartition des SVS entre les deux populations bactériennes s'est faite en considérant un rapport (= 0,2) entre les bactéries acidogènes X_1 et la biomasse totale $X_1 + X_2$, tiré de la littérature [SAN 94]. Les résultats sont présentés au tableau 3.4.

Paramètre	Unités	Valeur	Ecart type
k_1	g S_1 / g X_1	42,14	18,94
k_2	mmol S_2 / g X_1	116,50	113,60
k_3	mmol S_2 / g X_2	268,00	52,31
k_4	mmol CO_2 / g X_1	50,60	143,60
k_5	mmol CO_2 / g X_2	343,60	75,80
k_6	mmol CH_4 / g X_2	453,00	90,90

Tableau 3.4. Valeurs estimées des coefficients de rendement

3.6.5. L'analyse des résultats

La première étape essentielle après la calibration des paramètres est leur validation sur le jeu de données non utilisées pour la calibration. Celle-ci peut être visualisée sur les figures 3.5 et 3.6. On notera globalement une bonne adéquation entre le modèle et les données expérimentales en transitoire. Cela démontre l'intérêt d'une approche systématique pour l'identification des paramètres, encore plus dans un cas aussi difficile et complexe que celui-ci.

Mais il reste à évaluer la confiance que l'on peut avoir dans la valeur estimée des paramètres. Celle-ci correspond à l'écart type de chaque paramètre reprise dans les tableaux 3.2, 3.3 et 3.4. On notera en particulier l'augmentation de cette valeur lors du passage des rapports de coefficients de rendement à leur valeur individuelle. Cela est le reflet de la large incertitude liée à la détermination des concentrations en biomasse sur base des données de SVS et du rapport supposé de leur distribution. On notera aussi l'incertitude plutôt élevée des paramètres cinétiques, ce qui n'est pas tout à fait surprenant, étant donné le fait que les structures choisies sont heuristiques. Finalement, dans les rapports de coefficients de rendement, on note une plus mauvaise confiance pour les paramètres faisant intervenir k_1 : probablement un effet du caractère variable de la concentration d'alimentation S_{1in} .

3.7. Bibliographie

- [ABO 78] ABORHEY S., WILLIAMSON D., « State and parameter estimation of microbial growth processes », *Automatica*, vol. 14, p. 493-498, 1978.
- [BAL 94] BALTES M., SCHNEIDER R., STURM C., REUSS M., « Optimal experimental design for parameter estimation in unstructured growth models », *Biotechnol. Prog.*, vol. 10, p. 480-488, 1994.
- [BAS 90] BASTIN G., DOCHAIN D., *On-line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, 1990.

- [BEA 60] BEALE E.M.L., « Confidence regions in non-linear estimation », *J. Roy. Stat. Soc.*, vol. B22, p. 41-88, 1960.
- [BER 00] BERNARD O., HADJ-SADOK Z., DOCHAIN D., « Advanced monitoring and control of anaerobic treatment plants: II. Dynamical model development and identification », dans *Proc. Watermatex 2000*, p. 3.57-3.64, 2000.
- [BER 01] BERNARD O., HADJ-SADOK Z., DOCHAIN D., GENOVESI A., STEYER J.P., « Dynamical model development and parameter identification for an anaerobic wastewater treatment process », 2001 (soumis pour publication).
- [BOU 98] BOURREL S.V., BABARY J.P., JULIEN S., NIHTILÄ M.T., DOCHAIN D., « Modelling and identification of a fixed-bed denitrification bioreactor », *System Analysis Modelling Simulation (SAMS)*, vol. 30, p. 289-309, 1998.
- [CAS 85] CASTI J.L., *Non linear system theory*, Academic Press, Mathematics in Science and Engineering, vol. 175, Orlando, 1985.
- [DOC 95] DOCHAIN D., VANROLLEGHEM P.A., VAN DAELE M., « Structural identifiability of biokinetic models of activated sludge respiration », *Wat. Res.*, vol. 29, p. 2571-2579, 1995.
- [DRA 81] DRAPER N.R., SMITH H., *Applied Regression Analysis*, John Wiley, New York, 1981.
- [FLE 87] FLETCHER R., *Practical Methods of Optimization*, John Wiley, New York, 1987.
- [GEG 92] GEGENFURTNER K.R., « PRAXIS: Brent's algorithm for function minimization », *Behaviour Research Methods, Instruments, and Computers*, vol. 24, p. 560-564, 1992.
- [GHO 74] GHOSH S., POHLAND F., « Kinetics of substrate assimilation and product formation in anaerobic digestion », *J. Wat. Pollut. Control Fed.*, vol. 46, p. 748-759, 1974.
- [GOD 85] GODFREY K.R., DI STEFANO J.J. III, « Identifiability of model parameters », dans *Identification and System Parameter Estimation*, p. 89-114, Pergamon Press, Oxford, 1985.
- [GOL 89] GOLDBERG D.E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Menlo Park, Californie, 1989.
- [GOO 87] GOODWIN G.C., « Identification: Experiment design », dans Singh M. (dir.), *Systems and Control Encyclopedia*, vol. 4, p. 2257-2264, Pergamon Press, Oxford, 1987.
- [HOL 82a] HOLMBERG A., « On the practical identifiability of microbial growth models incorporating Michaelis-Menten type nonlinearities », *Math. Biosci.*, vol. 62, p. 23-43, 1982.

