



HAL
open science

Stratégie d'intensification des procédés

Florent Mathieu

► **To cite this version:**

Florent Mathieu. Stratégie d'intensification des procédés. Autre. Université de Lorraine, 2013. Français. NNT: 2013LORR0210 . tel-01750510

HAL Id: tel-01750510

<https://hal.univ-lorraine.fr/tel-01750510v1>

Submitted on 5 Feb 2024

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

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



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



ÉCOLE DOCTORALE RESSOURCES, PROCÉDÉS, PRODUITS ET ENVIRONNEMENT (RP2E)
SPÉCIALITÉ : GÉNIE DES PROCÉDÉS ET DES PRODUITS

ÉCOLE NATIONALE SUPÉRIEURE DES INDUSTRIES CHIMIQUES

MÉMOIRE PROVISoire DE THÈSE EN VUE DE L'OBTENTION DU TITRE DE
DOCTEUR DE L'UNIVERSITÉ DE LORRAINE

PRÉPARÉ AU
LABORATOIRE DE RÉACTIONS ET DE GÉNIE DES PROCÉDÉS

EN PARTENARIAT AVEC LE
CENTRE DE RECHERCHE ET D'INNOVATION DE SOLVAY À LYON

PRÉSENTÉ PAR
FLORENT MATHIEU

STRATÉGIE D'INTENSIFICATION DES PROCÉDÉS

SOUTENANCE PRÉVUE LE 5 DÉCEMBRE 2013, DEVANT LE JURY COMPOSÉ DE :

M. FRANCIS COURTOIS

PROFESSEUR, AGROPARISTECH (MASSY), RAPPORTEUR

M. LAURENT PRAT

PROFESSEUR, ENSIACET, LGC (TOULOUSE), RAPPORTEUR

M. JEAN-LÉON HOUZELOT

PROFESSEUR, ENSIC, LRGP (NANCY), MEMBRE

M. SÉBASTIEN LOMEL

DIRECTION SCIENTIFIQUE, SOLVAY (LYON), MEMBRE

M. KAMEL RAMDANI

DIRECTION SCIENTIFIQUE, SOLVAY (LYON), MEMBRE

M. FRANTZ FOURNIER

MAÎTRE DE CONFÉRENCES, ENSAIA (NANCY), LRGP, MEMBRE

M. JEAN-MARC COMMENGE

MAÎTRE DE CONFÉRENCES, ENSIC, LRGP (NANCY), CO-DIRECTEUR

M. LAURENT FALK

DIRECTEUR DE RECHERCHE, CNRS, LRGP (NANCY), DIRECTEUR

*« On fait la science avec des faits,
comme on fait une maison avec des
pierres : mais une accumulation de faits
n'est pas plus une science qu'un tas de
pierres n'est une maison. »*

Henri Poincaré

Sommaire

SOMMAIRE	5
NOMENCLATURE.....	9
INTRODUCTION.....	13
CHAPITRE 1 LES MÉTHODES D'OPTIMISATION DE PERFORMANCES	17
1.1 FORMALISME	17
1.1.1 <i>Les grandeurs mesurables</i>	18
1.1.2 <i>La structure des modèles</i>	20
1.1.3 <i>Les conditions expérimentales</i>	23
1.1.4 <i>Les paramètres et leur estimation</i>	26
1.1.5 <i>La performance</i>	32
1.1.6 <i>En résumé</i>	37
1.2 APPROCHE TENDANCIELLE	38
1.2.1 <i>Estimation des paramètres</i>	39
1.2.2 <i>Estimation de la région de confiance</i>	40
1.2.3 <i>Construction de la structure</i>	41
1.2.4 <i>Planification d'expériences</i>	42
1.2.5 <i>Exploitation des résultats</i>	43
1.3 APPROCHE PHÉNOMÉNOLOGIQUE	47
1.3.1 <i>Estimation des paramètres</i>	47
1.3.2 <i>Estimation de la région de confiance</i>	47
1.3.3 <i>Construction de la structure</i>	48
1.3.4 <i>Planification d'expériences</i>	53
1.3.5 <i>Exploitation des résultats</i>	56
1.4 APPROCHE MIXTE	61
1.4.1 <i>L'optimisation évolutionnaire</i>	62
1.4.2 <i>L'algorithme de Nancy</i>	62
1.4.3 <i>L'optimisation séquentielle « run-to-run »</i>	64
1.5 SYNTHÈSE.....	65
CHAPITRE 2 DÉVELOPPEMENT DE LA STRATÉGIE	67
2.1 CONSTRUCTION ET DÉTAIL DE LA STRATÉGIE	67
2.1.1 <i>Un nouveau degré de liberté : le choix technologique</i>	67
2.1.2 <i>Détail de la structure</i>	70
2.1.3 <i>Choix des outils numériques</i>	80
2.2 MISE EN PLACE DU SYSTÈME EXPÉRIMENTAL.....	85
2.2.1 <i>Conception du montage expérimental</i>	85

2.2.2	<i>Caractérisation</i>	87
2.3	CONCLUSION.....	99
CHAPITRE 3 APPLICATION VIRTUELLE DE LA STRATÉGIE		101
3.1	DESCRIPTION DES CAS D'ÉTUDES	103
3.2	RÉPÉTABILITÉ	108
3.2.1	<i>Détail des résultats pour le système n°1</i>	108
3.2.2	<i>Synthèse des systèmes n°2, 3 et 4</i>	110
3.2.3	<i>Conclusion sur la répétabilité</i>	114
3.3	TEST DE ROBUSTESSE DE LA STRATÉGIE	115
3.3.1	<i>Influence des grandeurs mesurées</i>	115
3.3.2	<i>Influence de la valeur des paramètres</i>	118
3.3.3	<i>Influence de l'erreur de mesure</i>	125
3.3.4	<i>Synthèse sur le test de robustesse</i>	127
3.4	L'AVANTAGE DE LA PRISE EN COMPTE DE PLUSIEURS TECHNOLOGIES	128
3.5	COMPARAISON À D'AUTRES STRATÉGIES	130
3.5.1	<i>Approche tendancielle</i>	130
3.5.2	<i>Algorithme de Nancy</i>	132
3.5.3	<i>Résultats</i>	132
3.6	CONCLUSION.....	134
CHAPITRE 4 APPLICATION À UN SYSTÈME RÉACTIONNEL COMPÉTITIF EN PHASE HOMOGENÈME 137		
4.1	DESCRIPTION DU CAS D'ÉTUDE.....	137
4.1.1	<i>La chimie</i>	137
4.1.2	<i>Méthodes d'analyse</i>	139
4.1.3	<i>Simplification du problème</i>	142
4.1.4	<i>Estimation indépendante</i>	143
4.2	VALIDATION EXPÉRIMENTALE DE L'ACQUISITION DE DONNÉES	145
4.2.1	<i>Réacteur batch</i>	146
4.2.2	<i>Réacteur piston</i>	148
4.2.3	<i>Réacteur semi-batch</i>	151
4.2.4	<i>Libre choix des réacteurs</i>	155
4.3	VALIDATION EXPÉRIMENTALE DE LA STRATÉGIE POUR L'OPTIMISATION DE PERFORMANCES.....	158
4.3.1	<i>Initialisation</i>	158
4.3.2	<i>Évolution de la stratégie</i>	159
4.4	CONCLUSION.....	164
CHAPITRE 5 APPLICATION À UN CAS INDUSTRIEL.....		165
5.1	DESCRIPTION DU CAS INDUSTRIEL	165
5.1.1	<i>La chimie</i>	165
5.1.2	<i>Objectifs et contraintes</i>	170

5.1.3	<i>Méthodes analytiques</i>	171
5.2	APPLICATION DE LA STRATÉGIE	178
5.2.1	<i>Domaines expérimentaux</i>	178
5.2.2	<i>Détermination du modèle</i>	179
5.2.3	<i>Déroulements de la stratégie</i>	186
5.3	DISCUSSION	193
5.4	CONCLUSION	197
CONCLUSIONS ET PERSPECTIVES		201
LES POINTS FORTS		202
	<i>Son efficacité</i>	202
	<i>Sa robustesse</i>	203
LES LIMITES.....		203
	<i>Choix de l'utilisateur</i>	203
	<i>Couplage de technologies</i>	204
LE POTENTIEL		205
	<i>Un outil d'aide à la conception</i>	205
	<i>L'automatisation complète et l'interfaçage graphique</i>	205
BIBLIOGRAPHIE		207

Nomenclature

Les variables sont en italique, les vecteurs et matrices sont en gras.

Lettres latines

<i>a</i>	Aire interfaciale spécifique	$[\text{m}^2/\text{m}^3]$
A_0	Facteur pré-exponentiel	$[(\text{m}^3/\text{mol})^{n-1}/\text{s}]$
<i>C</i>	Concentration	$[\text{mol}/\text{m}^3]$
C_p	Capacité thermique massique	$[\text{J}/\text{kg}/\text{K}]$
$d(\mathbf{a}, \mathbf{b})$	Préférence de a par rapport à b (PROMETHEE)	$[-]$
<i>D</i>	Diamètre	$[\text{m}]$
D_m	Constante de diffusion	$[1/\text{s}]$
E_a	Énergie d'activation	$[\text{J}/\text{mol}]$
<i>f</i>	Fonction représentant la structure du modèle	$[-]$
<i>F</i>	Débit molaire	$[\text{mol}/\text{s}]$
G	Jacobien de la fonction de performance γ	$[-]$
<i>j</i>	Critère à minimiser	$[-]$
<i>k</i>	Constante de vitesse réactionnelle	$[(\text{m}^3/\text{mol})^{n-1}/\text{s}]$
\mathcal{L}	Vraisemblance	$[-]$
ℓ	log-vraisemblance	$[-]$
<i>n</i>	Ordre de réaction	$[-]$
\mathcal{N}	Loi normale multivariée	$[-]$
<i>P</i>	Pression	$[\text{Pa}]$
r_i	Vitesse de la réaction <i>i</i>	$[\text{mol}/\text{m}^3/\text{s}]$
R_j	Vitesse de production de l'espèce <i>j</i>	$[\text{mol}/\text{m}^3/\text{s}]$
<i>R</i>	Dimension caractéristique, rayon	$[\text{m}]$
<i>R</i>	Constante des gaz parfait	$[\text{J}/\text{mol}/\text{K}]$
s	Sensibilité du modèle par rapport aux paramètres	$[-]$
<i>t</i>	Temps	$[\text{s}]$
<i>T</i>	Température	$[\text{K}]$
<i>U</i>	Coefficient de transfert thermique global	$[\text{W}/\text{m}^2/\text{K}]$
<i>V</i>	Volume	$[\text{m}^3]$
W	Matrice de pondération	$[-]$
y	Vecteur des grandeurs mesurées de dimension $(1 \times n_y)$	$[-]$

Lettres grecques

β	Biais du modèle	[–]
γ	Fonction de performance	[–]
Δ_r	Domaine opératoire du réacteur r	[–]
ϵ	Erreur expérimentale	[–]
θ	Vecteur des paramètres de dimension ($1 \times n_p$)	[–]
Θ	Domaine de recherche des paramètres	[–]
λ	Conductivité thermique	[W/m/K]
ν	Coefficient stœchiométrique	[–]
ξ	Vecteur de conditions opératoires de dimension ($1 \times n_x$)	[–]
Ξ	Matrice de conditions opératoires de dimension ($n_{pm} \times n_x$)	[–]
π	Densité de probabilité	[–]
ρ	Masse volumique	[kg/m ³]
Σ	Matrice de covariance	[–]
Φ	Flux de préférences net (PROMETHEE)	[–]
Ω_r	Espace expérimental du réacteur r	[–]

Indices et exposants

E	Se réfère à l'entrée du réacteur
i	Se réfère à la réaction i
j	Se réfère à l'espèce j
k	Se réfère à l'expérience k
p	Se réfère à la paroi du réacteur
r	Se réfère au réacteur r
S	Se réfère à la sortie du réacteur

Abbréviations

AA	AmidoAmine
AGly	Acide Glycolique
BET	Bétaïne
GAIA	Geometrical Analysis for Interactive Aid
MAUT	Multi Attribute Utility Theory
mc	Moindre carré
MCA	Acide MonoChloroAcétique
mle	Maximum Likelihood Estimator
PROMETHEE	Preference Ranking Organisation METHod for Enrichment Evaluations

Introduction

Dans un contexte industriel de plus en plus concurrentiel, il est primordial pour une entreprise de pouvoir mettre sur le marché de nouveaux produits le plus rapidement possible tout en ayant un outil de production le plus efficace possible. Ces deux objectifs sont souvent contradictoires, car plus l'on veut développer un outil de production efficace, plus sa mise au point est onéreuse et coûteuse en temps. Pour répondre à ces besoins, nous avons pour objectif de définir une méthodologie permettant la mise en œuvre de l'outil de production le plus performant possible dans les plus brefs délais.

La transformation de matières premières en molécules à haute valeur ajoutée passe par la mise en œuvre de procédés chimiques complexes. Ces procédés, qu'ils soient continus ou discontinus font intervenir une grande diversité d'opérations unitaires, telles que des mélanges, des réactions ou encore toutes sortes de séparations. Pour chacune de ces opérations unitaires, il existe souvent plusieurs solutions technologiques qui peuvent être mises en œuvre. Il faut donc dans un premier temps, choisir la technologie la plus adaptée, puis dans un second temps, trouver son optimum de fonctionnement, autrement dit son *design* optimal. Ainsi, pour trouver la meilleure technologie associée à son design optimal, on doit être en mesure d'estimer la performance globale du système pour chaque technologie. La performance globale s'évalue généralement selon plusieurs critères, qui peuvent être antagonistes, tels que la productivité, le coût d'investissement, le coût opératoire, la sécurité, ...

Pour faire le bon choix, nous devons être capables de prédire précisément les différents critères de performance de chaque technologie dans sa gamme de fonctionnement. La prédiction du comportement des différentes technologies passe ainsi par un modèle qui permet d'obtenir une représentation fiable du système étudié. L'objectif étant de quantifier un gain sur plusieurs critères, il est donc nécessaire d'utiliser des modèles quantitatifs.

Deux grandes classes de modèles quantitatifs existent. La première classe est la classe de modèles que l'on qualifie de « boîtes noires », car ils n'ont aucune réalité physique. Ces modèles sont souvent basés sur des considérations purement statistiques. Il existe de nombreuses méthodes qui les utilisent et dont l'efficacité n'est plus à prouver. Le lecteur pourra trouver dans le livre publié par Box (2005) des explications très complètes sur ces méthodes dont l'intérêt majeur réside en ce qu'elles ne requièrent pas de connaissance a priori du système. A contrario, la deuxième classe de modèles est qualifiée de « boîtes blanches » car ces modèles sont justement basés sur la description des phénomènes qui permettent d'expliquer et de prédire le comportement du système. L'introduction de connaissance a priori dans le modèle implique un traitement mathématique de ces modèles plus complexes. Pour l'heure, il n'existe quasiment pas de méthodes générales sur leur utilisation, mais seulement des principes généraux.

Quelle que soit la classe de modèle utilisée, le processus nécessaire à la prédiction du comportement d'un système, passe toujours par la définition du modèle utilisé, l'apport d'informations et enfin une phase d'interprétation. Concrètement, l'apport d'informations consiste en la réalisation de plusieurs expériences permettant de générer une base de données expérimentale, on appelle cette phase « l'acquisition de données ». Ainsi, dans la phase d'acquisition de données, on utilise le modèle pour qu'il représente au mieux ce qu'on observe, c'est son caractère explicatif. Tandis que dans la phase d'interprétation, le modèle est utilisé pour son caractère prédictif. La plupart du temps, la phase la plus longue et la plus onéreuse est l'acquisition de données. En effet, lors de cette phase, il faut mobiliser de nombreux moyens expérimentaux et analytiques. L'objectif industriel de réduire le temps de mise sur le marché se trouve fortement impacté par de l'effort expérimental à fournir. Rationnaliser cet effort est donc un enjeu crucial et la stratégie que nous avons développée tente d'y apporter une réponse.

Considérer le procédé dans son ensemble est un programme bien trop vaste, ainsi nous avons restreint notre champ d'étude à l'étape de réaction seule. Cette étape est particulièrement intéressante car elle est le cœur du procédé chimique. Au cours de l'étape de réaction, de nombreux phénomènes ont lieu en plus de la réaction elle-même, en particulier des phénomènes de transport de matière et de transport de chaleur. Le couplage de ces phénomènes rend leur identification et leur compréhension souvent très ardues. Nous nous limiterons aux réactions en phase liquide, sachant que les concepts et la stratégie développée pourront être étendus et transposés aux types de système réactionnel.

Un aspect important qui nous a conduits à nous focaliser sur l'étape de réaction est la diversité des technologies disponibles. Depuis la fin des années 90, de nouvelles technologies ont été développées dans le but d'intensifier les procédés. Ces technologies intensifiées ont été conçues pour réduire l'impact des phénomènes limitants qui peuvent considérablement ralentir un procédé si aucune précaution n'est prise. Elles permettent d'atteindre des conditions opératoires drastiques (forte température, forte pression, faible échelle, gravité artificielle...). Celles-ci ont pu voir le jour car elles nécessitent des technologies de fabrication avancées. Outre les conditions opératoires bien plus étendues que les technologies conventionnelles, ces technologies fonctionnent presque toutes dans un mode de production continu contrairement aux réacteurs batch largement utilisés dans la chimie de spécialité et pharmaceutique. Il est donc très important de pouvoir peser les pour et les contre aussi bien du gain engendré par l'extension des conditions opératoires que par le passage d'un mode discontinu à un mode continu.

Feuille de route

Ce travail propose une stratégie qui permet d'aider son utilisateur dans le choix de plusieurs technologies. Nous nous sommes focalisés sur l'étape de réaction en phase liquide des procédés chimiques.

Le premier chapitre est donc une revue des différentes approches basées sur des outils mathématiques permettant l'optimisation de performance. Ces approches passent toutes par la modélisation du système puis par l'exploitation du modèle. Elles diffèrent dans la manière de le concevoir et dans la manière d'acquérir des informations sur le système étudié.

Le chapitre 2 sera consacré au détail de la stratégie expérimentale développée et à l'argumentation du choix des outils mathématiques utilisés. De plus, dans ce chapitre est décrit le banc d'essai expérimental que nous avons conçu pour faciliter l'utilisation de plusieurs technologies.

Le chapitre 3 détaille les procédures de tests et leurs résultats. Ces tests réalisés grâce à des expériences virtuelles ont permis de démontrer le bienfondé de la stratégie et de mettre en lumière sa validité et ses limites.

Le chapitre 4 propose une mise en œuvre expérimentale de la stratégie sur un cas d'étude simplifié. Le système réactionnel a été choisi de telle sorte qu'il soit détaillé dans la littérature afin de nous affranchir de la difficulté qu'est la description du modèle.

Enfin, le cinquième et dernier chapitre donne une application industrielle de la stratégie. Le cas d'étude industriel utilisé est la synthèse d'un tensio-actif utilisé dans les produits de soin.

Chapitre 1

Les méthodes d'optimisation de performances

L'objectif de ce chapitre est de donner au lecteur les éléments théoriques nécessaires à la compréhension détaillée de la stratégie et aux arguments de sa construction. Nous décrirons des méthodes mathématiques ayant pour objectif de trouver l'optimum de performance d'un système, les performances devant être définies par l'utilisateur.

Les trois grandes approches sont l'approche tendancielle, l'approche phénoménologique et l'approche auto-adaptative. Chacune de ces approches est basée sur le même principe. On représente le comportement du système par un modèle mathématique, puis l'on cherche l'optimum prédit par le modèle. Ces approches se différencient par les hypothèses qui permettent la construction du modèle. Le même formalisme mathématique est utilisé pour ces trois approches et est détaillé dans le premier paragraphe. Puis, nous détaillerons les trois approches dans des paragraphes respectifs.

1.1 Formalisme

La modélisation a pour but non seulement de décrire mais aussi et surtout de prédire le comportement d'un système. Ainsi, un bon modèle permet de reproduire ou de prédire les grandeurs \mathbf{y} observables expérimentalement (concentration, viscosité,...) pour des conditions expérimentales données ξ^1 (temps d'opération, température,...). Ainsi, d'un point de vue mathématique, un modèle est une fonction vectorielle. On entre l'argument ξ , et le modèle renvoie les grandeurs observables estimées $\hat{\mathbf{y}}$:

$$\text{modèle} : \xi \mapsto \hat{\mathbf{y}} \quad (1.1)$$

Mais cette notation est incomplète. En effet, les grandeurs mesurées expérimentalement sont toujours entachées d'une erreur de mesure ϵ (qui est liée principalement à l'expérimentateur et à la précision des mesures, ...). Cette erreur de mesure est souvent supposée répondre à une loi de probabilité normale. Ainsi, les grandeurs mesurées vraies \mathbf{y}^* sont différentes des grandeurs observables mesurées $\tilde{\mathbf{y}}$:

¹Les symboles représentés en gras signifient qu'il s'agit d'un vecteur, et en majuscule, d'une matrice.

$$\tilde{\mathbf{y}} = \mathbf{y}^* + \boldsymbol{\epsilon} \quad (1.2)$$

Par ailleurs, lorsque l'on crée un modèle, il n'est pas certain que le modèle reflète bien la réalité des choses. Cet écart intrinsèque au modèle est appelé le biais $\boldsymbol{\beta}$.

$$\mathbf{y}^* = \hat{\mathbf{y}} + \boldsymbol{\beta} \quad (1.3)$$

Enfin, la construction d'un modèle repose sur deux composantes qui sont : la structure (ou l'architecture) du modèle qui représente les lois mathématiques et qui permettent d'obtenir une estimation numérique des grandeurs observables que nous noterons \mathbf{f} , et d'autre part, des constantes qui sont appelées « paramètres » et que nous noterons $\boldsymbol{\theta}$. Certains de ces paramètres sont connus, les autres doivent être estimés. Ces estimations peuvent se faire indépendamment du processus de modélisation (par exemple, c'est le cas lorsque les valeurs sont issues de la littérature) ou doivent être identifiés (nous reviendrons plus tard sur ce point) pour que le modèle fonctionne correctement. Ainsi, un modèle c'est l'ensemble du couple « structure-paramètres ».

On peut donc formaliser le processus de modélisation par :

$$\tilde{\mathbf{y}} = \mathbf{f}(\boldsymbol{\xi}, \boldsymbol{\theta}) + \boldsymbol{\epsilon} + \boldsymbol{\beta} \quad (1.4)$$

Pour résumer, la modélisation est un processus itératif qui consiste en une phase d'élaboration de la structure du modèle, suivie d'une phase d'identification paramétrique couplée à une phase d'acquisition de données et finalement une phase de validation. Si la validation échoue, alors il faut acquérir plus de données, ou améliorer la structure du modèle.