- [HOL 82b] HOLMBERG A., RANTA J., « Procedures for parameter and state estimation of microbial growth process models », *Automatica*, vol. 18, p. 181-193, 1982.
- [JEP 96] JEPSSON U., Modelling aspects of wastewater treatment processes, Thèse de doctorat, Dep. Ind. Elec. Eng. Aut., Lund Inst. Technol., Suède, 1996.
- [JUL 97] JULIEN S., Modélisation et estimation pour le contrôle d'un procédé boues activées éliminant l'azote des eaux résiduaires urbaines, Thèse de doctorat, LASS-CNRS, Toulouse, Rapport LAAS n° 97499, 1997.
- [JUL 98] JULIEN S., BABARY J.P., LESSARD P., « Theoretical and practical identifiability of a reduced order model in an activated sludge process doing nitrification and denitrification », *Wat. Sci. Tech.*, vol. 37, n° 12, p. 309-316, 1998.
- [LJU 87] LJUNG L., *System Identification—Theory for the User*, Prentice Hall, Englewood Cliffs, New Jersey, 1987.
- [LOB 91] LOBRY J.R., FLANDROIS J.P., « Comparison of estimates of Monod's growth model parameters from the same data set », *Binary*, vol. 3, p. 20-23, 1991.
- [LUB 91] LÜBBERT A., « Characterization of bioreactors », dans Schügerl K. (dir.), *Biotechnology: A Multi-volume Comprehensive Treatise*, vol. 4, *Measuring, Modelling, and Control*, p. 107-148, VCH, Weinheim, Allemagne, 1991.
- [MAR 63] MARQUARDT D.W., « An algorithm for least squares estimation of non-linear parameters », *J. Soc. Ind. Appl. Math.*, vol. 11, p. 431-441, 1963.
- [MAR 89] MARSILI-LIBELLI S., « Modelling, identification, and control of the activated sludge process », *Adv. Biochem. Eng. Biotechnol.*, vol. 38, p. 90-148, 1989.
- [MUN 89] MUNACK A., « Optimal feeding strategy for identification of Monod-type models by fed-batch experiments », dans Fish N., Fox R., Thornhill N. (dir.), *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*, p. 195-204, Elsevier, Amsterdam, 1989.
- [MUN 91] MUNACK A., « Optimization of the sampling », dans Schügerl K. (dir.), *Biotechnology: A Multi-volume Comprehensive Treatise*, vol. 4, *Measuring, Modelling, and Control*, p. 251-264, VCH, Weinheim, Allemagne, 1991.
- [PET 00a] PETERSEN B., Calibration, identifiability and optimal experimental design of activated sludge models, Thèse de doctorat, Fac. Agr. Appl. Biol. Sci. Ghent University, Belgique, 2000.
- [PET 00b] PETERSEN B., GERNAEY K., VANROLLEGHEM P.A., « Improved theoretical identifiability of model parameters by combined respirometric-titrimetric measurements: A generalisation of results », dans *Proc. IMACS 3rd MATHMOD Conference*, Vienne, Autriche, vol. 2, p. 639-642, 2000.