Une fois le modèle capable de prédire le comportement du système étudié, on peut utiliser ces prédictions pour calculer des critères de performances que l'on note \mathbf{j} , la fonction permettant le calcul des performances est noté γ . Cette fonction dépend des conditions expérimentales et des grandeurs observables. Ainsi, on a :

$$\hat{\mathbf{j}} = \gamma(\boldsymbol{\xi}, \hat{\mathbf{y}}) \quad (1.5)$$

Si les grandeurs observables ne sont pas estimées par le modèle mais mesurées expérimentalement, on obtient alors des performances mesurées :

$$\tilde{\mathbf{j}} = \gamma(\boldsymbol{\xi}, \tilde{\mathbf{y}}) \quad (1.6)$$

1.1.1 Les grandeurs mesurables

Les grandeurs que l'on mesure sont entachées d'une erreur expérimentale. Il est important pour la suite de pouvoir caractériser cette erreur. On admet que la valeur mesurée de la grandeur observable est régie par une loi de distribution de probabilité centrée sur la valeur théorique. Ainsi, en répétant un grand nombre de fois une même mesure, la moyenne des valeurs mesurées devrait être égale à la valeur théorique et la dispersion des écarts par rapport à la valeur théorique serait caractérisée par la variance de la distribution de probabilité. Dans le cas où plusieurs grandeurs observables sont

mesurées, il se peut qu'il y ait une interaction entre les différentes grandeurs, par exemple la mesure du pH dépend directement de la température, donc une erreur de mesure sur la température induira une erreur sur la mesure du pH. Cette interaction est décrite par les éléments non diagonaux de la matrice de covariance² des erreurs de mesure Σ_y .

$$\Sigma_y = \begin{bmatrix} \sigma_{y_1}^2 & \sigma_{y_1 y_2}^2 & \cdots & \sigma_{y_1 y_n}^2 \\ \sigma_{y_2 y_1}^2 & \sigma_{y_2}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{y_n y_1}^2 & \cdots & \cdots & \sigma_{y_n}^2 \end{bmatrix}$$

Dans certains cas, l'erreur de mesure dépend de la valeur elle-même. Cela est le cas pour la mesure de température par thermocouple : l'erreur de mesure augmente linéairement avec la température. Les modèles de variance les plus fréquemment utilisés sont détaillés dans le Tableau 1.1 où ω représente l'écart-type standard et γ le facteur hétéroscédastique :

Tableau 1.1 : Modèle de variance

Modèle de variance	Description	γ
<i>Constante</i>	$\sigma^2 = \omega^2$	0
<i>Constante relative</i>	$\sigma^2 = \omega^2(\hat{y}^2)$	1
<i>Hétéroscédastique</i>	$\sigma^2 = \omega^2(\hat{y}^2)^\gamma$	Dans l'intervalle [0 ; 1]
<i>Linéaire</i>	$\sigma^2 = \bar{\alpha} \cdot y + \bar{\beta}$	

Ainsi, la variance de la grandeur observable i lors du point de mesure k s'écrit :

$$\sigma_{y_{ik}}^2 = \omega_i^2 (\hat{y}_{i,k}^2)^{\gamma_i} \quad (1.7)$$

La valeur des grandeurs mesurables générées par le modèle contient aussi une incertitude : un modèle prédit une valeur avec un intervalle de confiance. L'intervalle de confiance dépend de la structure du modèle et de la connaissance que l'on a des paramètres du modèle. Plus les paramètres sont estimés précisément, plus l'intervalle de confiance du modèle sera réduit. Cette erreur du modèle est modélisée par une loi normale multivariée de matrice de covariance Σ_f .

² La matrice de covariance de l'erreur expérimentale est la généralisation à plusieurs variables de la variance.

1.1.2 La structure des modèles

Dans ce paragraphe, nous nous intéressons aux différents types de structure de modèles.

Modèles « boîtes noires »

Les modèles dits « boîtes noires » sont des modèles dont les lois mathématiques utilisées pour décrire la structure du modèle n'ont pas de signification physique connue. Les modèles polynomiaux, linéaires, multilinéaires, exponentiels, et bien d'autres en font partie. L'intérêt de ce type de modèle réside dans le fait qu'il n'est pas nécessaire d'avoir une connaissance physique du système que l'on souhaite décrire. Mais cet avantage a des implications, en particulier ces modèles ont un domaine de validité restreint, en d'autres termes, ils ne sont, en général, pas extrapolables.

Un autre type de modèle « boîte noire » sont les réseaux de neurones. Les réseaux de neurones peuvent être vus comme des modèles auto-adaptatifs, où chaque neurone, qui est une fonction élémentaire, va s'exprimer plus ou moins. Le nombre, la manière dont sont enchevêtrés et le type de fonctions de chaque neurone décrit la structure du réseau. Les poids des différents flux d'entrée et de sortie de chaque neurone lui permettra de s'adapter, c'est-à-dire de représenter des données en fonction de conditions. Pour trouver ces poids, il est préalablement nécessaire de définir un critère de performance caractérisant de la « régression paramétrique » (ou *fitting*), puis grâce à un optimiseur on peut trouver les poids optimaux.

Avantages et inconvénients

La principale difficulté lorsque l'on utilise des modèles boîtes noires est la définition de la structure du modèle. Si la structure est trop simple, le modèle ne représentera pas bien l'allure de la courbe et possèdera un biais important. Au contraire, s'il est trop complexe, il sera suffisamment souple pour passer par tous les points (le biais sera très faible), et sera alors influencé par le bruit de la base de données (forte variance). Il y a donc un compromis à trouver sur la complexité du modèle.

La Figure 1.1 et Figure 1.2 montrent que plus le modèle est complexe (degré du polynôme et respectivement le nombre de neurones) plus le modèle est souple. Au-delà d'une certaine souplesse, le modèle est trop souple et devient très fortement influencé par le bruit de mesure. Le compromis entre souplesse et représentativité est appelé le compromis biais/variance. Pour les réseaux de neurones, la recherche d'algorithme permettant de trouver la complexité optimale fait l'objet de nombreuses recherches.

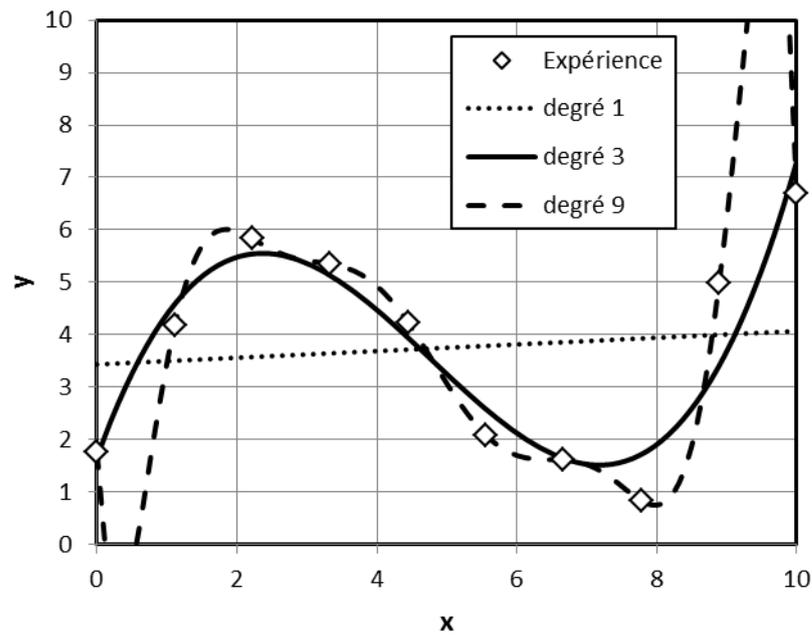


Figure 1.1 : Illustration du compromis biais-variance avec des polynômes.

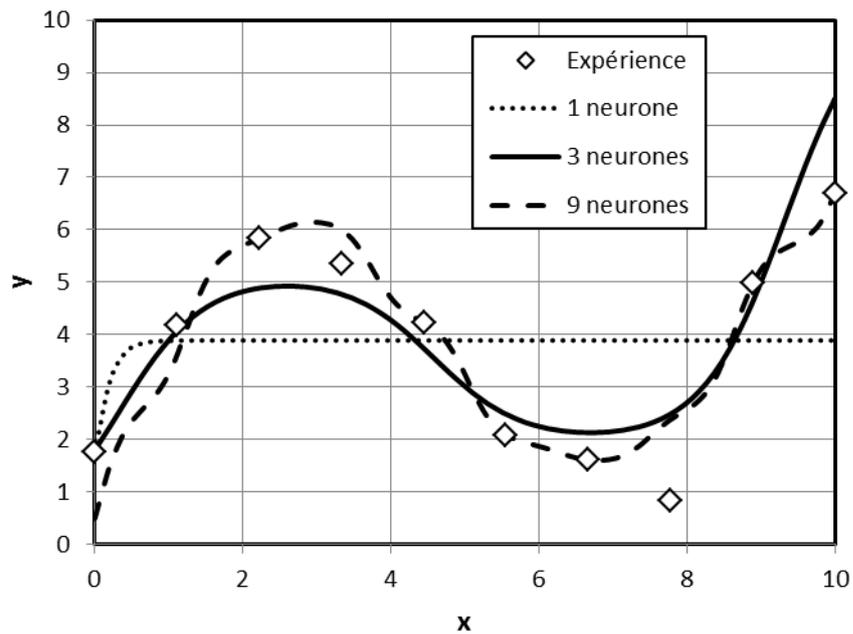


Figure 1.2 : Illustration du compromis biais-variance avec des réseaux de neurones.

1.1.2.1 Modèles « boîtes blanches »

A l'opposé des modèles « boîtes noires », les modèles « boîtes blanches » sont basés sur la compréhension physique des phénomènes décrits. Pour cette raison, ils sont plus couramment appelés modèles de connaissance, modèles de comportement ou encore modèles phénoménologiques. L'intérêt de ce type de modèle se trouve principa-

lement dans leur caractère prédictif. Mais, ces modèles sont souvent difficiles à mettre en œuvre pour plusieurs raisons :

- La complexité avec laquelle les phénomènes interagissent ne permet pas toujours de les identifier.
- Le nombre de phénomènes mis en jeu peut-être très important, ce qui requiert de faire de lourdes hypothèses afin de rendre opérant le modèle.
- Les lois mathématiques qui régissent certains phénomènes ne sont pas toujours connues.
- Chaque loi mathématique fait intervenir de nombreux paramètres qui rendent leur identification difficile.
- La résolution d'un modèle peut nécessiter des moyens de calculs élevés.

Lorsque le modèle est suffisamment simple, il est possible de le résoudre de manière analytique, par exemple le mouvement d'un corps dans un champ de gravitation uniforme. On peut donc utiliser la solution analytique directement, on dira alors que le modèle est décrit de manière explicite.

Malheureusement, les modèles que l'on sait résoudre analytiquement ne représentent qu'une infime partie. Ainsi, la plupart des modèles sont décrits de manière implicite, par exemple par un système d'équations différentielles, et nécessitent des méthodes de résolution mathématiques poussées, ce qui rend leur exploitation plus difficile. Cependant, il est important de préciser que la résolution numérique d'un système d'équations différentielles ordinaires (ODEs) est de loin plus facile à mettre en œuvre que la résolution d'équations aux dérivées partielles. Dans certains cas, la résolution du modèle est en elle-même un challenge.

Le choix des phénomènes à prendre en compte est tout à fait primordial et est défini par l'utilisateur uniquement : il n'existe pas de méthode systématique pour construire la structure du modèle. Ce choix doit être un compromis entre une simplification trop forte qui réduit la précision et le détail trop avancé qui empêche la résolution ou l'estimation des paramètres.

Dans certains cas, les phénomènes ayant lieu ne sont pas connus. L'une des réponses à ce problème est l'utilisation de la discrimination d'hypothèses, développé par Box dans (Box et Hill 1967). Cette méthode, basée sur la notion d'entropie d'information, consiste à postuler plusieurs modèles et à chercher les expériences qui permettent de les départager.

Pour les cas spécifiques où les stœchiométries ne sont pas connues, on peut faire appel à une méthode développée à l'école polytechnique fédérale de Lausanne par Bonvin et Rippin (1990) nommé Target Factor Analysis (TFA). La TFA permet, à partir de bilans de matière, de retrouver le nombre de réactions indépendantes et de quantifier la probabilité qu'une réaction donnée ait lieu.

1.1.3 Les conditions expérimentales

Les conditions expérimentales se définissent par deux concepts : les fenêtres opératoires et la flexibilité d'échantillonnage.

1.1.3.1 Fenêtres opératoires

Les fenêtres expérimentales pour un réacteur sont définies par deux types de limitation :

- Les limitations intrinsèques. Par exemple, un réacteur constitué d'une cuve agitée en verre est limité en température par la température d'ébullition du solvant à l'air libre. En pressurant ce réacteur, on peut élargir cette fenêtre en augmentant le point d'ébullition. De la même manière, le temps de réaction accessible dans une cuve agitée est de quelques minutes à quelques heures, tandis qu'un microréacteur permet d'atteindre des temps de réaction inférieure à la seconde mais ne permet pas d'avoir des temps de séjour trop longs. La Figure 1.3 montre ce concept.

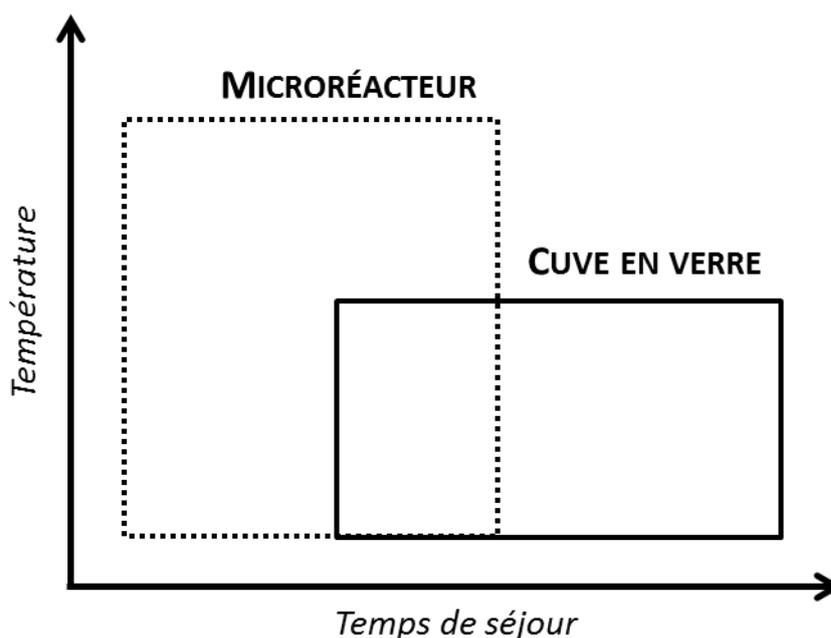


Figure 1.3 : Exemple de limitation intrinsèque du domaine opératoire de deux technologies

- Les limitations opératoires sont fixées par l'utilisateur et dépendent du système chimique étudié. Par exemple, à cause de la perte de charge dans un réacteur continu tubulaire d'une longueur donnée, des temps de séjour trop courts ou des viscosités trop élevées ne peuvent pas être mises en œuvre. Dans

ce cas, une partie de la fenêtre opératoire ne peut pas être opérée (voir Figure 1.4). Une autre limitation opératoire importante est liée à l'exothermie des réactions mises en œuvre. Chaque technologie a une capacité de d'évacuer l'énergie thermique plus ou moins élevée. Si le système chimique étudié est exothermique, en changeant de technologie, il conviendra d'être très attentif au fait de ne pas se placer dans des conditions où il y a un risque de limitation de transfert de chaleur qui induirait l'accumulation de chaleur et donc un emballement de la réaction. Sur la Figure 1.5 sont reportées les gammes de temps caractéristiques de transfert de matière et de chaleur pour des équipements usuels. Lorsque le temps de transfert thermique est supérieur au temps de transfert de matière, il y a un risque (zone grisée sur la figure).

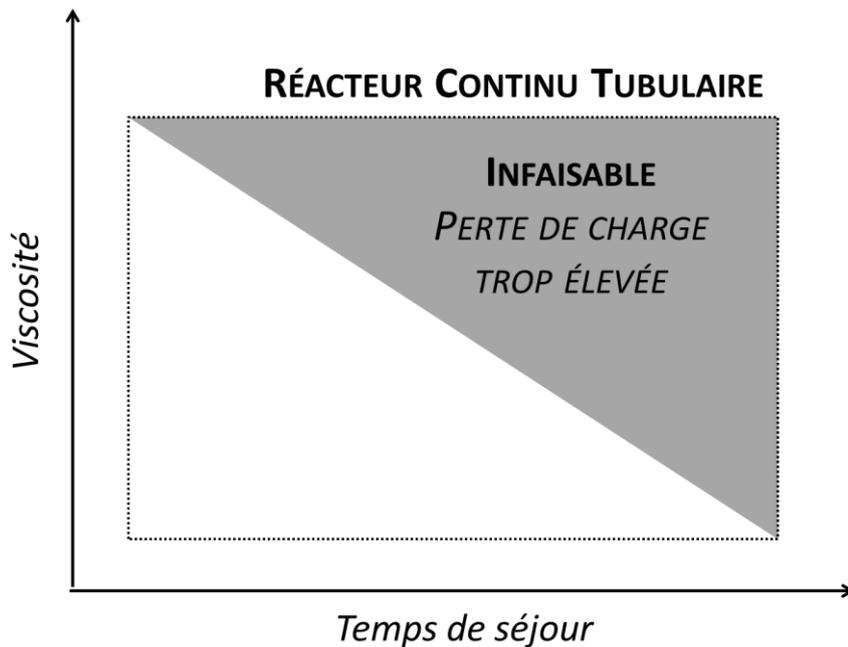


Figure 1.4 : Exemple de limitation extrinsèque liée à la viscosité.

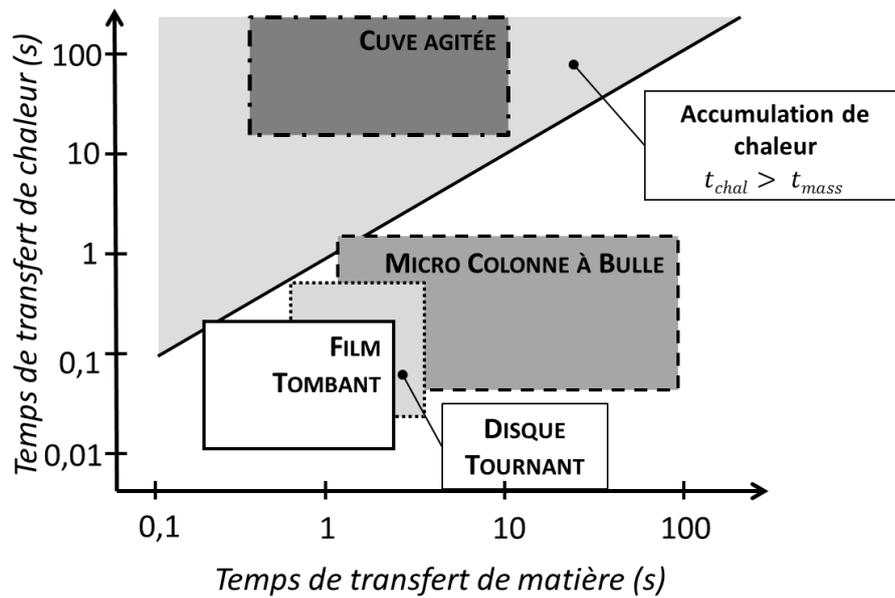


Figure 1.5 : Exemple de limitation extrinsèque liée à la possible accumulation de chaleur.

Ainsi, pour chaque réacteur utilisé, il est nécessaire de définir leur fenêtre expérimentale en prenant compte des limitations intrinsèques et des limitations opératoires. La fenêtre expérimentale (ou domaine expérimental) d'un réacteur r est notée Δ_r .

1.1.3.2 Échantillonnage

Lorsque l'on réalise des expériences en laboratoire, il est fréquent de pouvoir générer plusieurs échantillons en une seule expérience. Ceci est facilement concevable pour le cas des réacteurs en cuve agitée discontinue où il suffit de prélever une partie du réacteur à différents instants, sous réserve de pouvoir analyser directement ou de pouvoir figer l'échantillon. Avec un réacteur continu tubulaire, dans la limite où l'hydrodynamique reste constante, il est possible de faire varier le débit des pompes de manière à modifier le temps de séjour et le ratio des réactifs à l'entrée du réacteur. Dans la mesure où le temps de séjour est relativement court, donc que le temps de mise en régime du réacteur soit de quelques dizaines de minutes, il est possible de tester plusieurs conditions de temps de séjour et de ratio dans un intervalle de temps similaire à ce que requiert une expérience en réacteur batch. La Figure 1.6 montre comment peuvent être représentées des expériences générant plusieurs échantillons sur la fenêtre opératoire de deux réacteurs classiques.

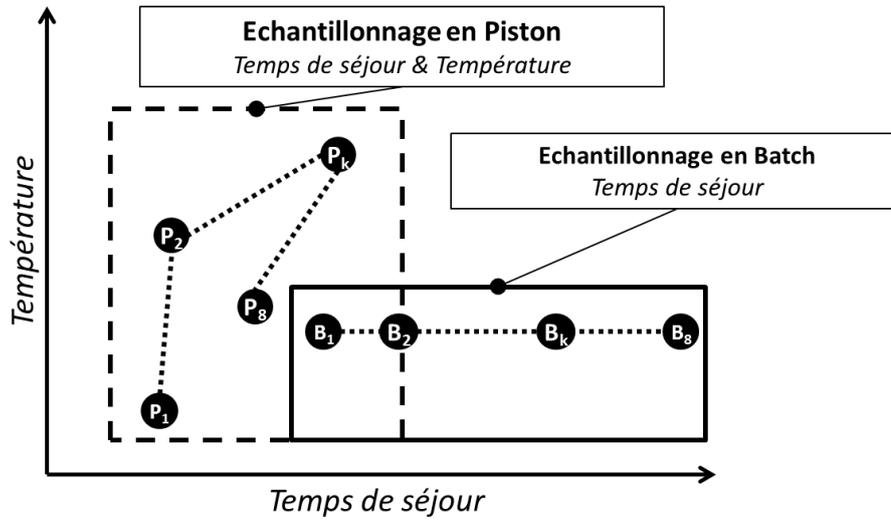


Figure 1.6 : Exemple d'échantillonnage en réacteur batch et en réacteur piston.

Pour formaliser le concept d'échantillonnage, on introduit la notion d'expérience Ξ_k telle qu'une expérience k contienne n_{pm} points de mesure :

$$\Xi_k = \{ \xi_{k,1}, \dots, \xi_{k,n_{pm}} \} \quad (1.8)$$

Par exemple, une expérience réalisée en batch, où les concentrations initiales en réactifs sont C_A^0 et C_B^0 , avec trois points de mesure s'écrirait :

$$\Xi = \begin{bmatrix} t_1 & C_A^0 & C_B^0 \\ t_2 & \vdots & \vdots \\ t_3 & C_A^0 & C_B^0 \end{bmatrix} \quad (1.9)$$

Compte tenu des problèmes expérimentaux, il existe des contraintes d'échantillonnage. Par exemple, il est très difficile de prélever deux échantillons à moins de 5 minutes d'intervalle. Ces contraintes doivent être définies par l'utilisateur. Pour résumer, l'espace expérimental d'un réacteur r , noté Ω_r , contient d'une part la fenêtre expérimentale Δ_r décrite plus haut et, d'autre part les contraintes d'échantillonnage.

1.1.4 Les paramètres et leur estimation

Lorsque la structure du modèle est fixée, il faut estimer les paramètres du modèle de telle sorte que les prédictions soient les plus représentatives possible des mesures. Pour cela, on utilise un critère qui permet de quantifier la bonne adéquation entre mesures et prédictions. Il existe plusieurs critères selon le type de problème. Nous ne détaillerons ici que les plus importants ; pour plus de détails, se reporter à Walter et Pronzato (1994). Les paramètres estimés $\hat{\theta}$ au sens du critère d'adéquation choisi $j(\cdot | \cdot)$ s'obtiennent en minimisant ledit critère sur un domaine de recherche donné Θ qui dépend des données expérimentales obtenues dans les conditions Ξ :

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} j(\boldsymbol{\theta}|\Xi) \quad (1.10)$$

Sauf dans certains cas particuliers, il n'existe pas de solution analytique qui permette de trouver les valeurs des paramètres qui minimisent le critère d'adéquation, on utilisera donc des méthodes d'optimisation numérique que nous ne détaillerons pas ; pour plus d'information se référer à Pronzato et Walter (1988).

1.1.4.1 Critères quadratiques

Les critères quadratiques, aussi appelés critères des moindres carrés j_{mc} , sont les plus couramment utilisés. Et pour cause, ils sont intuitifs, possèdent des propriétés de dérivation intéressantes pour l'optimisation et peuvent être construits indépendamment de tout critère statistique. Ces critères sont la somme pondérée par $\{w_{ki}\}$ des carrés des écarts entre mesures et estimations en chaque point de mesure et pour chaque grandeur observable :

$$j_{mc}(\boldsymbol{\theta}|\Xi) = \sum_{k=1}^{n_{pm}} \sum_{i=1}^{n_y} w_{ki} (y_{ki} - f_i(\boldsymbol{\xi}_k, \boldsymbol{\theta}))^2 \quad (1.11)$$

Sous forme matricielle :

$$j_{mc}(\boldsymbol{\theta}|\Xi) = \sum_{k=1}^{n_{pm}} [\mathbf{y}_k - \mathbf{f}(\boldsymbol{\xi}_k, \boldsymbol{\theta})]^T \mathbf{W}_k [\mathbf{y}_k - \mathbf{f}(\boldsymbol{\xi}_k, \boldsymbol{\theta})] \quad (1.12)$$

Ce qui va différencier les différents critères quadratiques c'est la définition de la matrice de poids \mathbf{W} . Voici quelques exemples :

- $w_{ki} = 0$ permet d'éliminer une donnée considérée douteuse.
- $w_{ki} = 1/y_{ki}^2$ permet de rendre l'écart prédiction/mesure relatif. Ainsi, on donnera plus d'importance aux petites valeurs. Cela peut s'avérer très utile lorsqu'une grandeur observable varie de plusieurs ordres de grandeur.
- $\mathbf{W} = \boldsymbol{\Sigma}_y^{-1}$ permet de favoriser les grandeurs observables les plus fiables. Plus la variance d'une grandeur est faible, plus son poids est important, et vice et versa. Cela suppose de connaître la matrice de covariance $\boldsymbol{\Sigma}_y$, ce qui n'est pas toujours le cas.

1.1.4.2 Maximum de vraisemblance

La matrice de covariance Σ_y n'est pas toujours connue, elle peut être estimée³ mais cela requiert un grand nombre de données répétitives. La méthode du maximum de vraisemblance permet de s'affranchir de cette difficulté.

La Figure 1.7 montre un échantillon de n observations indépendantes, ainsi que deux fonctions de distribution de probabilité $\pi(\cdot | \theta_1)$ et $\pi(\cdot | \theta_2)$. Ces deux fonctions sont identiques, seuls θ_1 et θ_2 sont différents et représentent leur position sur l'axe des abscisses.

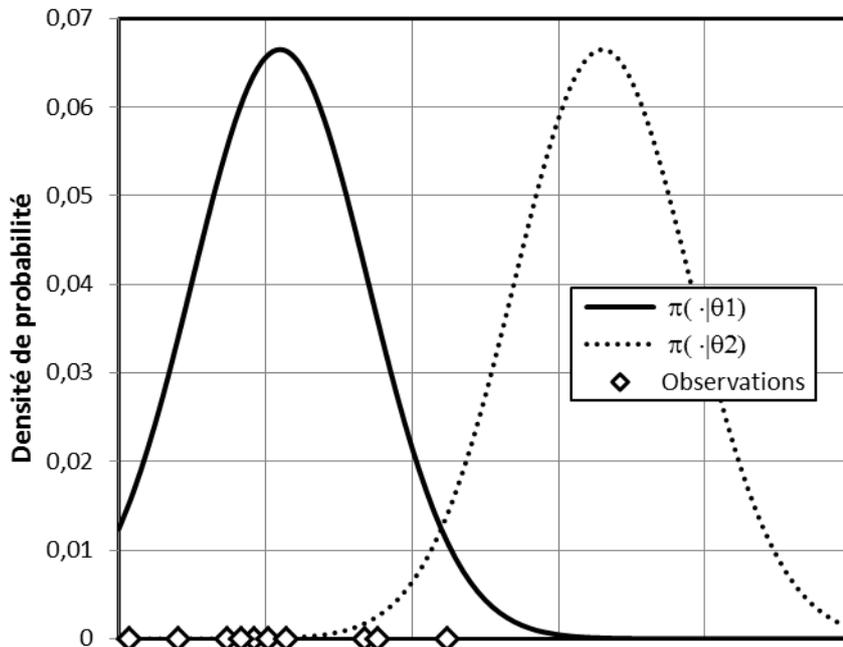


Figure 1.7 : Comparaison de deux fonctions de distribution de probabilité par rapport à plusieurs observations.

Il est fortement vraisemblable que la distribution paramétrée par θ_1 soit celle qui soit à l'origine des échantillons observés. En effet, même s'il n'est pas impossible que $\pi(\cdot | \theta_2)$ en soit à l'origine, il est invraisemblable que toutes les observations apparaissent dans des régions de faible densité de probabilité de $\pi(\cdot | \theta_2)$. Afin de quantifier cette intuition, on peut définir la vraisemblance de la distribution $\pi(\cdot | \theta)$ pour l'échantillon x_1, \dots, x_n par :

³ Soit n_{rep} le nombre de répétitions et \bar{y} la moyenne des mesures de chaque grandeur observable $\Sigma_y = \frac{1}{n_{rep}} \sum_{k=1}^{n_{rep}} (y_k - \bar{y})(y_k - \bar{y})^T$

$$\text{Vraisemblance}^4 = \mathcal{L}(\theta | x_1, \dots, x_n) = \prod_{k=1}^n \pi(x_k | \theta) \quad (1.13)$$

Il est clair que la vraisemblance ne sera élevée que si toutes les observations sont dans des régions où la fonction de distribution ne prend pas de très petites valeurs. On conçoit aussi aisément, qu'il est possible de chercher le paramètre optimal $\hat{\theta}$ qui maximise la vraisemblance. L'estimation du paramètre obtenu par cette méthode⁵ dépend bien évidemment de l'échantillon x_1, \dots, x_n et du domaine de recherche Θ . On note alors :

$$\hat{\theta}_{mle} = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta | x_1, \dots, x_n) \quad (1.14)$$

La vraisemblance étant définie comme un produit, il est généralement difficile de la maximiser. On peut la transformer en somme en utilisant le logarithme tout en conservant ses propriétés (le logarithme étant une fonction monotone croissante), ce nouveau critère est appelé la log-vraisemblance et sera notée ℓ :

$$\ell = \ln \mathcal{L} \quad (1.15)$$

On peut étendre la notion de maximum de vraisemblance à l'estimation des paramètres d'un modèle. En effet, si l'on admet que le modèle ne possède pas de biais ($\beta = 0$) et que l'erreur commise sur une mesure au point de mesure k ($\epsilon_k = y_k - f(\xi_k, \theta)$) suit une distribution de probabilité multivariée $\pi_y(\cdot)$, la vraisemblance, pour la base de données expérimentales Ξ , devient :

$$\mathcal{L}(\theta | \Xi) = \prod_{k=1}^{n_{pm}} \pi_y(\epsilon_k) \quad (1.16)$$

En faisant l'hypothèse que l'erreur de mesure est une variable aléatoire indépendante suivant une loi normale multivariée de covariance constante Σ_y , on a :

$$\pi_y(\epsilon_k) = \mathcal{N}(\mathbf{0}, \Sigma_y) = [(2\pi)^{n_y} \det \Sigma_y]^{-1/2} \exp \left[-\frac{1}{2} \epsilon_k^T \Sigma_y^{-1} \epsilon_k \right] \quad (1.17)$$

Cette hypothèse introduit la matrice de covariance Σ_y qu'il faut à présent estimer au même titre que le vecteur de paramètres θ . Ainsi la log-vraisemblance devient :

⁴ La notation \mathcal{L} fait référence au terme anglais « Likelihood », et est universellement adoptée.

⁵ On utilisera l'abréviation MLE issu du terme anglais « Maximum Likelihood Estimator ».

$$\begin{aligned} \ell(\boldsymbol{\theta}, \boldsymbol{\Sigma}_y | \mathfrak{E}) = & -\frac{n_y n_{pm}}{2} \ln(2\pi) - \frac{n_{pm}}{2} \ln(\det \boldsymbol{\Sigma}_y) \\ & - \frac{1}{2} \sum_{k=1}^{n_{pm}} [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})]^T \boldsymbol{\Sigma}_y^{-1} [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})] \end{aligned} \quad (1.18)$$

En utilisant les conditions d'optimalité du premier ordre et les équations de dérivation, il est possible de définir un critère⁶ ne dépendant plus de $\boldsymbol{\Sigma}_y$ comme l'ont montré Goodwin et Payne (1977) :

$$j_{mle}(\boldsymbol{\theta} | \mathfrak{E}) = \ln \det \left\{ \sum_{k=1}^{n_{pm}} [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})] [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})]^T \right\} \quad (1.19)$$

On peut alors retrouver $\boldsymbol{\Sigma}_y$ par :

$$\hat{\boldsymbol{\Sigma}}_y = \frac{1}{n_{pm}} \sum_{k=1}^{n_{pm}} [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})] [\mathbf{y}_k - f(\boldsymbol{\xi}_k, \boldsymbol{\theta})]^T \quad (1.20)$$

Ce critère implique que toutes les grandeurs mesurées le soient de manière synchrone, ce qui veut dire qu'il ne doit pas avoir de données manquantes. Si c'est le cas, en ajoutant l'hypothèse que la matrice $\boldsymbol{\Sigma}_y$ est diagonale, l'estimation des paramètres peut se faire en minimisant le critère suivant :

$$j_{mle}(\boldsymbol{\theta} | \mathfrak{E}) = \sum_{i=1}^{n_y} \frac{n_{pm_i}}{2} \ln \left\{ \frac{2\pi}{n_{pm_i}} \sum_{k=1}^{n_{pm_i}} [y_{ki} - f_i(\boldsymbol{\xi}_k, \boldsymbol{\theta})]^2 \right\} \quad (1.21)$$

où n_{pm_i} est le nombre de points de mesure de la $i^{\text{ème}}$ grandeur observable.

1.1.4.3 Difficultés et précautions

À la recherche des paramètres doit précéder une analyse d'estimabilité. En effet, il n'est pas rare que certains paramètres du modèle n'aient que très peu, voire pas du tout d'influence sur la prédiction des grandeurs observables comme l'ont montré Walter et Pronzato (1990). C'est souvent le cas pour des modèles phénoménologiques complexes faisant intervenir de nombreux phénomènes. Pour identifier les paramètres qui focaliseront l'attention, le lecteur pourra se référer aux travaux de Yao (2003) et de Jacques (1985) qui proposent des méthodes d'analyse d'estimabilité.

Une autre difficulté est la corrélation entre deux paramètres. Pour illustrer le problème, prenons un exemple extrême très simple, soit $f(\boldsymbol{\xi}, \boldsymbol{\theta}) = (\theta_1 + \theta_2) \boldsymbol{\xi}$. Ici, le mo-

⁶ On remarque au passage que s'il n'y a qu'une seule grandeur mesurée, le critère de maximum de vraisemblance équivaut au critère quadratique décrit plus haut.

dèle est une droite par rapport à ξ , on pourra donc estimer la pente de la droite (la somme des paramètres) sans difficulté, mais on ne pourra pas discerner θ_1 de θ_2 . Dans cet exemple, l'indiscernabilité des paramètres est liée directement à la structure même du modèle, il faudra simplifier celui-ci. Dans d'autres cas, la corrélation entre les paramètres peut venir du manque de données disponibles, il faudra envisager d'acquérir plus d'information, autrement dit, de faire plus d'expériences.

Enfin, il reste une difficulté numérique pour l'estimation des paramètres. En effet, les routines d'optimisation fonctionnent très efficacement pourvu que les valeurs des paramètres recherchés soient proches de l'unité et que les dérivées partielles du critère à optimiser par rapport aux paramètres soient de même ordre de grandeur (il ne faut pas qu'un paramètre soit de plusieurs ordres de grandeur plus sensible qu'un autre). Par exemple, la loi d'Arrhenius pose souvent problème. En effet, l'énergie d'activation Ea intervient dans une exponentielle, le modèle en est donc très sensible, alors que le facteur pré-exponentiel A_0 intervient linéairement. De plus, les ordres de grandeur de ces deux paramètres sont très différents (de 0 à 400 kJ/mol pour l'énergie d'activation et de 10^0 à 10^{20} SI pour le facteur pré-exponentiel). Pour rendre la tâche plus facile à la routine d'optimisation, on peut utiliser des techniques de reparamétrisation (Agarwal et Brisk 1985), ainsi, au lieu de chercher Ea et A_0 on peut rechercher deux autres paramètres θ_1 et θ_2 tel que :

$$k(T) = A_0 \exp\left\{-\frac{Ea}{RT}\right\} = \exp\{\theta_1\} \exp\left\{-\left(\frac{1}{T} - \frac{1}{T^*}\right)\right\} \exp\{\theta_2\} \quad (1.22)$$

avec T^* une température de référence arbitrairement choisie, $\theta_1 = \ln(k(T^*))$ soit le logarithme de la vitesse de réaction à la température de référence et $\theta_2 = \ln(Ea/R)$ la pente de la droite que suit le logarithme de la constante de vitesse en fonction de l'inverse de la température. Ainsi, les valeurs des paramètres recherchés se trouvent en valeur absolue dans une gamme assez restreinte, au plus quelques dizaines d'unités, et leurs sensibilités deviennent similaires. Bien que cette technique rende l'estimation des paramètres plus aisée numériquement, elle ne permet cependant pas d'obtenir une meilleure précision sur l'énergie d'activation et le facteur pré-exponentiel comme l'ont souligné MM. Rippin et Rimensberger (1986).

Une autre reparamétrisation très efficace est possible pour faciliter la mise en œuvre numérique de l'estimation des paramètres. L'idée est de définir deux températures (souvent prises aux extrémités du domaine de température étudié T_{min} et T_{max}) et de chercher les constantes de vitesses à chacune de ces deux températures (k_{min} et k_{max}). Pour éviter le problème de l'ordre de grandeur on prend souvent le logarithme des constantes de vitesse. On obtient ainsi :

$$\ln k(T) = \ln k_{min} + \frac{1/T - 1/T_{min}}{1/T_{max} - 1/T_{min}} (\ln k_{max} - \ln k_{min}) \quad (1.23)$$

Les deux paramètres jouent le même rôle, ils ont donc des sensibilités équivalentes. De plus, lors de l'estimation des paramètres, il est beaucoup plus facile de fixer les bornes du domaine de recherche Θ puisque l'on peut relier aisément les constantes de vitesses au temps caractéristiques de la réaction et donc avoir un ordre de grandeur de k_{min} et k_{max} .

1.1.5 La performance

La performance s'évalue en général par plusieurs critères qui peuvent être le rendement, la consommation énergétique, la productivité,... Ces critères sont estimés à l'aide d'une fonction vectorielle γ qui dépend des conditions expérimentales et des grandeurs observables (seulement dans certains cas spécifiques des grandeurs observables peuvent être confondues avec des performances). Si le calcul des performances est basé sur les grandeurs observables mesurées, alors on parlera de performances expérimentales. Si au contraire, les grandeurs observables sont prédites par le modèle, on parlera alors de performances estimées (voir équations (1.5) et (1.6)).

1.1.5.1 Région de confiance

Dans les deux cas, il est intéressant de connaître l'erreur commise sur le calcul des performances, où la matrice de covariance qui la caractérise Σ_j . Cette erreur est liée à l'erreur provenant des grandeurs mesurées définies par Σ_y ($\hat{\Sigma}_j$ pour les grandeurs prédites). Pour estimer Σ_j connaissant Σ_y on peut utiliser la relation suivante :

$$\Sigma_j = \begin{bmatrix} \frac{\partial \gamma_1}{\partial y_1} & \dots & \frac{\partial \gamma_1}{\partial y_{n_y}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \gamma_{n_j}}{\partial y_1} & \dots & \frac{\partial \gamma_{n_j}}{\partial y_{n_y}} \end{bmatrix} \Sigma_y \begin{bmatrix} \frac{\partial \gamma_1}{\partial y_1} & \dots & \frac{\partial \gamma_{n_j}}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \gamma_1}{\partial y_1} & \dots & \frac{\partial \gamma_{n_j}}{\partial y_{n_y}} \end{bmatrix} = \mathbf{G}^T \Sigma_y \mathbf{G} \quad (1.24)$$

où \mathbf{G} est la matrice jacobienne de la fonction de performance γ par rapport à \mathbf{y} . Il est important de noter que le calcul des erreurs de performance est indispensable pour la fiabilité de l'optimisation du procédé.

Une fois que l'on a évalué les critères de performances, il reste une étape de choix car, il n'est pas rare que certains critères soient antagonistes. Par exemple si l'on améliore la conversion, ce sera probablement au détriment de la sélectivité et vis-et-versa. Il existe une frontière infranchissable, appelée « front de Pareto », qui représente l'ensemble des meilleurs compromis possibles. En d'autres termes, lorsque l'on tente d'améliorer un seul critère d'un point situé sur le front de Pareto, cela entraîne automatiquement une perte sur au moins un autre critère.

Outre le problème des compromis, il y a le problème des spécifications. Les spécifications sont des contraintes de performances que l'on doit respecter obligatoirement (par exemple, la concentration en réactif restant). La Figure 1.8 est une illustration du front de Pareto et des spécifications. Il faut noter que l'objectif se situe en bas à droite ;

en effet, par convention, nous cherchons toujours à minimiser le critère (on minimise ici l'opposé de la conversion et de la sélectivité).

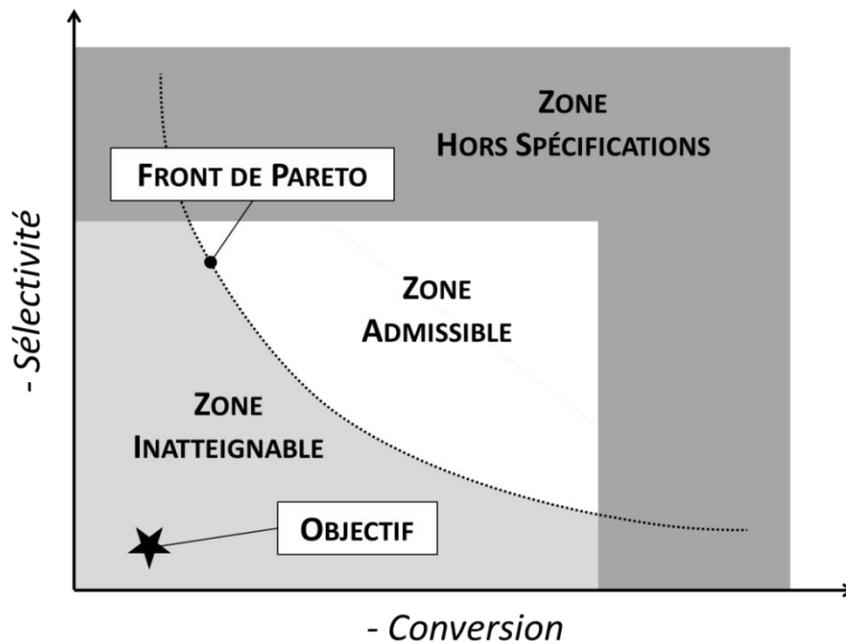


Figure 1.8 : Illustration du front de Pareto et des contraintes de spécifications.

Ainsi, l'optimum de performance d'un procédé est donc souvent un compromis à trouver. Pour trouver le meilleur compromis tout en respectant les spécifications deux types de méthodes peuvent se trouver dans la littérature. La première est basée sur l'agrégation *a priori* de plusieurs critères en un seul. Elle a le désavantage de faire intervenir la subjectivité de l'utilisateur avant de connaître ce qui est possible (la zone admissible sur la Figure 1.8). La seconde, a contrario, essaie de trouver tous les compromis possibles, puis a posteriori de classer les différents scénarii (alternatives, ou encore compromis) les uns par rapport aux autres.

1.1.5.2 Les méthodes *a priori*

Les méthodes *a priori* passent par l'agrégation des critères en un seul (la performance est caractérisée par une unique note). L'intérêt est de rendre l'optimisation beaucoup plus facile car les méthodes classiques d'optimisation numérique peuvent s'appliquer aisément pour trouver les conditions expérimentales qui améliorent au mieux le critère de performance. La méthode d'agrégation la plus intuitive est la pondération des critères. L'inconvénient majeur de la pondération est le choix des poids (1.25). De plus, un tel critère ne tient pas compte des spécifications.

$$j_{perf} = \sum_{i=1}^{n_j} \omega_i \gamma_i(\xi, \mathbf{y}) \quad (1.25)$$

Le critère présenté ci-dessous (1.26) permet de tenir compte des spécifications. La Figure 1.9 schématise la construction de cette manière d'agréger.

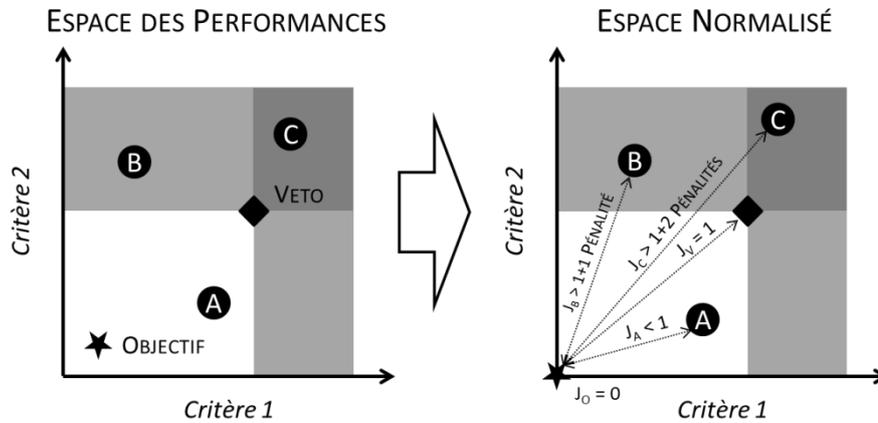


Figure 1.9 : Exemple d'agrégation de critères de performances.