- [PET 00c] PETERSEN B., GERNAEY K., VANROLLEGHEM P.A., « Practical identifiability of model parameters by combined respirometric-titrimetric measurements », dans *Proc. Watermatex 2000*, 2000.
- [POH 78] POHJANPALO H., « System identifiability based on the power series expansion of the solution », *Math. Biosci.*, vol. 41, p. 21-33, 1978.
- [POS 90] POSTEN C., MUNACK A., « On-line application of parameter estimation accuracy to biotechnical processes », dans *Proc. ACC*, p. 2181-2186, 1990.
- [PRE 86] PRESS W.H., FLANNERY B.P., TEUKOLSKY S.A., VETTERLING W.T., *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge, Royaume-Uni, 1986.
- [RAK 85] RAKSANYI A., LECOURTIER Y., WALTER E., VENOT A., « Identifiability and distinguishability testing via computer algebra », *Math. Biosci.*, vol. 77, p. 245-266, 1985.
- [RAT 86] RATKOWSKY D.A., « A suitable parameterization of the Michaelis-Menten enzyme reaction », *Biochem. J.*, vol. 240, p. 357-360, 1986.
- [ROB 85] ROBINSON J.A., « Determining microbial parameters using nonlinear regression analysis: Advantages and limitations in microbial ecology », *Adv. Microb. Ecol.*, vol. 8, p. 61-114, 1985.
- [SAN 94] SANCHEZ J., ARIJO S., MUNOZ M., MORINIGO M., BORREGO J., « Microbial colonisation of different support materials used to enhance the methanogenic process », *Appl. Microbiol. Biotechnol.*, vol. 41, p. 480-486, 1994.
- [SCH 98] SCHUETZE M., *Integrated simulation and optimum control of the urban wastewater system*, Thèse de doctorat, Imperial College, Londres, Royaume-Uni, 1998.
- [SPE 00] SPÉRANDIO M., PAUL E., « Estimation of wastewater biodegradable COD fractions by combining respirometric experiments in various SO/XO ratios », *Wat. Res.*, vol. 34, p. 1233-1246, 2000.
- [VAJ 89] VAJDA S., GODFREY K.R., RABITZ H., « Similarity transformation approach to identifiability analysis of nonlinear compartmental models », *Math. Biosci.*, vol. 93, p. 217-248, 1989.
- [VAN 96] VANROLLEGHEM P.A., KEESMAN K.J., « Identification of biodegradation models under model and data uncertainty », *Wat. Sci. Tech.*, vol. 33, n° 2, p. 91-105, 1996.
- [VAN 98] VANROLLEGHEM P.A., DOCHAIN D., « Bioprocess model identification », dans Van Impe J.F., Vanrollegheem P.A., Iserentant D. (dir.), *Advanced Instrumentation, Data Interpretation, and Control of Biotechnological Processes*, p. 251-318, Kluwer, Amsterdam, 1998.

- [VER 97] VERSYCK K.J., CLAES J.E., VAN IMPE J.F., « Practical identification of unstructured growth kinetics by application of optimal experimental design », *Biotechnol. Prog.*, vol. 13, p. 524-531, 1997.
- [VER 00] VERSYCK K.J., VAN IMPE J.F., « Optimal design of system identification experiments for bioprocesses », *Journal A*, vol. 41, n° 2, p. 25-34, 2000.
- [VIA 85] VIALAS C., CHERUY A., GENTIL S., « An experimental approach to improve the Monod model identification », dans Johnson A. (dir.), *Modelling and Control of Biotechnological Processes*, p. 155-160, Pergamon, Oxford, 1985.
- [WAL 82] WALTER E., LECOURTIER Y., « Global approaches to identifiability testing for linear and nonlinear state space models », *Math. Computers Sim.*, vol. 24, p. 472-482, 1982.
- [WAL 94] WALTER E., PRONZATO L., *Identification de modèles paramétriques à partir de données expérimentales*, Masson, Paris, 1994.
- [WAL 95] WALTER E., PRONZATO L., « On the identifiability and distinguishability of nonlinear parametric models », dans *Proc. Symp. on Appl. of Mod. and Cont. in Agric. and Bioind., IMACS, V.A.3.1-8*, Bruxelles, Belgique, 1995.