L'idée réside dans le fait d'ajouter des pénalités au critère lorsque celui-ci ne respecte pas sa spécification. Quant à la pondération des critères, elle se fait automatiquement. En effet, après avoir défini un point objectif (qui peut être inatteignable) et un point veto (définissant les spécifications), la valeur du critère est normalisée entre l'objectif et le veto de sorte que, si le critère est satisfait, le critère normalisé est compris entre zéro et un. Si la spécification du critère n'est pas satisfaite, alors il sera supérieur à un. Si au contraire, un objectif est dépassé, alors la valeur ne pourra pas être négative et sera donc égale à zéro. Ensuite, l'agrégation se fait en prenant la distance à l'origine du point dans l'espace des critères normalisés. Cette distance est normalisée par la racine carrée du nombre de critères pour que le point veto ait une note égale à un. Enfin, pour éviter que des points ne respectant pas tous les critères aient quand même une note supérieure à un, un point de pénalité est ajouté pour chaque spécification non respectée. L'équation suivante formalise l'expression de ce critère :

$$j_{perf} = \frac{1}{\sqrt{n_j}} \sqrt{\sum_{i=1}^{n_j} (\gamma'_i)^2} + \sum_{i=1}^{n_j} p_i \quad (1.26)$$

$$\text{Avec } \gamma'_i = \begin{cases} \gamma_i - o_i & \text{si } \gamma_i \geq o_i \\ v_i - o_i & \text{sinon} \end{cases} \quad \text{et } p_i = \begin{cases} 0 & \text{si } \gamma_i \leq v_i \\ 1 & \text{sinon} \end{cases}$$

L'inconvénient des méthodes a priori est de faire intervenir la subjectivité du décideur avant même que l'ensemble des meilleures alternatives possibles ne soit mis au jour. Pour pallier ce défaut, des méthodes dites *a posteriori* ont été développées.

1.1.5.3 Les méthodes a posteriori

Le but des méthodes *a posteriori* est de trouver l'ensemble des meilleurs compromis possibles (le front de Pareto) puis de décider quel compromis (ou alternative) correspond au mieux au choix du décideur.

Pour trouver le front de Pareto, il existe plusieurs méthodes numériques d'optimisation multicritères dont les plus utilisées sont l'optimisation par essais particuliers (Particule Swarm Optimization, en anglais) (Reyes-Sierra et Coello 2006) et les algorithmes évolutionnaires (Coello, Lamont, et Van Veldhuizen 2007). L'optimisation multicritère fonctionne de manière assez similaire à l'optimisation classique monocritère, à ceci près, que la solution trouvée n'est pas un unique point mais un ensemble de points. L'ensemble des points solutions est donc une approximation du front de Pareto qui lui comporte une infinité de points. Une fois le front de Pareto approximé, seules quelques alternatives significatives sont conservées puis classées grâce à des méthodes de prise de décision multicritères. Nous présentons ici seulement deux méthodes : MAUT et PROMETHEE. Pour plus d'information sur les méthodes de prise de décision multicritère, on peut se reporter au très bon état de l'art dans le livre de Greco (2004) et à des exemples d'application dans les travaux de Pohekar (2004).

La méthode MAUT (pour Multi-Attribute Utility Theory) est basée sur l'agrégation des critères en un seul. La note unique attribuée à chaque individu prend en compte l'intérêt et la tolérance que l'on a vis-à-vis de chaque critère. Ces deux aspects sont quantifiés par la fonction suivante qui associe une note U à chaque individu \mathbf{a} (j_i^{min} et j_i^{max} sont respectivement la meilleure et la pire alternatives vis-à-vis du critère i) :

$$U(\mathbf{a}) = \sum_{i=1}^{n_j} \omega_i u_i(\mathbf{a}) \quad (1.27)$$

$$\text{avec } u_i(\mathbf{a}) = \left(\frac{y_i(\mathbf{a}) - j_i^{min}}{j_i^{max} - j_i^{min}} \right)^{\alpha_i}$$

L'intérêt que l'on porte au critère i est représenté par le poids ω_i et la tolérance par α_i . La Figure 1.10 montre comment α_i impact sur la fonction d'utilité u_i .

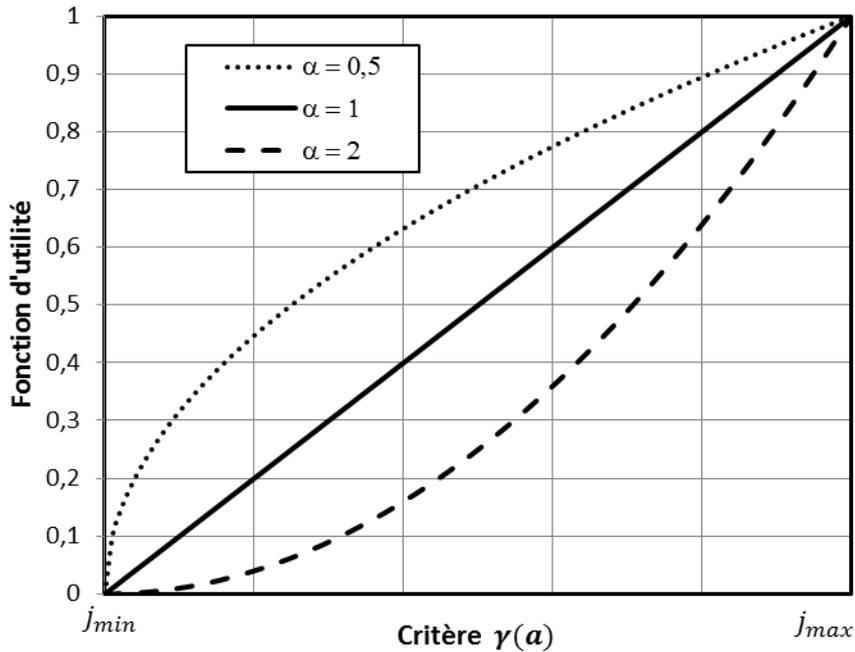


Figure 1.10 : Fonction d'utilité de la méthode d'agrégation MAUT.

La méthode PROMETHEE (pour Preference Ranking Organization Method for Enrichment Evaluation) est une méthode de surclassement développée par une équipe de l'université de Bruxelles (Brans, Vincke, et Mareschal 1986). Elle permet de comparer les différentes alternatives les unes par rapport aux autres. Toutes les alternatives sont comparées deux-à-deux sur chacun des n_j critères :

$$d_i(\mathbf{a}, \mathbf{b}) = \gamma_i(\mathbf{a}) - \gamma_i(\mathbf{b}) \tag{1.28}$$

La méthode utilise alors une fonction de préférence strictement croissante. Habituellement, cette fonction renvoie 0 lorsque la différence est négative ou inférieure à un seuil d'indifférence q_j . Si la différence dépasse le seuil de préférence p_j , on la considère comme significative et la fonction renvoie 1. Entre les deux seuils, la fonction fait une interpolation linéaire. Il existe d'autre type de fonctions de préférences qui permettent par exemple de traiter des critères discrets (comme bien, mauvais, très mauvais,...).

$$P_i(x) \begin{cases} 0 & \text{si } x \leq q_i \\ \frac{x - q_i}{p_i - q_i} & \text{si } q_j < x < p_i \\ 1 & \text{si } x \geq p_i \end{cases} \tag{1.29}$$

Puis, l'on calcule pour chaque couple (\mathbf{a}, \mathbf{b}) la fonction de préférence globale, qui est définie comme étant la somme pondérée des préférences par rapport à chaque critère :

$$\pi(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n_j} \omega_i \cdot P_i(d_i(\mathbf{a}, \mathbf{b})) \tag{1.30}$$

Enfin, on calcule deux fonctions, le flux positif et le flux négatif définis par :

$$\Phi^+(\mathbf{a}) = \frac{1}{n-1} \sum_{\mathbf{b} \in A} \pi(\mathbf{a}, \mathbf{b}) \quad (1.31)$$

$$\Phi^-(\mathbf{a}) = \frac{1}{n-1} \sum_{\mathbf{b} \in A} \pi(\mathbf{b}, \mathbf{a}) \quad (1.32)$$

Le score global, appelé flux net, est la différence du flux positif et du flux négatif.

$$\Phi(\mathbf{a}) = \Phi^+(\mathbf{a}) - \Phi^-(\mathbf{a}) \quad (1.33)$$

L'intérêt principal de cette méthode réside dans les critères d'indifférences et de préférences qui sont des grandeurs réelles contrairement aux poids et aux tolérances requises par la méthode MAUT. Par ailleurs, on peut définir des flux nets pour chacun des critères, que l'on appelle flux net unicritère, tel que :

$$\Phi_i(\mathbf{a}) = \frac{1}{n-1} \sum_{\mathbf{b} \in A} \{P_i(d_i(\mathbf{a}, \mathbf{b})) - P_i(d_i(\mathbf{b}, \mathbf{a}))\} \quad (1.34)$$

En appliquant une analyse en composante principale à l'ensemble des alternatives dans l'espace des flux net unicritères, on peut obtenir le plan GAIA (pour Geometrical Analysis for Interactive Aid) qui permet une visualisation aisée des compromis.

1.1.6 En résumé

Voici un résumé de l'ensemble des variables utilisées dans le formalisme décrit ci-dessus :

- ξ_k est un vecteur de n_x conditions expérimentales au point de mesure k ,
- y_k est un vecteur de n_y grandeurs observables obtenues dans les conditions ξ_k ,
- θ est un vecteur de n_p paramètres,
- f est une fonction vectorielle qui représente la structure du modèle, aussi appelée architecture,
- ϵ est un vecteur de n_y erreurs expérimentales obtenues dans les conditions ξ_k ,
- γ est une fonction vectorielle de n_j critères qui permet le calcul des performances.
- Σ_y est une $(n_y \times n_y)$ matrice de covariance de l'erreur expérimentale,
- Σ_f est une $(n_y \times n_y)$ matrice de covariance de l'erreur de prédiction du modèle,
- Σ_j est une $(n_j \times n_j)$ matrice de covariance de l'erreur de calcul de la performance,

- σ_{ij}^2 est l'élément de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne de Σ_y ,
- $\Xi = \{\xi_1, \dots, \xi_p, \dots, \xi_{n_{pm}}\}$ est l'ensemble de n_e échantillons,
- Δ_r est la fenêtre expérimentale du réacteur r ,
- Ω_r est l'espace expérimental du réacteur r ,
- Θ est le domaine de recherche des paramètres.

Toutes les variables décrites dans ce paragraphe s'adaptent à n'importe quel modèle. Mais, selon le choix de la structure du modèle, la manière d'utiliser ces variables est complètement différente. Ainsi, à chaque type de structure de modèle existe un type d'approche particulier. La première approche utilise des modèles de tendance qui n'ont aucun sens physique mais qui présentent des propriétés mathématiques intéressantes. La deuxième approche se base sur des modèles dits phénoménologiques, car ils s'appuient sur la description de phénomènes physiques. Enfin la dernière approche ouvre la possibilité de modifier la structure du modèle en cours de route.

1.2 Approche tendancielle

Les approches tendanciennes sont basées sur l'idée que le système à optimiser est trop complexe pour être modélisé par un modèle phénoménologique et qu'un modèle de tendance suffit à trouver l'optimum du système. Le modèle de tendance permet de décrire l'influence d'un facteur ou d'une combinaison de facteurs sur les performances du système. Les facteurs sont les leviers d'action que l'utilisateur peut modifier, ce peut être par exemple la température, la concentration, la pression,... La première phase de cette approche consiste à cibler parmi tous les facteurs potentiellement intéressants ceux qui ont de l'importance sur les performances, elle est nommée « criblage ». Ensuite, intervient une phase d'affinement, dite de modélisation, qui permet de quantifier les interactions potentielles entre différents facteurs. Enfin, lors de la phase d'optimisation, seul un nombre restreint de facteurs est conservé pour les ajuster du mieux possible.

Ce type d'approche utilise des modèles interpolateurs qui n'ont pas de sens physique. Les modèles utilisés sont des modèles faciles de manipulation, en général ils sont linéaires par rapport aux paramètres (modèles LP). L'équation (1.35) présente des exemples de modèles LP et non-LP :

$$\begin{aligned} f(\xi, \theta) &= \theta_1 \xi^2 + \theta_2 \xi + \theta_3 \text{ est LP} \\ g(\xi, \theta) &= \theta_1 \xi + \ln(\theta_2 + \xi) \text{ est non-LP} \end{aligned} \tag{1.35}$$

Le fait que la structure des modèles utilisés soit LP permet une estimation aisée des paramètres. De plus, la planification d'expériences est relativement simple et les plans d'expériences à appliquer sont souvent prêts à l'emploi. Par ailleurs, l'estimation de l'incertitude des paramètres est également facilitée.

1.2.1 Estimation des paramètres

L'écriture généralisée des modèles les plus couramment utilisés est la suivante :

$$y = \theta_0 + \sum \theta_i \xi_i + \sum \theta_{ij} \xi_i \xi_j + \sum \theta_{ii} \xi_i^2 + \epsilon \quad (1.36)$$

où y est la grandeur observable par l'expérimentateur. Lorsque plusieurs grandeurs observables doivent être modélisées, on écrira un modèle séparément pour chaque grandeur observable. Les ξ_i sont les facteurs que l'expérimentateur peut faire varier et dont le système semble dépendre *a priori*, ϵ est l'erreur commise lors de la mesure, $\theta_0, \theta_i, \theta_{ij}$ et θ_{ii} sont respectivement les coefficients des termes constants, linéaires, d'interactions et quadratiques du modèle.

Le vecteur \mathbf{y} de dimension $(n_{pm} \times 1)$ est donc l'ensemble des mesures obtenues en chaque point de mesures de la base de données expérimentale Ξ :

$$\mathbf{y} = \Phi(\Xi) \boldsymbol{\theta} + \epsilon \quad (1.37)$$

où Φ représente une fonction vectorielle qui renvoie n_t termes. Par exemple⁷ :

$$\Phi(\Xi) = [\phi_1(\Xi), \phi_2(\Xi), \phi_3(\Xi), \phi_4(\Xi), \phi_5(\Xi)] \quad (1.38)$$

Souvent, la fonction Φ renvoie une valeur normalisée comprise en -1 et 1 de manière à pouvoir comparer deux facteurs expérimentaux qui varient sur des gammes complètement différentes.

Vu l'expression du modèle, $\boldsymbol{\theta}$ est donc un vecteur de dimension $(n_t \times 1)$. Pour obtenir $\boldsymbol{\theta}$ grâce à une base de données expérimentale Ξ contenant n_{pm} échantillons, on utilise le maximum de vraisemblance sous forme matricielle :

$$\mathcal{L}(\boldsymbol{\theta}|\Xi) = \sqrt{2\pi\sigma_y^2} \exp\left\{-\frac{1}{2\sigma_y^2} [\mathbf{y} - \Phi\boldsymbol{\theta}]^T [\mathbf{y} - \Phi\boldsymbol{\theta}]\right\} \quad (1.39)$$

Avec \mathbf{y} le vecteur de données expérimentales recueillies dans les conditions Ξ , σ_y l'écart-type supposé connu et

$$\Phi = \Phi(\Xi) = \begin{bmatrix} \phi_1(\xi_1) & \cdots & \phi_{n_t}(\xi_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\xi_{n_{pm}}) & \cdots & \phi_{n_t}(\xi_{n_{pm}}) \end{bmatrix} \quad (1.40)$$

⁷ Dans l'équation (1.36) la fonction Φ renvoie un terme constant, un terme linéaire par rapport à chaque facteur, un terme quadratique par rapport au premier facteur et un terme d'interaction.

Au maximum de vraisemblance, les dérivées partielles de la log-vraisemblance par rapport au vecteur de paramètres $\boldsymbol{\theta}$ s'annulent :

$$\left. \frac{\partial \ell}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = -\frac{1}{\sigma_y^2} \boldsymbol{\Phi}^T [\mathbf{y} - \boldsymbol{\Phi} \boldsymbol{\theta}] = \mathbf{0} \quad (1.41)$$

On en déduit, sous réserve que $\boldsymbol{\Phi}^T \boldsymbol{\Phi}$ s'inverse, la valeur estimée du vecteur de paramètres :

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \mathbf{y} \quad (1.42)$$

1.2.2 Estimation de la région de confiance

On peut aussi déterminer analytiquement la région de confiance des paramètres. Le maximum de vraisemblance est un estimateur non biaisé de $\boldsymbol{\theta}^*$ tel que si l'on pouvait répéter un nombre infini la même expérience, l'estimée des paramètres serait égale à la vraie valeur :

$$\mathbb{E}_{\mathbf{y}|\boldsymbol{\theta}^*}(\hat{\boldsymbol{\theta}}_{mle}) = \boldsymbol{\theta}^* \quad (1.43)$$

Soit $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ la matrice de covariance de cet estimateur. Puisqu'il est non biaisé, on a :

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathbb{E}_{\mathbf{y}|\boldsymbol{\theta}^*} \left([\hat{\boldsymbol{\theta}}_{mle} - \boldsymbol{\theta}^*][\hat{\boldsymbol{\theta}}_{mle} - \boldsymbol{\theta}^*]^T \right) \quad (1.44)$$

D'après l'inégalité de Cramér et Rao (Radhakrishna Rao 1945), on sait que :

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} \geq \mathbf{F}^{-1}(\boldsymbol{\theta}^*) \quad (1.45)$$

où \mathbf{F} est la matrice d'information de Fisher qui est donnée par :

$$\begin{aligned} \mathbf{F}(\boldsymbol{\theta}) &= \mathbb{E}_{\mathbf{y}|\boldsymbol{\theta}^*} \left(\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\boldsymbol{\Xi}) \right] \left[\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\boldsymbol{\Xi}) \right]^T \right) \\ &= - \mathbb{E}_{\mathbf{y}|\boldsymbol{\theta}^*} \left(\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \ell(\boldsymbol{\theta}|\boldsymbol{\Xi}) \right) \end{aligned} \quad (1.46)$$

Pour rappel, $\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\boldsymbol{\Xi})$ et $\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \ell(\boldsymbol{\theta}|\boldsymbol{\Xi})$ sont le gradient et le Hessian de la log-vraisemblance. En dérivant l'équation (1.46) par rapport à $\boldsymbol{\theta}$, on obtient :

$$\mathbf{F}(\boldsymbol{\theta}) = -\frac{1}{\sigma_y^2} \boldsymbol{\Phi}^T \boldsymbol{\Phi} \quad (1.47)$$

On peut montrer que, dans le cas où le modèle est LP et que le bruit est normal à variance stationnaire, l'estimateur du maximum de vraisemblance devient efficace, c'est à dire que :

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathbf{F}^{-1}(\boldsymbol{\theta}^*) \quad (1.48)$$

En utilisant le test de Fisher, on obtient l'équation de l'hypersurface définissant la région de confiance qui dans ce cas, a la forme d'un ellipsoïde :

$$[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}]^T \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} [\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}] = F_{n_p; n_{pm} - n_p} (1 - \alpha) \quad (1.49)$$

où α est l'indice de confiance souvent choisi à $\alpha = 5\%$.

1.2.3 Construction de la structure

La manière dont est définie la fonction Φ décrit totalement la structure du modèle. À chaque étape, criblage, modélisation et optimisation, doit donc être définie une structure appropriée basée sur les deux considérations suivantes :

- Comme nous l'avons vu au paragraphe 1.1.2, la complexité du modèle ne doit pas être trop grande car le caractère prédictif du modèle s'en trouverait diminué. La structure du modèle doit donc être adaptée aux seuls effets les plus influents, les termes d'interactions et les termes quadratiques doivent être en nombre limité.
- La seconde considération à prendre en compte est le nombre d'échantillons à générer. Ce nombre doit être au moins égal au nombre de paramètres, qui est lui-même égal au nombre de variables du modèle. Si l'on souhaite faire seulement quatre expériences, on ne pourra tester au plus que trois facteurs (un paramètre pour chaque facteur et un pour la constante). Le Tableau 1.2 montre comment évolue le nombre de paramètres par type de termes en fonction du nombre de facteurs considérés :

Tableau 1.2 : Évolution du nombre des différents termes en fonctions du nombre de facteurs.

Nombre de facteurs n_u	Linéaires ou Quadratiques	Interactions doubles	Toutes les interactions
<i>expression</i>	n_u	$\frac{n_u(n_u - 1)}{2}$	$2^{n_u} - n_u$
2	2	1	1
3	3	3	4
4	4	6	11
5	5	10	26
6	6	15	57
7	7	21	120
8	8	28	247

Ainsi, généralement, lorsque le nombre de facteurs pris en compte dans le plan d'expériences est supérieur à 10, on évitera de faire intervenir des termes d'interaction. Seuls des termes linéaires seront pris en compte. On parle de « *plan de criblage* ». À l'issue du plan de criblage, il est possible de déterminer quels sont les facteurs les plus influents sur le comportement du système. On éliminera ainsi les facteurs les moins influents afin d'affiner le modèle. Ayant moins de facteurs, on pourra prendre en compte des termes d'interactions. Ainsi, lorsque le nombre de facteurs étudiés est compris entre 4 et 10 on peut utiliser un modèle contenant certains termes d'interactions. Cette étape est l'étape de « *modélisation* ». À l'issue de la modélisation, très souvent, seuls trois ou quatre facteurs sont réellement significatifs et on pourra alors introduire des termes quadratiques. Les modèles ayant des termes quadratiques ne doivent pas dépasser plus de trois ou quatre facteurs pour conserver un nombre restreint d'expériences à faire.

L'objectif de procéder par étape est de limiter le nombre d'expériences totales. Ainsi, au fur et à mesure des étapes, on se concentre sur les termes importants. À chacune de ces trois étapes, l'objectif est de ne pas dépasser plus d'une vingtaine d'expériences. Si, dès le départ, le nombre de facteurs est inférieur à dix, on peut directement se tourner vers un plan de modélisation, de même s'il est inférieur ou égal à quatre, on utilisera directement un plan d'optimisation.

1.2.4 Planification d'expériences

En fonction de la nature du modèle choisi, donc des termes utilisés, le choix des expériences à réaliser est différent. C'est là que le choix d'utiliser des modèles linéaires aux paramètres à toute son importance. En effet, puisqu'il est possible d'estimer analytiquement la valeur des paramètres et leur incertitude en fonction des expériences réalisées (1.48), il faut choisir judicieusement à l'avance les expériences à faire pour minimiser l'incertitude des paramètres et ainsi rendre plus fiable le modèle.

Les trois principaux critères sur lesquels est basée la planification d'expériences pour les modèles linéaires aux paramètres sont :

- Le critère de **O-optimalité**. O pour « *Orthogonal* ». Lorsque la matrice Φ est orthogonale alors la matrice $\Phi^T\Phi$ est une matrice diagonale, donc la matrice de covariance des paramètres estimés l'est aussi. Ainsi, les paramètres seront à coup sûr décorrélés : on estimera chacun des paramètres de manière indépendante.
- Le critère de **D-optimalité**. D pour "*Determinant*". Ce critère cherche à minimiser la variance de tous les paramètres en même temps, cela revient à maximiser le déterminant de la matrice $\Phi^T\Phi$.
- Le critère de **G-optimalité**. G pour "*General Variance*". L'objectif de ce critère est trouver le plan d'expériences qui permet de

répartir au mieux l'erreur de prédiction du modèle sur l'ensemble du domaine étudié. L'idée est donc de calculer la variance de prédiction en chacun des points de mesure du plan d'expériences et d'en minimiser la plus grande. Concrètement, cela revient à minimiser le plus grand terme de la diagonale de la matrice $\Phi[\Phi^T\Phi]\Phi^T$.

On peut montrer que les plans obtenus par la matrice de Hadamard sont les plans optimaux (vis-à-vis la O- et de la D- optimalité) pour les modèles ne contenant que des termes linéaires. Lorsque l'on introduit des termes d'interaction, les plans factoriels à deux niveaux sont les plus adaptés. Lorsque le modèle prend en compte des termes quadratiques, on utilisera des plans de Box-Behnken conçus pour satisfaire le critère G-optimal ou des plans composites centrés qui favorisent le critère d'orthogonalité. Comme son nom l'indique, les plans composites sont la superposition de plusieurs plans, ainsi lorsqu'un plan factoriel a déjà été réalisé, il est possible de réutiliser certaines expériences déjà faites. Par ailleurs, il existe des plans introduisant un caractère aléatoire permettant de balayer de manière homogène le domaine expérimental, le plus utilisé est le plan latin (McKay, Beckman, et Conover 1979). De manière générale, le nombre d'expériences à réaliser par rapport au nombre de paramètres à estimer varie dans un rapport de 1,5 à 2,7 selon le type de plan utilisé.

Dans les cas où la fenêtre étudiée n'est pas un hypercube, ce qui peut être le cas lorsqu'il y a des contraintes opératoires, on utilisera des méthodes permettant de trouver le plan D-optimal adapté aux contraintes (Fedorov 1972 ; Mitchell 1974 ; Cook 1980). Le Tableau 1.3 résume les approches tendanciennes :

Tableau 1.3 : Résumé des différents types de plan d'expérience en fonction du nombre de facteurs

	$n_u > 10$	$10 \geq n_u > 3$	$3 \geq n_u$
<i>Étape</i>	Criblage	Modélisation	Optimisation
<i>Type de modèle</i>	Linéaire	Linéaires et interactions	Linéaires, interactions et quadratiques
<i>Type de plan</i>	Hadamard	Factoriels	Box-Behnken, composite central, Latin

1.2.5 Exploitation des résultats

Une fois que les expériences ont été réalisées, une étape d'exploitation des résultats intervient. Pour illustrer le propos, nous nous appuyons sur un exemple très simple. Nous souhaitons modéliser le volume occupé par un gaz en fonction de la pression, de la température et du nombre de moles de gaz en utilisant un modèle de tendance (on fait comme si l'on ne connaissait pas la loi des gaz parfaits). On a donc trois facteurs, ce qui nous autorise à utiliser un modèle quadratique complet (termes linéaires, d'interaction et quadratiques). Le plan choisi est un plan de Box-Behnken et la matrice des expé-

riences avec les mesures correspondantes est donnée dans le Tableau 1.4. De plus, nous supposons que la mesure du volume est entachée d'une erreur expérimentale ayant pour écart type 1L.

Tableau 1.4 : Exemple de plan d'expériences appliqué à la mesure du volume d'un gaz en fonction de la pression, de la température et du nombre de moles.

N°	P [bar]	T [°C]	N [mol]	V [L]
1	1,0	25	0,15	1,66
2	1,0	150	0,15	4,58
3	2,0	25	0,15	1,98
4	2,0	150	0,15	1,36
5	1,0	87,5	0,1	2,93
6	1,0	87,5	0,2	5,06
7	2,0	87,5	0,1	1,05
8	2,0	87,5	0,2	1,77
9	1,5	25	0,1	1,47
10	1,5	25	0,2	3,88
11	1,5	150	0,1	3,08
12	1,5	150	0,2	6,14
13	1,5	87,5	0,15	3,69
14	1,5	87,5	0,15	2,56
15	1,5	87,5	0,15	4,17

Après normalisation de la matrice expérimentale entre -1 et 1, on estime les paramètres et leur région de confiance (voir paragraphes 1.2.1 et 1.2.2). La Figure 1.11 montre les valeurs obtenues pour chaque terme du modèle quadratique complet ainsi que leur incertitude à 95% de confiance. On remarque que certains termes ne sont pas significatifs car leur intervalle de confiance englobe zéro (interactions $P.n$, $T.n$ et quadratiques T^2 , n^2). Ces termes peuvent être supprimés du modèle afin de rendre ce dernier plus prédictif, on obtient alors la Figure 1.12.

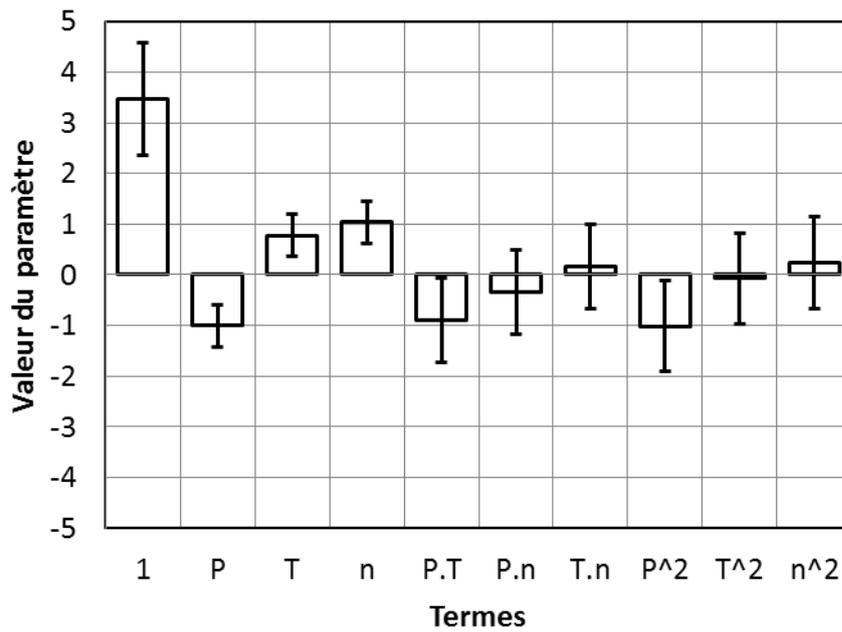


Figure 1.11 : Valeurs estimées des paramètres pour un modèle quadratique complet.

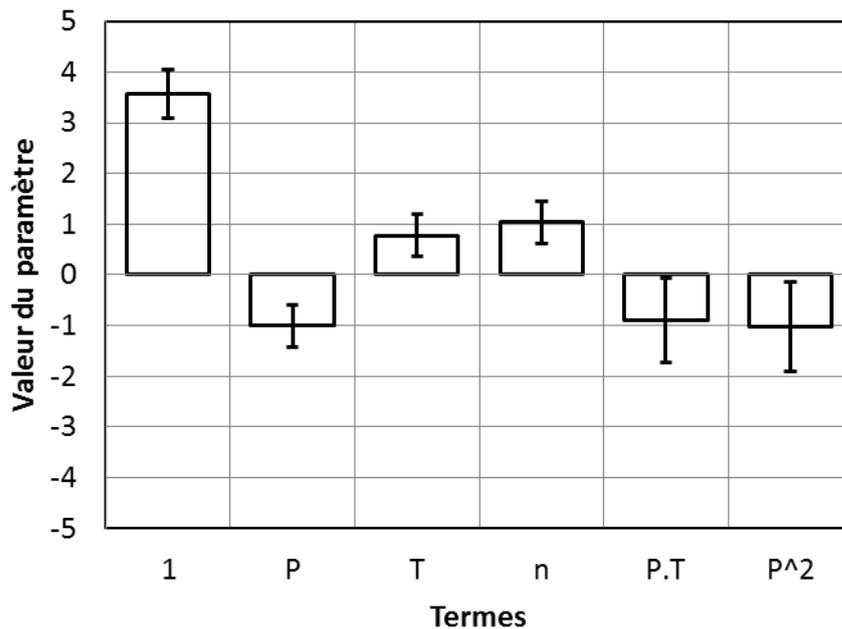


Figure 1.12 : Valeurs estimées des paramètres pour un modèle quadratique simplifié.

Le fait que l'on considère que les paramètres soient estimés avec une incertitude, implique que les prédictions du modèle ont eux aussi une incertitude. La propagation de l'incertitude des paramètres vers l'incertitude de prédiction est proportionnelle à la sensibilité du modèle vis-à-vis des paramètres. Si le modèle est peu sensible à un paramètre, même si l'incertitude de ce même paramètre est grande, la prédiction du modèle n'en sera que peu affectée. Inversement, si le modèle est très sensible à un paramètre,

alors une petite incertitude impliquera une forte incertitude de prédiction. Ainsi, on peut exprimer l'incertitude des prédictions du modèle $\hat{\Sigma}_y$ au point de mesure ξ par sa matrice jacobienne G :

$$\hat{\Sigma}_y = G \Sigma_\theta G^T \quad (1.50)$$

Or le modèle est linéaire au paramètre, donc, dans ce cas, $G = \Phi$, et on obtient :

$$\hat{\Sigma}_y = \Phi \Sigma_\theta \Phi^T \quad (1.51)$$

Il est alors possible d'estimer l'intervalle de confiance des estimées du modèle pour chaque expérience (voir Figure 1.13 rond). Il est alors possible de voir si certaines données sont aberrantes ou que le modèle ne prédit pas les observations (ce n'est pas le cas ici, aucun résidu n'est en dehors de l'intervalle de confiance expérimental en pointillé). De même, on peut voir quelles sont les zones où le modèle est le plus fiable (les trois dernières expériences numérotées de 13 à 15, qui sont situées au centre du domaine expérimental, ont un intervalle de confiance réduit, ce qui signifie que le modèle est très fiable au centre).

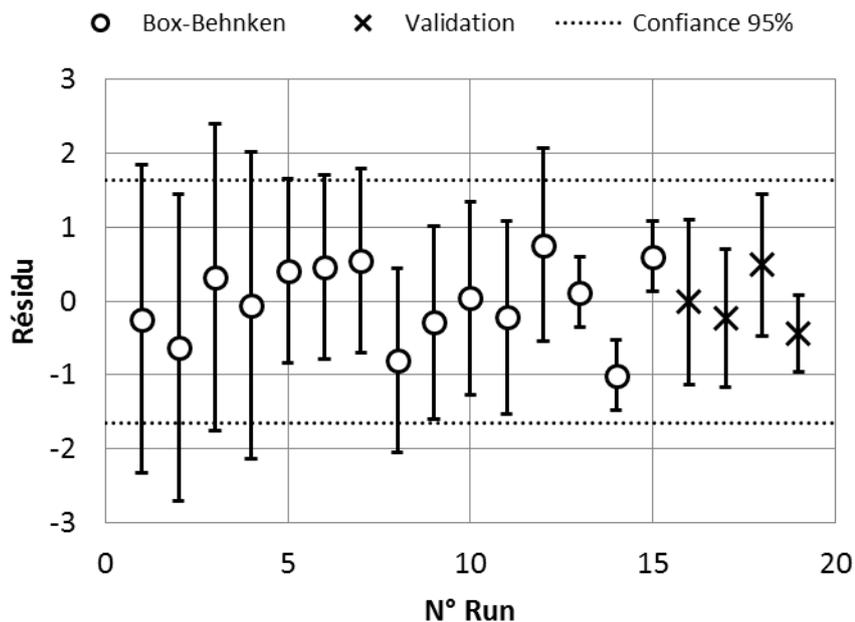


Figure 1.13 : Intervalles de confiance sur les résidus (barres d'erreur) par rapport à la confiance expérimentale (pointillés)

Enfin, la dernière des choses à vérifier avant de pouvoir utiliser le modèle est de tester des points au hasard dans le domaine expérimental afin de valider sa prédictivité. Les points de validation sont estimés avec succès puisque les résidus sont à l'intérieur de l'intervalle de confiance expérimental.

1.3 Approche phénoménologique

Contrairement aux approches tendanciennes, les approches phénoménologiques utilisent des connaissances a priori du système étudié. Le système est donc décrit par l'enchevêtrement de plusieurs phénomènes qui sont eux-mêmes décrits par des lois mathématiques. La représentation mathématique de ces phénomènes fait la plupart du temps intervenir des relations non-linéaires. Ces non-linéarités induisent une exploitation des modèles beaucoup plus complexe que vu précédemment dans la section 1.2. De plus, selon la complexité du modèle mis en œuvre, la résolution en elle-même peut être un challenge. Ainsi, il existe, dans de très rares cas, une expression explicite du modèle : le modèle est décrit par une expression analytique. Dans la plupart du temps, on ne dispose que d'une expression implicite. Sa résolution peut alors être très complexe, c'est le cas des modèles faisant intervenir des équations aux dérivées partielles. En revanche, la résolution des équations différentielles est relativement plus aisée à mettre en œuvre. Les techniques statistiques présentées dans ce paragraphe sont bien adaptées aux modèles explicites et aux modèles décrits par des équations différentielles ordinaires (ODEs).

1.3.1 Estimation des paramètres

L'introduction de non-linéarités dans le modèle rend l'estimation des paramètres beaucoup plus complexe : la résolution analytique n'étant plus possible (sauf certains cas très isolés). Pour obtenir l'estimation des paramètres on utilise des méthodes d'optimisation numérique qui permettent de minimiser un critère de vraisemblance (voir paragraphe 1.1.4).

1.3.2 Estimation de la région de confiance

De même, l'estimation de la région de confiance devient beaucoup plus ardue. Cependant plusieurs techniques ont tout de même été développées, nous ne détaillerons que l'estimation asymptotique (qui est une approximation), et l'estimation par la recherche de point appartenant à une isocritère.

Approximation de la région de confiance

Historiquement, la première idée a été développée par G.E.P. Box (Box et Lucas 1959) et consiste à faire une approximation linéaire du modèle par rapport aux paramètres dans la région qui avoisine la vraie valeur des paramètres θ^* :

$$f(\xi, \theta) \approx f(\xi, \theta^*) + (\theta - \theta^*) \left. \frac{\partial f}{\partial \theta} \right|_{\xi, \theta^*} \quad (1.52)$$

Grâce à cette approximation, on peut alors calculer la matrice d'information de Fisher :

$$\mathbf{F}(\boldsymbol{\theta}|\mathfrak{E}) = \sum_{i=1}^{n_y} \sum_{j=1}^{n_y} \tilde{\sigma}_{ij} \mathbf{Q}_i^T \mathbf{Q}_j \quad (1.53)$$

où $\tilde{\sigma}_{ij}$ est l'élément (i, j) de l'inverse de $\boldsymbol{\Sigma}_y$ et \mathbf{Q}_i la matrice de sensibilité de la $i^{\text{ème}}$ grandeur observée qui est définie par :

$$\mathbf{Q}_i = \begin{bmatrix} \left. \frac{\partial f_i}{\partial \theta_1} \right|_{\xi_1, \boldsymbol{\theta}} & \dots & \left. \frac{\partial f_i}{\partial \theta_{n_p}} \right|_{\xi_1, \boldsymbol{\theta}} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial f_i}{\partial \theta_1} \right|_{\xi_{n_{pm}}, \boldsymbol{\theta}} & \dots & \left. \frac{\partial f_i}{\partial \theta_{n_p}} \right|_{\xi_{n_{pm}}, \boldsymbol{\theta}} \end{bmatrix} \quad (1.54)$$

En faisant l'hypothèse que $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathbf{F}^{-1}$, on peut approximer la région de confiance des paramètres par l'équation suivante :

$$[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}]^T \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1} [\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}] = \chi_{n_p}^2(\alpha) \quad (1.55)$$

La région de confiance ainsi estimée doit être prise avec beaucoup de précautions car :

- lorsque l'on dispose de peu de données, il n'est pas rare que l'estimateur du maximum de vraisemblance soit biaisé, et que l'égalité $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathbf{F}^{-1}$ ne soit pas vérifiée,
- on assimile $\mathbf{F}(\boldsymbol{\theta}^*|\mathfrak{E})$ à $\mathbf{F}(\hat{\boldsymbol{\theta}}_{mle}|\mathfrak{E})$.

Région de confiance exacte

On peut estimer la région de confiance exacte par l'ensemble des points $\boldsymbol{\theta}$ qui satisfassent le test du ratio de vraisemblance :

$$\begin{aligned} & j_{mle}(\boldsymbol{\theta}^*|\mathfrak{E}) - j_{mle}(\boldsymbol{\theta}|\mathfrak{E}) \\ & \leq \frac{n_e(n_p + n_y - 1)}{n_e - n_p - n_y + 1} F_{1-\alpha; n_p+n_y-1; n_e-n_p-n_y+1} \end{aligned} \quad (1.56)$$

Pour déterminer l'ensemble des points de la région de confiance, on peut, par exemple, utiliser un algorithme d'optimisation génétique ayant comme condition d'arrêt le test du ratio de vraisemblance (Benyahia 2009). La population solution obtenue après convergence de l'algorithme donne une idée de la forme de la région de confiance.

1.3.3 Construction de la structure

Il n'existe pas de méthode systématique pour la construction de modèles phénoménologiques. C'est donc à l'utilisateur de choisir quels sont les phénomènes à décrire, quels sont ceux à négliger. De plus, l'utilisateur ne sait pas toujours comment représen-

ter mathématiquement tel ou tel phénomène. Enfin, il n'est pas rare de ne même pas connaître l'existence de certains phénomènes.

En génie chimique, pour ce qui est des réactions, la plupart des modèles de réacteurs sont basés sur des bilans de matières et des bilans enthalpiques. La Figure 1.14 montre le schéma d'un réacteur agité.

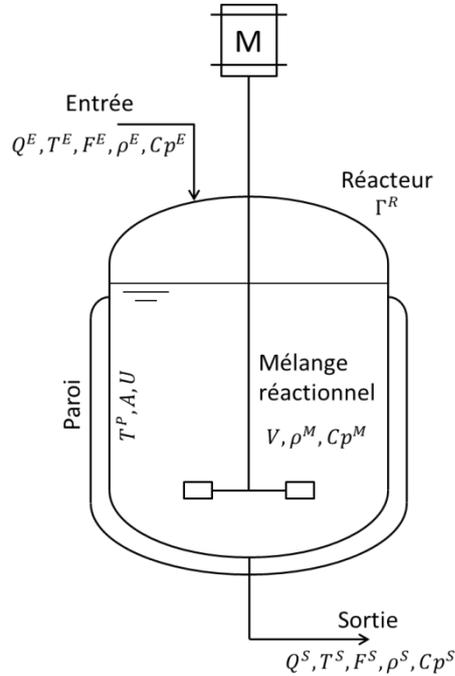


Figure 1.14 : Schéma de principe d'un réacteur cuve agitée en régime quelconque.

L'écriture généralisée des modèles de réacteur repose sur un bilan de matière qui s'écrit :

$$\frac{dn_j}{dt} = F_j^E - F_j^S + F_j^R \quad (1.57)$$

où F_j^E désigne le débit d'entrée, F_j^S le débit de sortie et F_j^R le débit de réaction du $j^{ième}$ composé. Le débit de réaction s'exprime par :

$$F_j^R = V \sum_{i=1}^{n_r} \nu_{ij} r_i \quad (1.58)$$

Le bilan enthalpique s'écrit de la manière suivante :

$$(V \rho^M C_p^M + \Gamma_R) \frac{dT}{dt} = \Phi^E - \Phi^S + \Phi^P + \Phi^R \quad (1.59)$$

où :

- $V \rho^M C_p^M$ représente la capacité thermique du mélange réactionnel,
- Γ_R la capacité thermique du corps du réacteur,

- le flux de chaleur à l'entrée est $\Phi^E = Q^E \rho^E C p^E (T^E - T)$,
- le flux de chaleur à la sortie $\Phi^S = Q^S \rho^S C p^S (T^S - T)$,
- le flux à la paroi $\Phi^P = UA(T^P - T)$,
- le flux de chaleur provenant des réactions $\Phi^R = V_R \sum_{i=1}^{NR} r_i \Delta H_i$.

Ainsi, la structure globale des modèles de réacteur est connue, exceptés deux points fondamentaux :

- La matrice de coefficients stœchiométriques v_{ij} . Cette matrice est souvent connue, mais ce n'est pas toujours le cas. Pour l'identifier il faut faire appel à la méthode TFA (Box et Hill, 1967).
- Les lois cinétiques de chaque réaction r_i sont généralement inconnues et c'est là la principale difficulté. La technique la plus utilisée est de postuler des lois cinétiques connues puis de les départager grâce à la technique de discrimination d'hypothèse (Box et Hill 1967).

Actuellement en développement, certaines techniques consistent à considérer l'identification des lois cinétiques comme un problème inverse mal-posé (Brendel, Bonvin, et Marquardt 2006). Pour cela, on réécrit le bilan de matière (1.57) sous forme matricielle :

$$\frac{d}{dt}(V(t)\mathbf{C}(t)) = Q^E(t)\mathbf{C}^E(t) - Q^S(t)\mathbf{C}(t) - \mathbf{F}^R(t) \quad (1.60)$$

avec la matrice des coefficients stœchiométriques \mathbf{N} :

$$\mathbf{N} = \begin{bmatrix} v_{1,1} & \cdots & v_{1,n_c} \\ \vdots & \ddots & \vdots \\ v_{n_r,1} & \cdots & v_{n_r,n_c} \end{bmatrix} \quad (1.61)$$

et le flux de réaction \mathbf{F}^R :

$$\mathbf{F}^R(t) = V(t) \mathbf{r}(t) \mathbf{N} \quad (1.62)$$

La vitesse de réaction est une fonction vectorielle \mathbf{m} qui dépend des concentrations $\mathbf{C}(t)$ et des paramètres $\boldsymbol{\theta}$:

$$\mathbf{r}(t) = \mathbf{m}(t, \mathbf{C}(t), T(t), \boldsymbol{\theta}) \quad (1.63)$$

Ainsi, toute la difficulté est de retrouver l'expression de la fonction \mathbf{m} à partir des flux de réaction que l'on peut retrouver en intégrant l'équation (1.60) avec $n_j = C_j V$:

$$n_j(t) = n_j(t_0) + \int_{\tau=t_0}^{\tau=t} [F_j^E(\tau) - F_j^S(\tau)] d\tau + \int_{\tau=t_0}^{\tau=t} F_j^R(\tau) d\tau \quad (1.64)$$

Soit y_j le titre molaire et étant égale à la quantité :

$$y_j(t) = n_j(t) - n_j(t_0) - \int_{\tau=t_0}^{\tau=t} [F_j^E(\tau) - F_j^S(\tau)] d\tau \quad (1.65)$$

On obtient en introduisant (1.64) dans (1.65) :

$$F_j^R(t) = \frac{dy_j(t)}{dt}, \quad y_j(t_0) = 0 \quad (1.66)$$

D'après l'équation (1.62), on aboutit finalement à :

$$\mathbf{r}(t) = V^{-1}(t)\mathbf{F}^R(t)\mathbf{N}^\dagger \quad (1.67)$$

où $\mathbf{N}^\dagger = \mathbf{N}^T(\mathbf{N}\mathbf{N}^T)^{-1}$ est la pseudo-inverse de Moore-Penrose de \mathbf{N} .

La dérivée $dy_i(t)/dt$ est estimée en interpolant par une fonction continue les titres molaires mesurés aux n_{pm} points de mesure $y_j(t_k)$, $k = 1..n_{pm}$. Pour cela, plusieurs techniques ont été mises à l'épreuve : la régularisation de Tikhonov (Mhamdi et Marquardt 1999), les splines lissantes (Craven et Wahba 1978), les réseaux de neurones (MacKay 1992) et d'autres encore. Afin de diminuer le bruit de mesure on peut utiliser une technique de réconciliation basée sur des bilans atomiques (Bonvin et Rippin 1990). Quelle que soit la technique utilisée, cela nécessite un grand nombre de mesures souvent obtenues par une analyse en ligne.

Ainsi, la démarche proposée par Brendel, Bonvin, et Marquardt (2006) est de formuler plusieurs lois cinétiques m_i pour chacune des réactions considérées et de choisir celle qui se rapproche le plus des vitesses de réactions mesurées, obtenues par l'équation **Erreur ! Source du renvoi introuvable.** Pour que le choix de l'expression de la loi cinétique soit le plus discernable possible, il est recommandé d'acquérir les mesures des vitesses de réactions dans des régimes dynamiques que l'on obtient le plus souvent en utilisant des profils de coulées, de soutirage et de température complexe.

La Figure 1.15 montre l'exemple de la cinétique de réaction entre le dihydrogène gazeux et le dibrome gazeux dont l'expression de la loi de vitesse r_A est complexe. Une simulation, que nous avons réalisée, a permis de générer des mesures expérimentales virtuelles de chacune des trois espèces. Ces concentrations ont ensuite été lissées grâce à des fonctions splines (voir Figure 1.15). La dérivée des courbes lissées a permis de remonter à la vitesse de réaction.

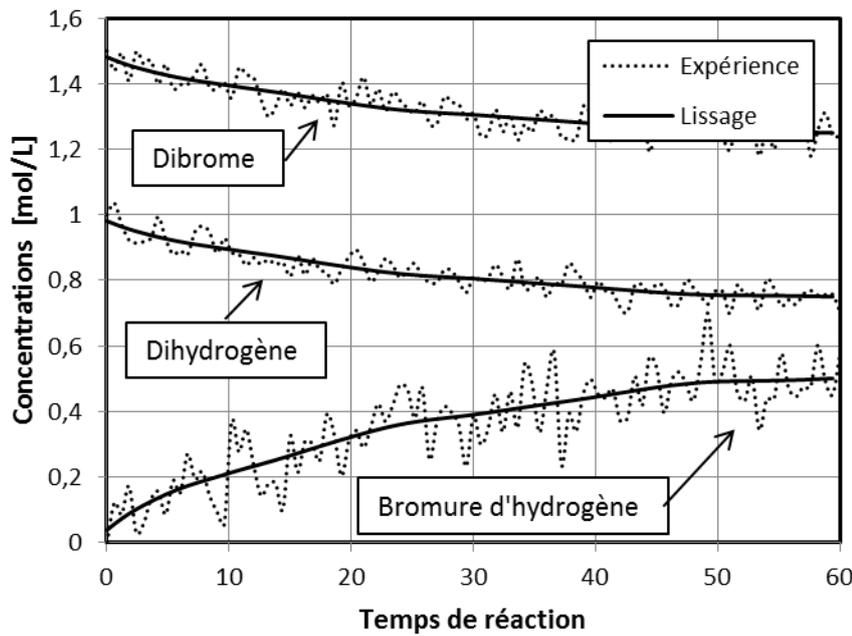


Figure 1.15 : Exemple de lissage par des fonctions splines de mesures bruitées.

Ensuite, la vitesse de réaction en fonction des concentrations de chaque espèce a été obtenue grâce aux équations présentées plus haut. Cela a permis la comparaison (voir Figure 1.16) de deux lois cinétiques A et B dont les expressions sont les suivantes :

$$\begin{aligned}
 r_A &= m_A(\mathbf{C}(t), k_A, k_A') = \frac{k_A[H_2][Br_2]^{3/2}}{[Br_2] + k_A'[HBr]} \\
 r_B &= m_B(\mathbf{C}(t), k_B) = k_B[H_2][Br_2]
 \end{aligned}
 \tag{1.68}$$

Ainsi, la Figure 1.16 montre clairement qu'une loi cinétique classique de type B ne représente pas très bien les résultats obtenus par les expériences virtuelles. Cette méthode est particulièrement efficace dans les régimes transitoires obtenus par l'addition de réactif avec des profils de coulée complexes. Par contre, cette méthode ne permet pas de proposer de manière systématique des lois cinétiques.

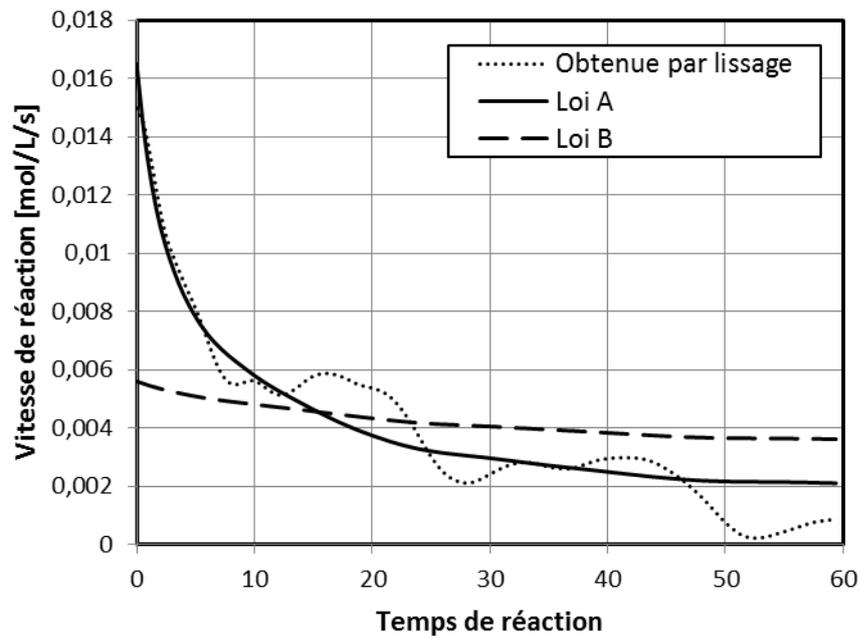


Figure 1.16 : Comparaison de deux types de lois cinétiques par rapport à l'expérience.

1.3.4 Planification d'expériences

La plupart des planifications d'expériences dans les cas non-linéaires reposent sur l'approximation qui consiste à linéariser localement le modèle par rapport aux paramètres afin de pouvoir estimer la région de confiance approchée. Contrairement aux cas linéaires, la région de confiance approchée dépend de la valeur des paramètres. Cela implique que la planification d'expériences doit se faire de manière itérative, c'est-à-dire qu'après chaque expérience, il faut ré-estimer les paramètres et l'expérience optimale du point de vue de l'acquisition de données comme le montre le schéma de la Figure 1.17.

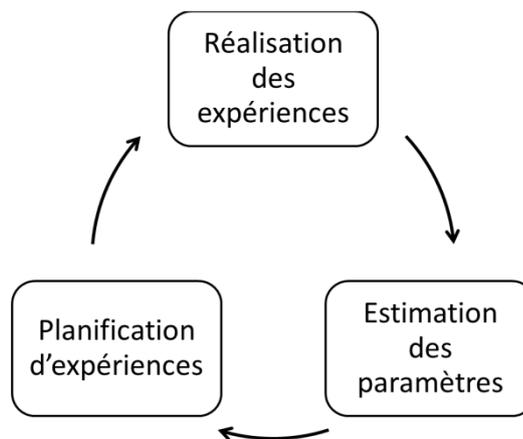


Figure 1.17 : Schéma de principe de l'approche phénoménologique.

Les critères d'optimalité sont comparables à ceux présentés pour les modèles linéaires (§ 1.2.4), ce sont des fonctions métriques, notée $\Phi(\cdot)$ de la matrice de covariance des paramètres, elle-même obtenue par le calcul de la matrice d'information de Fisher ($\Sigma_{\theta} = \mathbf{F}^{-1}$).

Concrètement, cela revient à chercher l'expérience $\Xi = \{\xi_1, \dots, \xi_{n_{pm}}\}$ contenant n_{pm} échantillons, qui maximise un critère d'optimalité étant données les n_{exp} expériences déjà réalisées.

$$\Xi_{n_{exp}+1} = \arg \min_{\Xi \in \mathbb{X}} \Phi \left(\Sigma_{\theta} \left(\theta \left\{ \Xi_1, \dots, \Xi_{n_{exp}}, \Xi \right\} \right) \right) \quad (1.69)$$

avec

$$\Sigma_{\theta}(\theta|\{\Xi, \xi\}) = \left[\mathbf{F}(\theta|\Xi) + \sum_{k=1}^{n_{exp}} \mathbf{F}(\theta|\Xi_k) \right]^{-1} \quad (1.70)$$

Les quatre métriques $\Phi(\cdot)$ les plus couramment utilisées sont résumé par la Figure 1.18. Ces quatre critères sont :

- La **D-optimalité** minimise le volume de l'ellipsoïde de confiance. Elle s'obtient lorsque $\Phi(\cdot) = \det(\cdot)$, c'est le critère historique proposé par Box et Lucas (1959). L'inconvénient majeur de ce critère est de favoriser l'estimation des paramètres les plus sensibles, alors que les incertitudes sur les autres paramètres restent grandes.
- La **E-optimalité** introduite par Hosten (1974) minimise le plus grand axe de l'ellipsoïde de confiance ($\Phi(\cdot) = \min \lambda$). Ce critère n'étant qu'un critère de forme, il tend à rendre plus sphérique l'ellipsoïde de confiance. Il ne diminue pas l'incertitude globale sur les paramètres.
- La **A-optimalité** minimise l'hypercube enveloppant l'ellipsoïde de confiance ($\Phi(\cdot) = \text{tr}(\cdot)$). Comme ce critère ne prend pas en compte les éléments non diagonaux de la matrice de covariance Σ_{θ} , cela induit une perte d'information notable, particulièrement quand les paramètres sont fortement corrélés (Franceschini 2008). L'intérêt de ce critère est de ne pas favoriser les paramètres les plus sensibles ce qui est intéressant lorsque le nombre de paramètres est relativement élevé (supérieur à 4)
- La **C-optimalité**, quant à elle, tend à diminuer la corrélation globale des paramètres. Ce critère proposé par Pritchard et Bacon (1978) s'obtient en prenant la racine carrée de la somme des coefficients de corrélation des paramètres pris deux à deux :

$$\Phi(\cdot) = \sqrt{\sum_{\substack{i,j \\ i \neq j}} \frac{Corr_{ij}^2}{n_p^2 - n_p}}, \quad Corr_{ij} = \frac{v_{ij}}{\sqrt{v_{ii}v_{jj}}} \quad (1.71)$$

où v_{ij} est l'élément (i, j) de la matrice Σ_{θ} .

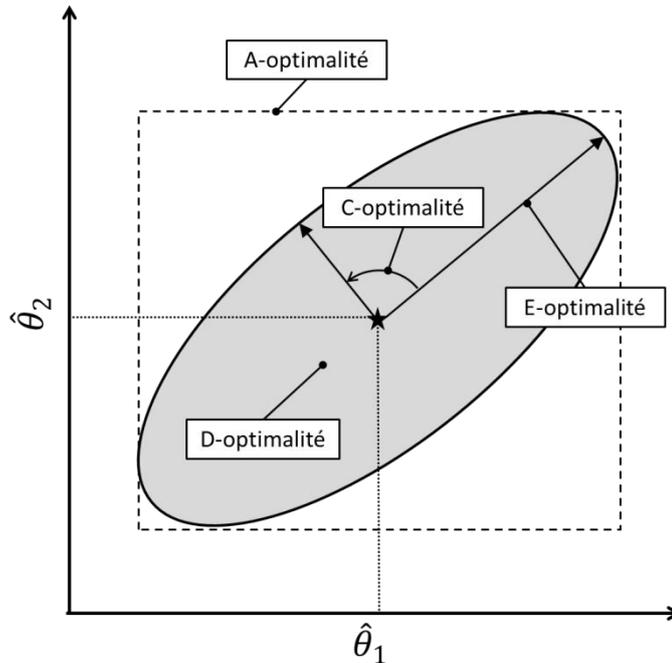


Figure 1.18 : Illustration de l'action des quatre principaux critères de planification d'expériences sur la matrice de covariance des paramètres estimés.

Il existe de nombreux autres critères qui ont été développés dans le but de pallier aux défauts des critères abordés plus haut. Le lecteur pourra se reporter à l'état de l'art publié par Franceschini et Macchietto (2008).

Ces techniques sont très efficaces et l'on peut trouver de nombreux cas d'étude dans la littérature. Cependant, ces techniques de planification d'expériences sont dites *locales*, car l'approximation linéaire, sur laquelle reposent ces techniques, suppose que l'estimation des paramètres est proche de leur vraie valeur. Or, lorsque le nombre d'expériences est restreint, il n'est pas rare que l'estimation des paramètres soit, de fait, assez éloignée de la vraie valeur. Ainsi, des méthodes dites *robustes* ont été mises au point pour éviter de converger vers un optimum local.

L'idée initialement proposée par Pronzato et Walter (1988), appelée approche « maxmin », est de chercher quel est le jeu de paramètres pour lequel le critère optimal de la meilleure expérience planifiée est le pire possible. En d'autres termes, on veut trouver l'expérience qui optimise le critère D-optimal dans le pire des cas. Cela donne formellement :

$$\Xi_{n_{exp}+1} = \max_{\Xi \in \mathbb{X}} \min_{\theta \in \Theta} \det \left(\mathbf{F} \left(\theta \mid \{ \Xi_1, \dots, \Xi_{n_{exp}}, \Xi \} \right) \right) \quad (1.72)$$

Cette approche est généralement moins efficace dans les premières itérations, mais beaucoup plus efficace par la suite. Toutefois, les ressources calculatoires nécessaires pour mettre en œuvre cette approche sont considérables dès que le modèle utilisé est un peu complexe.

1.3.5 Exploitation des résultats

L'exemple qui suit, réalisé à partir d'expériences que nous avons simulées, illustre le processus d'optimisation de la performance. L'objectif est de maximiser le rendement du produit intermédiaire R de deux réactions consécutives.



On suppose alors une loi cinétique du premier ordre pour chacune des réactions, on obtient alors une formulation implicite du problème :

$$\begin{aligned} \frac{d[A]}{dt} &= -k_1[A] \\ \frac{d[R]}{dt} &= k_1[A] - k_2[R] \\ \frac{d[S]}{dt} &= k_2[R] \end{aligned} \quad (1.74)$$

Ce système d'équations différentielles peut se résoudre analytiquement, on obtient donc la formulation explicite suivante (en supposant qu'à l'instant $t = 0$ seul A est introduit dans le réacteur) :

$$\begin{aligned} [A](t) &= [A]_0 \exp(-k_1 t) \\ [R](t) &= [A]_0 \frac{k_1}{k_2 - k_1} (\exp(-k_1 t) - \exp(-k_2 t)) \\ [S](t) &= [A]_0 \left(1 - \frac{k_2 \exp(-k_1 t) - k_1 \exp(-k_2 t)}{k_2 - k_1} \right) \end{aligned}$$

Supposons que seul R soit mesuré, que l'on ne soit pas capable de prélever plusieurs échantillons et que l'on connaisse l'écart-type de l'erreur expérimentale $\sigma_{[R]} = 0.025 \text{ mol/L}$. On a donc un modèle où les conditions expérimentales sont la concentration initiale de $[A]_0$ et le temps de réaction t . Pour simplifier, nous fixons $[A]_0 = 1 \text{ mol/L}$, il n'y a au final qu'une seule condition expérimentale. Les paramètres du modèle sont les constantes cinétiques k_1 et k_2 . Les stratégies expérimentales dédiées aux modèles non linéaires vont donc permettre de les déterminer du mieux possible afin de pouvoir prédire le maximum de rendement, qui dans ce cas, vaut :

$$Y_{R/A}^{max} = \left(\frac{k_2}{k_1} \right)^{\frac{k_2}{k_1 - k_2}} \quad (1.75)$$

et de prédire aussi les conditions expérimentales qui permettent de maximiser ce rendement :

$$t^{max} = \frac{1}{k_2 - k_1} \ln \frac{k_2}{k_1} \quad (1.76)$$

Afin d'amorcer la méthode, il est nécessaire d'avoir une estimation préliminaire des paramètres ($\widehat{k}_1 = 0,562 \text{ min}^{-1}$ et $\widehat{k}_2 = 0,327 \text{ min}^{-1}$) qui dans notre cas est obtenue grâce à trois points de mesures voir la Figure 1.19. Il est important de remarquer que l'intervalle de confiance à 95% sur les prédictions est très large malgré le fait que le modèle semble passer parfaitement par les points mesurés. Cela est lié à l'insuffisance des données utilisées pour l'estimation paramétrique.

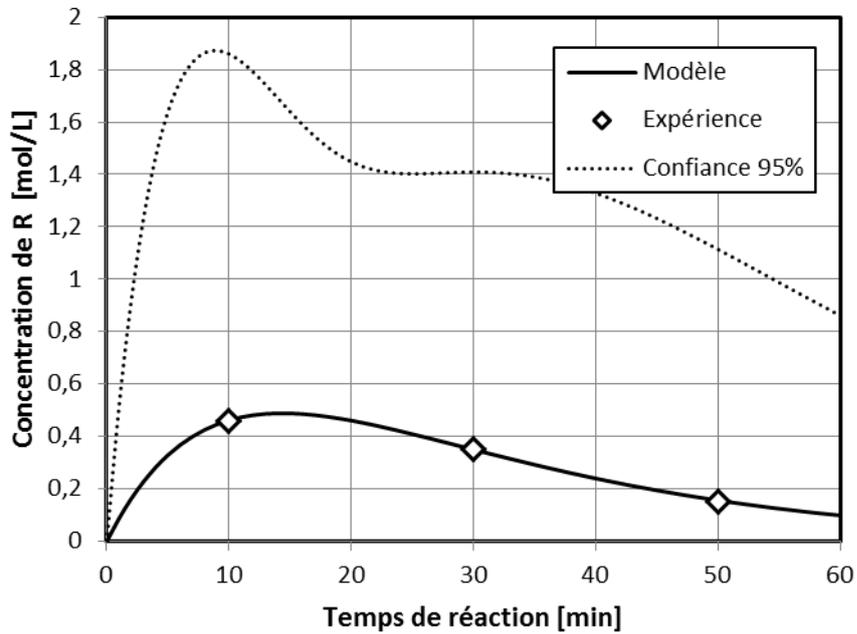


Figure 1.19 : Exemple de régression paramétrique et de l'intervalle de confiance des prédictions du modèle.

Pour savoir où faire le nouveau point de prélèvement, on cherche quel est le temps de prélèvement qui permet de minimiser le critère D-optimal, c'est-à-dire de minimiser le déterminant de la matrice de covariance des paramètres Σ_{θ} :

$$t_{n_{exp}+1} = \arg \min_{t \in [0,60]} \det(\Sigma_{\theta}(\{k_1, k_2\} | \{t_1, t_2, t_3, t\})) \quad (1.77)$$

Pour rappel, on obtient la matrice de covariance des paramètres par la relation $\Sigma_{\theta} = \left[\sum_{i=1}^{n_y} \sum_{j=1}^{n_y} \tilde{\sigma}_{ij} \mathbf{Q}_i^T \mathbf{Q}_j \right]^{-1}$ où la matrice de sensibilité \mathbf{Q}_i dépend des expériences déjà faites et de l'expérience à faire :

$$Q_i = \begin{bmatrix} \left. \frac{\partial[R]}{\partial k_1} \right|_{t_1, \widehat{k}_1, \widehat{k}_2} & \left. \frac{\partial[R]}{\partial k_2} \right|_{t_1, \widehat{k}_1, \widehat{k}_2} \\ \vdots & \vdots \\ \left. \frac{\partial[R]}{\partial k_1} \right|_{t_3, \widehat{k}_1, \widehat{k}_2} & \left. \frac{\partial[R]}{\partial k_2} \right|_{t_3, \widehat{k}_1, \widehat{k}_2} \\ \left. \frac{\partial[R]}{\partial k_1} \right|_{t, \widehat{k}_1, \widehat{k}_2} & \left. \frac{\partial[R]}{\partial k_2} \right|_{t, \widehat{k}_1, \widehat{k}_2} \end{bmatrix} \quad (1.78)$$

La Figure 1.20 montre la valeur du critère D-optimal pour chaque temps de prélèvement. On voit clairement deux optima, un local et un global. Par conséquent, on choisira de faire la prochaine expérience à l'optimum global ($t_4 = 8 \text{ min}$). Dans le cas où le système est plus complexe, en particulier, lorsque le nombre de conditions opératoires est supérieur à deux, il n'est plus possible de tracer l'évolution du critère D-optimal sur une figure 2D, il faut donc utiliser un algorithme d'optimisation.

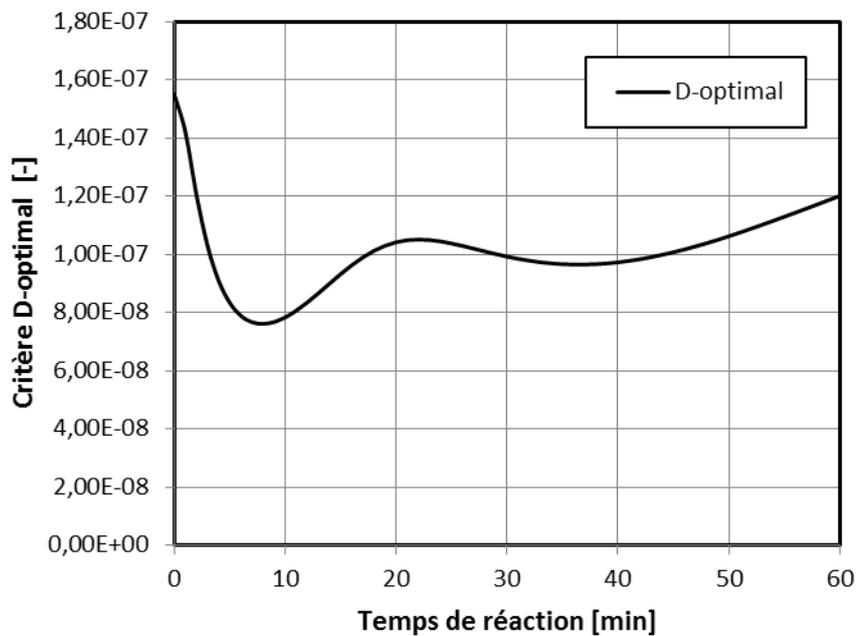


Figure 1.20 : Exemple d'évolution du critère D-optimal en fonction du temps de prélèvement.

La nouvelle expérience permet d'enrichir la base de données et d'améliorer la qualité de l'estimation des paramètres. La Figure 1.21, montre l'ellipsoïde de confiance de l'estimation des paramètres après l'initialisation, la 3^{ème} et la 10^{ème} itération D-optimale. On observe bien un resserrement de l'ellipsoïde de confiance au fur et à mesure des expériences.

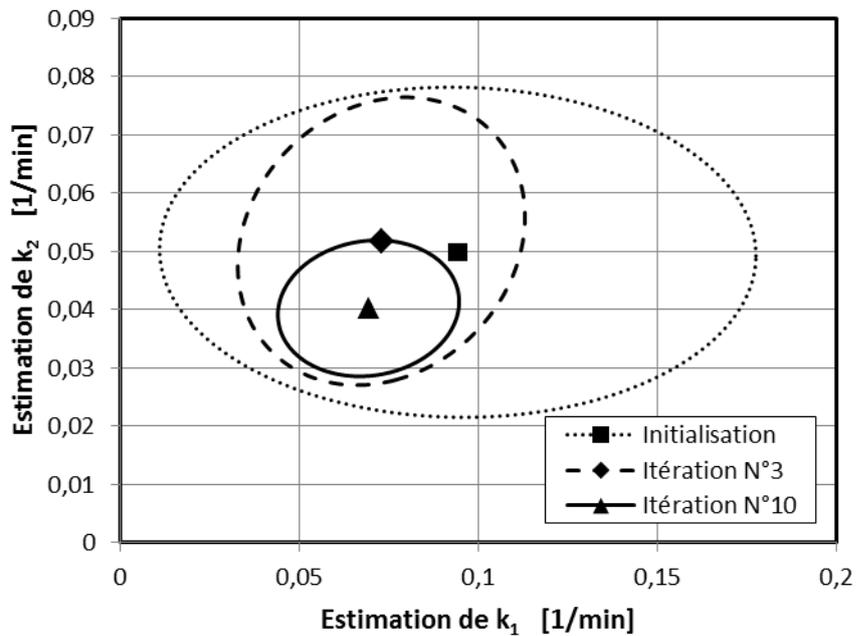


Figure 1.21 : Exemple d'évolution de la précision de l'estimation des paramètres k_1 et k_2 en fonction du nombre d'expériences D-optimales.

Le resserrement de la région de confiance des paramètres implique que le modèle soit plus fiable, ce que montre la Figure 1.22 où l'intervalle de confiance des prédictions est relativement étroit. Par ailleurs, cette figure montre comment sont répartis les points de mesure D-optimaux après 10 itérations : il apparait clairement que ces points de mesure ne sont pas répartis de manière aléatoire, ni de manière homogène sur le domaine expérimental :

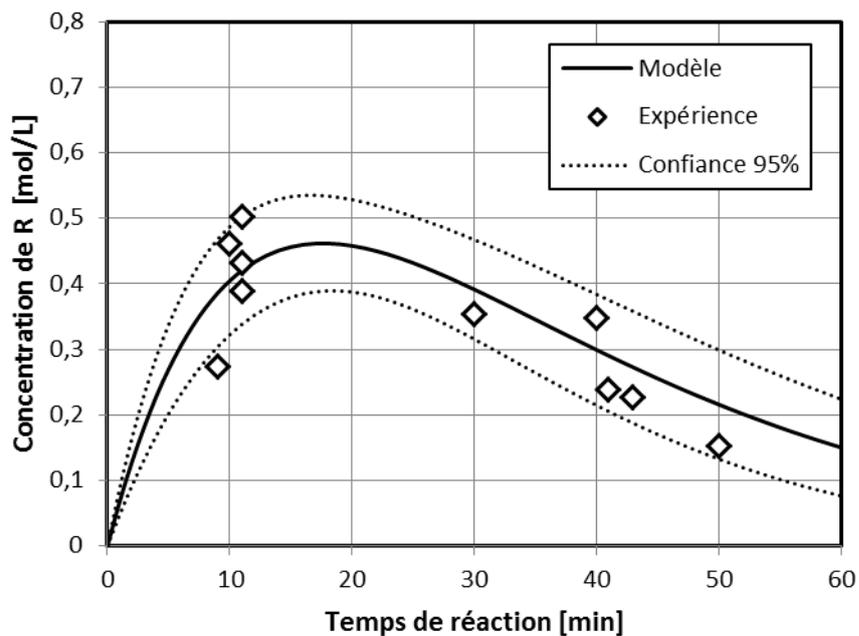


Figure 1.22 : Exemple de répartition des temps de prélèvement D-optimaux et l'intervalle de confiance de la prédiction du modèle.

Mais l'estimation des paramètres n'est pas une fin en soi, ce qui intéresse généralement l'utilisateur, c'est bien de prédire les meilleures performances d'un procédé et dans quelles conditions elles sont atteintes. D'après les équations (1.75) et (1.76), il est possible de calculer le rendement maximal théorique et le temps de réaction associé à ce rendement. Les Figure 1.23 et Figure 1.24 montrent l'évolution des performances prédites à chaque itération. De même que pour les prédictions du modèle, les prédictions de performance sont de plus en plus fiables au fur et à mesure des itérations, même si les valeurs prédites dès l'initialisation sont proches des valeurs prédites après 10 itérations.

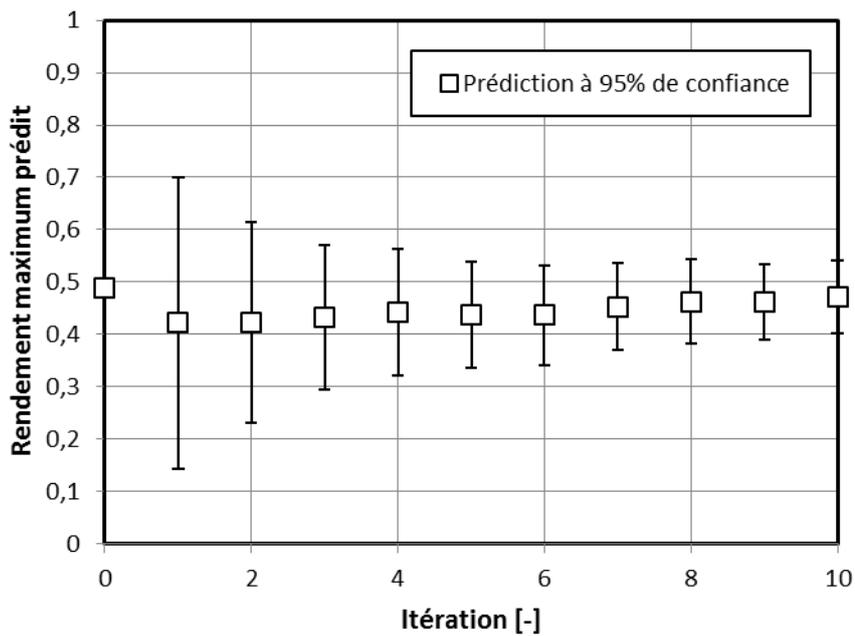


Figure 1.23 : Évolution de la prédiction du rendement maximum et de son intervalle de confiance associé.

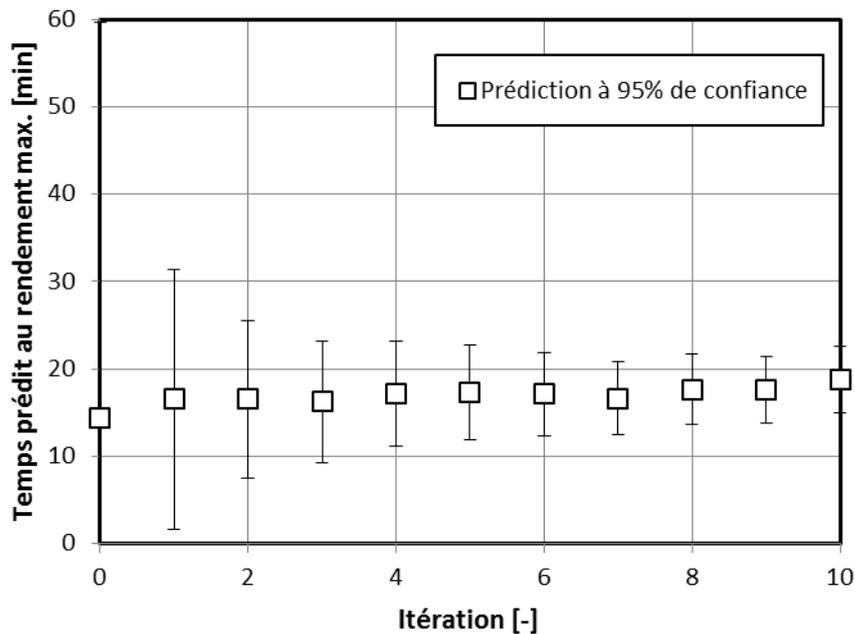


Figure 1.24 : Évolution de la prédiction du temps de séjour au rendement maximum et de son intervalle de confiance associé.

Ainsi, avec un modèle phénoménologique, dans la plupart des cas, seul un nombre restreint d'expériences est nécessaire pour obtenir une valeur optimale de performance assez fiable. Ceci est lié à l'introduction de connaissance a priori. En revanche, si la structure du modèle était fautive, il serait fort possible que l'optimum trouvé ait été assez éloigné des résultats trouvés ici. L'élaboration de la structure du modèle est primordiale lorsque l'on utilise des modèles phénoménologiques.

1.4 Approche mixte

Ce type d'approche, itérative tout comme l'approche phénoménologique, part du constat que cette dernière est extrêmement dépendante du choix de la structure du modèle. Nous présenterons ici trois approches, l'optimisation évolutionnaire, l'algorithme de Nancy et l'optimisation séquentielle (en anglais, run-to-run optimization) développée à l'école polytechnique de Lausanne.

L'idée principale de ces stratégies est de minimiser l'impact de la structure du modèle sur la réussite de la stratégie. Ainsi, l'optimisation évolutionnaire utilise un nouveau modèle local de type *boîte noire* à chaque itération, l'algorithme de Nancy permet de faire évoluer un schéma réactionnel apparent et l'optimisation séquentielle utilise un modèle phénoménologique relativement simplifié.

Contrairement aux deux approches abordées précédemment, cette approche fonde la phase d'acquisition de données (i.e. de planification d'expérience) non pas sur la région de confiance des paramètres mais sur l'optimisation du critère de performance.

1.4.1 L'optimisation évolutionnaire

Cette approche ne nécessite pas de modèle particulier. L'idée est d'appliquer un algorithme d'optimisation numérique sur la fonction de coût mesurée. Habituellement, les algorithmes d'optimisation minimisent une fonction numérique calculée par un programme. Dans le cas de l'optimisation évolutionnaire, la fonction est mesurée expérimentalement. Ainsi, on peut utiliser la plupart des algorithmes classiques tels que les algorithmes du gradient ou du simplexe. Malheureusement, lorsque le nombre de conditions opératoires est grand (par exemple profil de coulée ou de température), le nombre d'expériences à faire augmente considérablement et peut devenir rédhibitoire. En outre, il est possible que les conditions optimales ainsi trouvées soient en réalité un optimum local.

Cette méthode, bien connue depuis plusieurs dizaines d'années et décrite par Box (1957), a été appliquée avec succès dans de nombreux autres domaines industriels. Elle a récemment été appliquée pour optimiser une synthèse réalisée dans un réacteur intensifié, le Corning Advanced Flow Reactor (McMullen et Jensen 2010). Les auteurs ont pu facilement utiliser cette approche, car leur système expérimental était complètement automatisable et piloté par ordinateur.

1.4.2 L'algorithme de Nancy

Cet algorithme développé au Laboratoire des Sciences du Génie Chimique à Nancy à la fin des années 80 est une approche hybride (Fillipi-Bossy 1989 et Marchal-Brassely et al. 1992), qui combine modèle de tendance et modèle phénoménologique. L'idée est d'utiliser un modèle basé sur des bilans de matière mais dont le schéma réactionnel est réactualisé à chaque itération.

Élaboration du schéma réactionnel et des lois cinétiques

La première étape consiste à construire le schéma réactionnel grâce à des bilans de matière entre l'entrée et la sortie d'un réacteur parfaitement agité. En réarrangeant l'équation (1.57), on a :

$$\int_{\tau=t_0}^{\tau=t} F_j^R(\tau) d\tau = n_j(t) - n_j(t_0) - \int_{\tau=t_0}^{\tau=t} [F_j^E(\tau) - F_j^S(\tau)] d\tau \quad (1.79)$$

En exprimant le terme de production grâce aux avancements de réactions et en utilisant le titre molaire défini à l'équation (1.65), on obtient l'égalité suivante :

$$y_j(t) = \int_{\tau=t_0}^{\tau=t} F_j^R(\tau) d\tau = \sum_{i=1}^{n_r} \nu_{ij} X_i(t) \quad (1.80)$$

Ce qui sous forme matricielle donne :

$$\mathbf{Y} = \mathbf{X} \mathbf{N} \quad (1.81)$$

L'objectif est donc de déterminer la matrice des coefficients stœchiométriques \mathbf{N} de dimension $n_r \times n_c$ et la matrice des avancements de chaque réaction \mathbf{X} en chaque point de mesure de dimension $n_r \times n_{pm}$. Le nombre de réactions est fixé par l'utilisateur, il peut être amené à évoluer au fur et à mesure des itérations. Il y a donc $n_r \times n_c + n_{pm} \times n_r$ inconnues, et cela impose une condition sur le nombre minimum de points de mesures à effectuer (remarque: si l'on dispose d'un certain nombre d'expériences, c'est alors le nombre de réactions qui sera limité).

$$n_{pm} \times n_c \geq n_r \times n_c + n_{pm} \times n_r \Leftrightarrow n_{pm} \geq \frac{n_r \times n_c}{n_r + n_c} \quad (1.82)$$

On peut prendre plusieurs points de mesure sur une seule et même expérience, mais il est préférable de faire plusieurs expériences pour que l'estimation de \mathbf{N} et \mathbf{X} soit plus précise. Pour les estimer, on utilise une technique itérative où l'on estime successivement \mathbf{X} puis \mathbf{N} . Pour amorcer l'estimation on utilise tout d'abord une matrice des coefficients stœchiométriques préliminaire \mathbf{N}_0 contenant une seule et unique réaction et choisie de telle sorte que les composés qui sont majoritairement consommés aient leur coefficient égal à -1, ceux qui sont produits à +1 et les autres à 0. Ensuite, on cherche par une méthode des moindres carrés les avancements de réaction :

$$\mathbf{X}_1 = (\mathbf{N}_0 \mathbf{N}_0^T)^{-1} \mathbf{N}_0 \mathbf{Y}^T \quad (1.83)$$

Ayant une nouvelle valeur des avancements, on peut alors réajuster la matrice des coefficients stœchiométriques toujours par une méthode des moindres carrés :

$$\mathbf{N}_1 = (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{Y} \quad (1.84)$$

Cette nouvelle matrice ne possède pas de coefficients entiers mais des nombres réels. On arrondi alors à l'entier le plus proche chacun des coefficients. L'estimation successive de \mathbf{N} et \mathbf{X} se fait jusqu'à convergence de la somme des résidus obtenue par $SCE = \text{tr}([\mathbf{Y} - \mathbf{XN}]^T[\mathbf{Y} - \mathbf{XN}])$.

Lorsque \mathbf{X} et \mathbf{N} ont convergé, on ajoute une nouvelle réaction si le nombre de points de mesure est suffisant ou si l'utilisateur le juge opportun. Pour amorcer les coefficients de la nouvelle réaction, on calcule la matrice des titres molaires non expliqués par la première réaction qui n'est autre que la matrice des résidus ($\mathbf{R} = \mathbf{Y} - \mathbf{XN}$), puis l'on applique le même amorçage : si R_j est globalement négatif (resp. positif) alors $\nu_{2j} = -1$ (resp. $\nu_{2j} = +1$). Puis les étapes itératives d'estimation de \mathbf{X} et \mathbf{N} décrites plus haut sont répétées jusqu'à convergence. Lorsqu'il n'est plus possible de rajouter des réactions ou que l'utilisateur juge qu'ajouter une nouvelle réaction est superflu, on obtient la matrice stœchiométrique du schéma réactionnel. Il est important de noter que le schéma obtenu est apparent et peut être dénué de sens physique.

Ensuite, à chaque réaction est associée une loi cinétique. D'après les auteurs de l'algorithme de Nancy, une simple loi cinétique ayant un ordre 1 par rapport à chaque réactif et utilisant la loi d'Arrhenius suffit pour que la stratégie fonctionne.

Optimisation du procédé

La structure du modèle étant fixée, il reste à estimer les paramètres cinétiques du modèle en utilisant les mêmes bilans de matière que pour l'élucidation du schéma réactionnel. Le modèle ainsi développé permet de prédire quelles sont les conditions expérimentales qui permettent d'optimiser un critère de performance choisi arbitrairement par l'utilisateur. Les conditions expérimentales que l'on peut faire varier sont la température, les concentrations initiales, les profils de coulées de réactifs.

L'expérience dans les conditions optimales de performance prédites est alors réalisée. On compare alors le critère estimé et le critère prédit. Si un écart important est constaté, cela signifie que, premièrement, la base de données des constantes cinétiques et des ordres de réaction utilisée pour l'estimation des cinétiques est insuffisante, et deuxièmement, que la structure même du modèle n'est pas adéquate. Les auteurs préconisent alors de recalculer une matrice de coefficients stœchiométriques et les paramètres cinétiques associés en augmentant si possible le nombre de réactions. Les données utilisées pour cela sont complétées par les résultats expérimentaux de l'expérience réalisée au maximum de performance. Lorsque le critère estimé et le critère mesuré sont concordants, l'algorithme s'arrête. La Figure 1.25 montre le principe de fonctionnement de l'algorithme de Nancy.

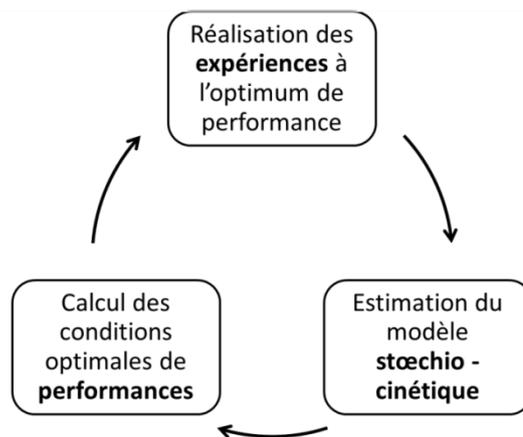


Figure 1.25 : Principe de fonctionnement de l'Algorithme de Nancy.

La principale différence entre l'algorithme de Nancy et les approches phénoménologiques est que l'on se sert des expériences à l'optimum de performance prédit pour enrichir la base de données expérimentales au lieu d'utiliser des méthodes de planification d'expériences. Ainsi, avec cette approche, on prend le risque de ne pas explorer des régions du domaine expérimental qui pourraient être intéressantes pour l'acquisition de données.

1.4.3 L'optimisation séquentielle « run-to-run »

Nous ne rentrerons pas dans les détails de cette technique développée à l'école polytechnique de Lausanne et le lecteur pourra se reporter directement aux travaux de (Srinivasan et al. 2001). L'idée sur laquelle l'optimisation séquentielle est basée est de

considérer l'optimisation de la conduite du réacteur comme un problème de commande en boucle fermée sous contraintes (débit d'entrée maximal, spécification produit, température limite, ...). Ainsi on assimile la répétition de nombreuses expériences batch à un procédé continu, où la mesure et l'ajustement des variables d'entrées se fait à la fin de chaque batch. À chaque nouvelle expérience, les conditions opératoires sont légèrement modifiées, de telle sorte que le critère d'optimisation de performance soit amélioré et que les contraintes soient respectées.

Pour calculer les nouvelles conditions, la stratégie « run-to-run » repose sur un modèle phénoménologique simplifié. Cela permet d'obtenir la sensibilité de la fonction de coût par rapport aux conditions expérimentales sans avoir à la mesurer (optimisation évolutionnaire). Connaissant ces sensibilités, il est possible de déterminer dans quelle direction faire varier les conditions opératoires pour améliorer la fonction de coût. La région de confiance des paramètres du modèle est utilisée pour que l'algorithme de commande ne propose pas de nouvelles conditions qui risqueraient de ne pas satisfaire les contraintes. Après chaque itération, les paramètres sont de nouveau estimés.

L'intérêt de cette technique est de pouvoir être implémenté directement sur une chaîne de production industrielle car cette stratégie satisfait les contraintes à chaque itération. Toutefois, la violation des contraintes n'est pas garantie et repose sur le choix judicieux des réglages de l'algorithme de commande. Avec de bons réglages, tous les lots intermédiaires, jusqu'à ce que l'optimum soit trouvé, respectent les contraintes et peuvent donc être commercialisés.

Cette technique n'a été appliquée actuellement qu'aux réacteurs discontinus. Tout comme l'algorithme de Nancy, la base de données s'enrichit avec les expériences réalisées à l'optimum de performance, les zones riches en information sur les paramètres ne sont pas explorées.

1.5 Synthèse

Nous avons vu dans ce chapitre qu'il existe trois grandes approches expérimentales pour l'optimisation de critères de performances. Ces trois grandes approches sont toutes basées sur la création d'un modèle, la rationalisation des expériences à faire et la prédiction de l'optimum de performance (les critères de performances devant, dans tous les cas, être définis par l'utilisateur). Ce qui les différencie c'est la manière dont est construit le modèle.

Chacune de ces approches possèdent des inconvénients et des avantages (Tableau 1.5). Par exemple, l'approche tendancielle est très efficace lorsque l'on ne possède aucune connaissance phénoménologique du système étudié. Mais, elle est relativement mal adaptée dans les cas où l'espace des conditions opératoires n'est pas un hypercube. Or, le fait d'utiliser plusieurs technologies, induit des discontinuités dans l'espace de recherche. L'approche phénoménologique, plus lourde à mettre en œuvre, permet de mieux gérer ce problème de conditions opératoires. En outre, elle a

l'avantage d'être plus robuste vis-à-vis du changement d'échelle. Enfin, les approches auto-adaptatives, basées sur la construction d'un modèle dont la complexité s'adapte au fur et à mesure de l'avancement de la stratégie, sont efficaces pour des problèmes bien délimités.

Le chapitre 2 pèse les pour et les contres de chaque approche pour définir une nouvelle stratégie qui s'adapte au problème du choix technologique.

Tableau 1.5 : Comparaison des différentes approches.

Approches	Niveau de connaissance requis	Nombre de conditions expérimentales	Extrapolation
<i>Tendancielle</i>	Très faible	Dépend du nombre d'expériences disponibles	Très faible
<i>Phénoménologique</i>	Élevé	Élevée	Élevée
<i>Évolutionnaire</i>	Faible	Dépend du nombre d'expériences disponibles	Faible
<i>Algo. Nancy</i>	Moyen	Élevée	Faible
<i>Run-to-run</i>	Moyen	Élevée	Faible

Chapitre 2

Développement de la stratégie

Ce chapitre a pour but d'élaborer une stratégie expérimentale pour le choix de la meilleure technologie disponible. Au chapitre précédent, nous avons détaillé un certain nombre d'approches permettant l'optimisation de critère de performances. Dans un premier paragraphe, après avoir défini en détail le problème du choix technologique et ces implications dans la gestion de l'espace des conditions opératoire, nous donnerons ses arguments qui nous ont permis d'élaborer la stratégie Stanislas⁸. C'est à partir de la « boîte à outils » détaillée dans le premier chapitre que nous l'avons développée. Une fois la stratégie établie, nous utiliserons un cas d'étude très simple pour l'illustrer. Dans un second paragraphe, nous aborderons l'aspect expérimental de la stratégie. Nous définirons le cahier des charges du banc expérimental, puis nous caractériserons les réacteurs utilisés.

2.1 Construction et détail de la stratégie

Dans ce paragraphe, nous détaillerons la stratégie expérimentale que nous avons implémentée et nous justifierons notre choix. Enfin, nous détaillerons les outils numériques mis en place.

2.1.1 Un nouveau degré de liberté : le choix technologique

Comme nous l'avons vu dans l'introduction générale, les nouvelles technologies récemment développées offrent de nouvelles fenêtres expérimentales qui peuvent être intéressantes pour améliorer les performances d'un procédé, la problématique étant de choisir la technologie la plus adaptée au système réactionnel étudié. Le choix de la meilleure technologie peut se voir comme l'optimisation d'un système dont le domaine opératoire est complexe et possiblement discontinu. On peut donc appliquer les méthodes d'optimisation de performance décrites plus haut.

⁸ Stanislas fait référence au dernier Duc de Lorraine, Stanisław Leszczyński, qui régna de 1766 à 1766 sur les duchés de Lorraine et de Bar. Stanisław Leszczyński, réputé pour s'intéresser d'avantage à la littérature et aux sciences qu'aux arts de la guerre, avait pour coutume de signer ses œuvres par « le philosophe bienfaisant ». La célèbre place qui porte son nom au cœur Nancy, capitale de la Lorraine, les châteaux de Commercy, de Lunéville et de Bar-le-duc, la Société Royale des Sciences et Belles-Lettres font partie de l'héritage de ce bon roi philosophe.

1.1.1.1. *Gérer plusieurs fenêtres opératoires*

Le type d'approche le plus simple et le plus facile à mettre en œuvre est l'approche tendancielle. Dans le cas de plusieurs technologies, le domaine opératoire n'est pas un hypercube, mais une forme complexe pouvant être discontinue. Ceci rend assez difficile l'utilisation des modèles de tendance et ce pour les raisons suivantes :

- Quand le domaine n'est pas un hypercube, les plans optimaux (factoriel, Box-Behnken, composite central) ne sont plus applicables. Il faut passer par l'utilisation d'algorithmes de recherche de plan D-optimaux, tel que l'algorithme de Fedorov (voir paragraphe 1.2.4). Ce qui est une difficulté supplémentaire.
- Les modèles de tendance sont très performants quand la fenêtre opératoire est restreinte, puisque ce sont des modèles d'interpolation. Or, lorsque l'on utilise plusieurs réacteurs, il est nécessaire d'avoir un modèle robuste qui puisse prendre en compte les différentes conditions opératoires des réacteurs et prédire le comportement des différents équipements. Les modèles de tendances sont donc mal adaptés à la stratégie recherchée. Les conditions opératoires et les modes de fonctionnement peuvent être différents d'une technologie à l'autre.
- Le nombre de facteurs peut être très grand si l'on considère des profils de coulée ou de température. Cela implique automatiquement un très grand nombre d'expériences à faire.

Pour pouvoir quand même utiliser les modèles de tendance, il serait préférable de réaliser un plan d'expériences pour chaque technologie puis de comparer les performances obtenues.

Les modèles phénoménologiques permettent de représenter le système sur un large domaine opératoire. De plus, planifier des expériences en prenant en compte des profils de coulées ou de température complexes n'induit pas un nombre d'expériences à faire plus grand puisque, contrairement aux modèles de tendance, le nombre de paramètres à estimer est indépendant du nombre de conditions expérimentales.

L'utilisation de modèles phénoménologiques semble donc plus appropriée pour gérer plusieurs technologies simultanément.

1.1.1.2. *L'échantillonnage*

L'autre aspect à prendre en compte pour le choix de l'approche est la flexibilité d'échantillonnage propre à chaque technologie. Là encore, les modèles phénoménologiques semblent plus appropriés. Pour planifier des expériences comprenant plusieurs échantillons en utilisant des modèles de tendance, nous avons deux possibilités :

- Soit on échantillonne de la même manière pour toutes les expériences du plan. Ce qui pour un batch revient à échantillonner à chaque expérience toujours aux mêmes temps de prélèvement.
- Soit on établit un plan D-optimal sans prendre en compte la possibilité d'échantillonner, puis lorsque cela est possible, on prélève plusieurs échantillons en une seule expérience, comme le montre la Figure 2.1. En effet, cette figure montre un plan D-optimal de 27 expériences (certaines expériences sont répétées) pour un modèle quadratique qui prend en compte la possibilité d'utiliser plusieurs technologies. On remarque qu'en seulement 3 expériences on peut produire les 7 échantillons demandés par le plan.

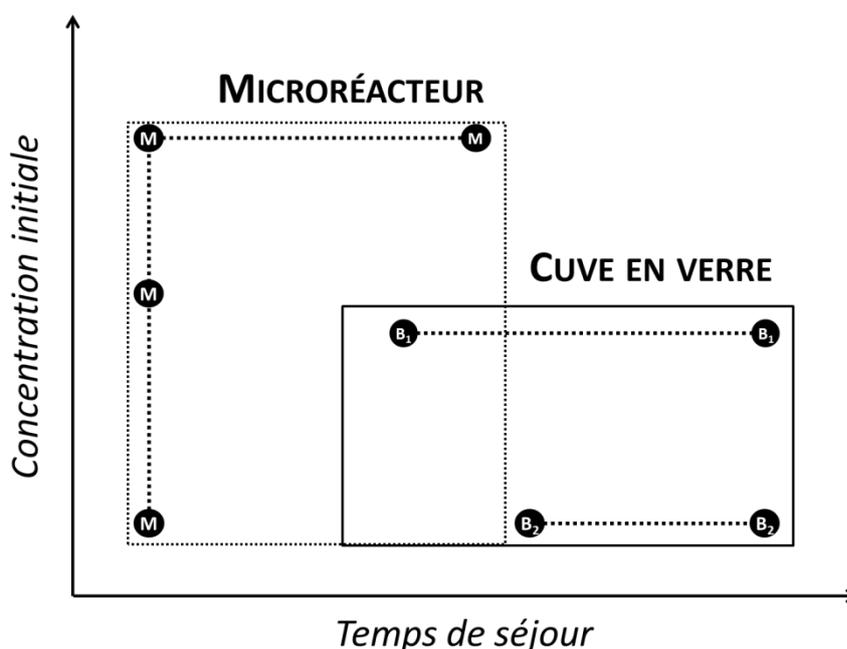


Figure 2.1 : Exemple de mise à profit de la capacité d'échantillonnage propre au microréacteur et à la cuve agitée afin de minimiser le nombre d'expériences.

De toute manière, même si l'on parvient à générer un plan D-optimal qui prend en compte les possibilités d'échantillonnage, les modèles linéaires aux paramètres ne sont pas très bien adaptés au fait que certains facteurs puissent être évalués plus de fois que d'autres.

L'utilisation de modèles phénoménologiques permet de tirer le meilleur parti de la flexibilité d'échantillonnage de chaque équipement, mais cela induit quelques difficultés numériques.

2.1.2 Détail de la structure

Pour les raisons détaillées au paragraphe 2.1.1 nous avons opté pour l'utilisation de modèles phénoménologiques dont une difficulté est de faire un compromis entre une bonne description des phénomènes sans avoir un modèle surparamétré, coûteux en acquisition de données. Nous avons donc opté pour notre stratégie à une structure assez similaire à celle de l'algorithme de Nancy basé sur des modèles phénoménologiques, dont la structure est itérative et qui a pour objectif de trouver un optimum de performance le plus rapidement possible. Cependant, contrairement à l'algorithme de Nancy, nous avons introduit une étape supplémentaire d'acquisition de données. Une fois la structure du modèle posée à l'aide d'un protocole d'initialisation (elle pourra être amenée à changer), notre stratégie alterne des phases d'acquisition de données et des phases de validation. Les trois étapes : initialisation, acquisition de données et validation sont détaillées dans les paragraphes qui suivent. La Figure 2.2 montre le schéma algorithmique de la stratégie :

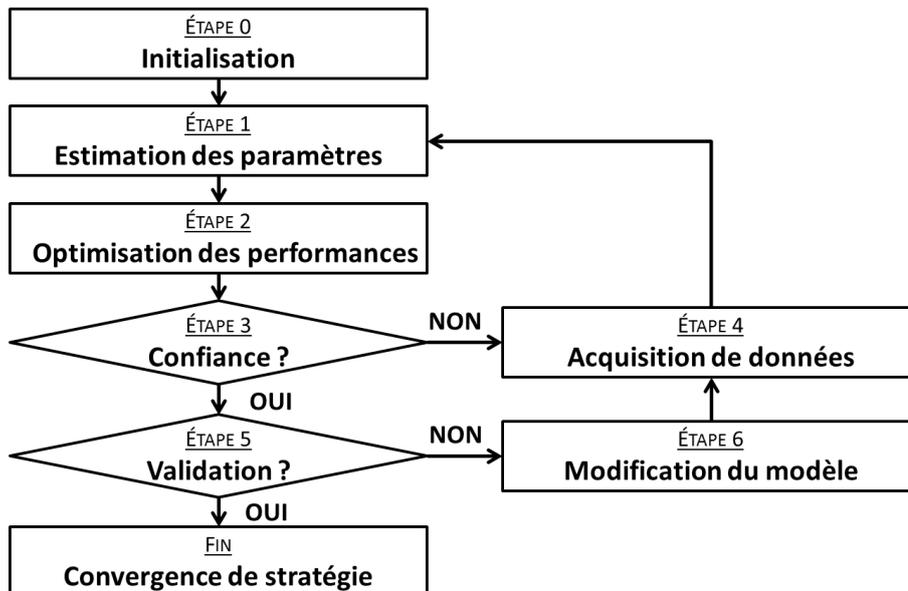


Figure 2.2 : Schéma algorithmique de la stratégie.

Afin de faciliter la compréhension, nous nous appuyons sur un cas d'étude virtuel élaboré à partir des travaux de (Issanchou, Cognet, et Cabassud 2003). Il s'agit de l'hydrolyse de l'acétate de pentyle (A) par la soude (B) dans un milieu liquide-liquide pour former du pentanol (C) et de l'acétate de sodium (D). Pour faire la réaction, on ajoute un fort excès d'acétate de pentyle (non soluble dans l'eau) de manière à obtenir une fraction volumique de la phase dispersée $\Phi_{disp} = 0,04$ et de la soude. Enfin, on suppose que deux technologies de réacteurs sont disponibles : un réacteur batch de 1L ($D_r = 10,8 \text{ cm}$ et $D_{stir} = 3,6 \text{ cm}$) et un microréacteur piston. Pour des questions pratiques, nous n'avons pas fait d'expériences réelles, mais nous les avons simulées en utilisant le modèle et les valeurs des paramètres utilisés par les auteurs de l'article. Dans le paragraphe d'introduction du chapitre 3, le lecteur trouvera une description détaillée de la manière dont sont générés les points expérimentaux simulés.

2.1.2.1 Protocole d'initialisation

Pour choisir les phénomènes à modéliser, nous avons mis au point un protocole qui permet en quelques expériences préliminaires de connaître la sensibilité de ces phénomènes. Ce protocole est basé sur l'analyse des temps caractéristiques (cf. Tableau 2.1). En faisant le rapport entre les valeurs d'un temps caractéristique obtenues dans deux conditions expérimentales distinctes, on peut éliminer des constantes physiques souvent inconnues et calculer assez facilement la dépendance de ce ratio par rapport à chacune des conditions expérimentales. Prenons l'exemple du temps caractéristique de transfert de chaleur pour une cuve agitée:

$$t_{heat} = \frac{\rho V C_p}{h S}$$

$$\text{avec } h = \frac{\lambda Nu}{D_r} \sim Re^{0,8} \quad (2.1)$$

$$\frac{S}{V} = \frac{\pi \frac{D_{cuve}^2}{4} + \pi H D_{cuve}}{\pi D_{cuve}^2 H} = \frac{5}{D_{cuve}} \quad \text{si } H = D_{cuve}$$

Soit :

$$t_{heat} \propto \rho^{0,2} \mu^{0,8} C_p D_{cuve} D_{agit}^{-1,6} N^{-0,8} \quad (2.2)$$

En faisant le ratio, pour deux conditions réacteurs a et b on obtient :

$$\frac{t_{heat}^a}{t_{heat}^b} \propto \left(\frac{D_{cuve}^a}{D_{cuve}^b} \right) \left(\frac{D_{agit}^b}{D_{agit}^a} \right)^{-1,6} \left(\frac{N^a}{N^b} \right)^{-0,8} \quad (2.3)$$

On peut faire de même pour la plupart des temps caractéristiques. Ainsi, en faisant varier chacune des conditions expérimentales par rapport à une expérience de référence, on peut estimer si un phénomène est limitant dans le système réactionnel. De plus, même si les valeurs des constantes physiques comme la viscosité, la masse volumique ne sont pas connues avec précision on peut toujours estimer un temps caractéristique avec l'ordre de grandeur des constantes. Si l'ordre de grandeur du temps caractéristique est très faible devant le temps caractéristique global (assimilable au temps de synthèse), on peut considérer que le phénomène n'est pas limitant et ne nécessite pas d'être pris en compte dans le modèle phénoménologique.

Tableau 2.1 : Principaux temps caractéristiques.

Temps caractéristiques	Symbole	Expression en cuve agitée	Expression en microcanal
Réaction	t_{reac}	$\frac{1}{k(T) (C^0)^{ordre-1}}$	idem
Macro-mélange	$t_{mix_{macro}}$	$\frac{4}{N} \left(\frac{D_a}{D_c}\right)^2$	–
Micro-mélange	$t_{mix_{micro}}$	$17 \left(\frac{\mu}{\rho \varepsilon}\right)^{0,5}$	$\approx 0,15 \varepsilon^{-0,45}$
Transfert de chaleur	t_{heat}	$\frac{\rho V Cp}{h S}$	$\frac{\rho Cp D_h^2}{\lambda Nu}$
Transfert à l'interface Liquide/Liquide	t_{LL}	$\frac{1}{k_L a}$	$\frac{1}{k_L a}$

Tableau 2.2 : Principaux nombres adimensionnels

Nombre adimensionnels	Symbole	Expression en cuve agitée	Expression en microcanal
Reynolds	Re	$\frac{\rho N D_a^2}{\mu}$	$\frac{\rho u D_h}{\mu}$
Weber	We	$\frac{\rho N^2 D_a^3}{\sigma}$	$\frac{\rho u^2 D_h}{\sigma}$
Newton	Np	$\frac{P}{\rho N^3 D_a^5}$	–
Nusselt	Nu	$\frac{h D_c}{\lambda}$	$\frac{h D_h}{\lambda}$
Capillarité	Ca	–	$\frac{u \mu}{\sigma}$

2.1. CONSTRUCTION ET DÉTAIL DE LA STRATÉGIE

Tableau 2.3: Principales grandeurs caractéristiques

Grandeur caractéristique	Symbole	Expression en cuve agitée	Expression en microcanal
Energie dissipée [W/kg]	ε	$\frac{Np N^3 D_a^5}{V}$	$\frac{32 \mu u^2}{\rho D_h^2}$
Aire interfaciale spécifique [m ² /m ³]	a	$\frac{6\Phi_{disp}}{d_{32}}$	$\frac{1}{D_h}$
Diamètre de Sauter des gouttes ⁹ [m]	d_{32}	$D_a Cte_1 (1 + Cte_2 \Phi_{disp}) We^{-0,6}$	$\approx \frac{D_{entrée}}{Ca}$
Coefficient d'échange à la paroi [W/m ² /K]	h	$\frac{\lambda Nu}{D_c} \approx Re^{0,8}$	$\frac{\lambda}{D_h}$
Augmentation de la température [K]	ΔT_{adiab}	$-\frac{\Delta_r H C^0}{\rho Cp}$	$-\frac{\Delta_r H C^0}{\rho Cp}$

Tableau 2.4 : Dépendance des temps caractéristiques en fonction des principales conditions expérimentales en Batch.

Temps caractéristique	Concentration	Température	Diamètre cuve	Diamètre agitateur	Vitesse d'agitation
Réaction	$\frac{1}{(C^0)^{ordre+1}}$	$\exp\left\{-\frac{Ea}{RT}\right\}$	—	—	—
Macro-mélange	—	—	$\frac{1}{D_c^2}$	D_a^2	$\frac{1}{N}$
Micro-Mélange	—	—	$D_c^{1,5}$	$\frac{1}{D_a^{2,5}}$	$\frac{1}{N^{1,5}}$
Transfert de chaleur	—	—	D_c	$\frac{1}{D_a^{1,6}}$	$\frac{1}{N^{0,8}}$
Transfert L/L	—	—	—	$\frac{1}{D_a^{0,8}}$	$\frac{1}{N^{1,2}}$

⁹ Les constantes Cte_1 et Cte_2 doivent être estimées pour chaque système. Cependant, on peut trouver de nombreuses références (Paul, Atiemo-Obeng, et Kresta 2004) permettant de couvrir une large gamme de conditions.

Tableau 2.5 : Dépendance des temps caractéristiques en fonction des principales conditions expérimentales en microréacteur.

Temps caractéristiques	Concentration	Température	Diamètre hydraulique	Vitesse du fluide
Micro-Mélange	–	–	$D_h^{0,9}$	$u^{-0,9}$
Transfert de chaleur	–	–	D_h^2	–
Transfert L/L	–	–	D_h	–

Reprenons l'exemple de l'hydrolyse de l'acétate de pentyle cité plus haut et rappelons qu'il s'agit d'expériences simulées. On peut supposer que la viscosité du milieu est proche de l'eau, de même pour la masse volumique. Avec ces informations, on peut estimer les temps caractéristiques de macro-mélange, micro-mélange et de transfert de chaleur pour une expérience de référence en cuve agitée ($C_B(t = 0) = 12 \text{ mol/m}^3, T = 25 \text{ °C}, N = 600 \text{ RPM}$) :

$$t_{mix_{macro}} = 4 \text{ s} \quad t_{mix_{micro}} = 0,1 \text{ s} \quad t_{heat} = 90 \text{ s} \quad (2.4)$$

Les temps de mélange micro et macro sont de l'ordre de la seconde. Or, le temps de réaction est de l'ordre d'une heure, il est donc peu probable que le mélange limite la réaction. Par ailleurs, en prenant une enthalpie de réaction $\Delta_r H = -50 \text{ kJ/mol}$ on trouve $\Delta T_{adiab} = 10^{-6} \text{ K}$ (avec comme concentration initiale $C_B^0 = 012 \text{ mol/m}^3$). Ainsi, on peut d'ores et déjà ne pas tenir compte de l'aspect thermique dans le modèle.

En revanche, la tension superficielle, les paramètres cinétiques et l'enthalpie de réaction sont complètement inconnus. Pour savoir si le transfert de matière à l'interface est important à modéliser, si le système est sensible à la température et si la réaction admet un ordre différent de 1, on peut par exemple faire varier la température et la vitesse d'agitation à un temps de synthèse fixé ($t_{synthese} = 30 \text{ min}$). On obtient les résultats suivants :

Tableau 2.6 : Exemple de conditions expérimentales pour le protocole d'initialisation de la stratégie.

N° exp	N [tr/min]	T [°C]	C_A^{aq} [mol/m ³]	X_B^{aq} [%]
REF	600	25	13	83
1	60	25	7	55.8
2	600	35	13	95.8

Le fait d'avoir diminué la vitesse d'agitation par dix a augmenté le temps de transfert à l'interface d'un facteur 16. Il se trouve que la conversion en B a fortement dimi-

nué, le transfert a donc un impact non négligeable sur le système, il faut le prendre en compte. De plus, si l'on considère le temps global de synthèse (temps qu'il aurait fallu pour atteindre 63% de conversion en supposant que le système soit régi par une décroissance exponentielle) qui s'exprime par :

$$X_B = 1 - \exp\left(-\frac{t_{\text{synthèse}}}{t_{\text{global}}}\right) \Rightarrow t_{\text{global}} = \frac{t_{\text{synthèse}}}{\ln(1 - X_B)} \quad (2.5)$$

on s'aperçoit qu'il passe de 40 à 85 min. Si la synthèse avait été entièrement déterminée par le transfert, il aurait été multiplié par 16. Donc le transfert ne détermine que partiellement le temps global. Par ailleurs, la température a un effet significatif puisque l'on passe de 83% de conversion en B à 96% entre l'expérience de référence et l'expérience 2. On peut donc postuler le modèle suivant décrit par la Figure 2.3 :

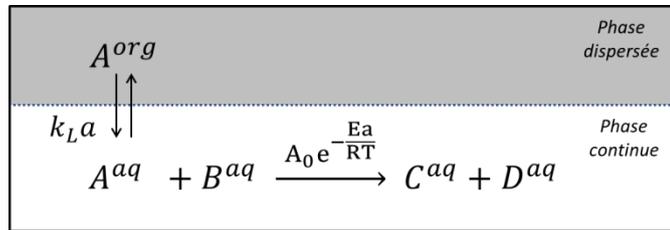


Figure 2.3 : Schéma du système réactionnel à l'issue du protocole d'initialisation.

Ce qui, sous forme d'équations différentielles, donne :

$$\begin{aligned} \frac{dC_A^{aq}}{dt} &= -A_0 \exp\left(\frac{-Ea}{RT}\right) C_A^{aq} C_B^{aq} + k_L a \left(\frac{C_A^{org}}{m} - C_A^{aq}\right) \\ \frac{dC_B^{aq}}{dt} &= -\frac{dC_C^{aq}}{dt} = -\frac{dC_D^{aq}}{dt} = -A_0 \exp\left(\frac{-Ea}{RT}\right) C_A^{aq} C_B^{aq} \end{aligned}$$

La résolution du modèle en réacteur batch et en microréacteur est identique puisque le fluide est incompressible et que l'écoulement en microréacteur est supposé piston. En revanche, ce qui les différencie, c'est la manière de calculer l'aire interfaciale spécifique. En supposant que la tension interfaciale est $\sigma = 0,05 \text{ N}\cdot\text{m}^{-1}$, on peut balayer une gamme d'aires interfaciales allant de 300 à 2100 m^2/m^3 en faisant varier la vitesse d'agitation de 100 à 1000 tr/min en cuve agitée et en microréacteur une gamme de 1000 à 10 000 m^2/m^3 en utilisant plusieurs microcanaux de différentes tailles.

2.1.2.2 Estimation paramétrique

On effectue une estimation paramétrique après chaque nouvelle expérience (voir Figure 2.2 page 70). À l'issue de cette étape, on dispose d'une nouvelle estimation des paramètres ainsi que de la matrice de covariance des paramètres. Pour le détail de l'estimation des paramètres, le lecteur se reportera à la section 1.1.4 page 26. Le choix de l'utilisation du quadratique ou du maximum de vraisemblance revient à l'utilisateur.

S'il connaît de manière indépendante la matrice de covariance de l'erreur de mesure il utilisera le critère quadratique, dans les autres cas, le maximum de vraisemblance.

Reprenons l'exemple illustratif d'une synthèse liquide-liquide décrite plus haut. Le modèle a été posé au paragraphe précédent. Les conditions expérimentales que l'on peut faire varier sont le temps de synthèse, la concentration initiale en B, la température et l'aire interfaciale spécifique (en jouant sur l'agitation en batch et sur le diamètre du canal en microréacteur). Les paramètres à estimer sont le facteur pré-exponentiel, l'énergie d'activation, le coefficient de transfert à l'interface et le coefficient de partage de A. Supposons que l'on ait déjà fait les trois expériences préliminaires pour la détermination de la structure du modèle et que chacune de ces expériences a permis de générer 6 échantillons aux temps de prélèvements suivants : 10, 20, 30, 40, 50 et 60 minutes. Grâce à ces premières données expérimentales, une première estimation des paramètres du modèle a été réalisée (voir Tableau 2.7).

Tableau 2.7 : Estimation paramétrique initiale.

<i>Paramètre</i>		Estimation	Intervalle de confiance à 95%
A_0	[mol/s]	$9,72 \times 10^4$	$\pm 2,01 \times 10^6$
Ea	[J/mol]	$5,17 \times 10^4$	$\pm 5,16 \times 10^4$
k_L	[m ² /s]	$4,04 \times 10^{-5}$	$\pm 3,88 \times 10^{-5}$
m	[–]	$5,76 \times 10^{-2}$	$\pm 2,13 \times 10^{-2}$

On peut remarquer que l'intervalle de confiance du facteur pré-exponentiel est très supérieur à 100%, et que les autres paramètres sont estimés avec un intervalle de confiance très large aussi. Cela signifie que les valeurs des paramètres trouvés sont à prendre avec précaution. Pour autant, ce qui nous intéresse, c'est bien la performance et il se peut que l'imprécision sur les paramètres ne soit pas si problématique pour la confiance sur les prédictions de performance. C'est ce que nous allons voir au paragraphe suivant.

2.1.2.3 Optimisation des performances

Après avoir terminé l'estimation paramétrique, la stratégie proposée effectue une optimisation des performances basée sur les nouvelles valeurs des paramètres. L'optimisation des performances permet le choix de la meilleure technologie et des conditions optimales associées.

Puisque les valeurs des paramètres ne sont que des estimations, elles sont associées à un intervalle de confiance. Or, on peut, comme vu dans le paragraphe 1.1.5, reporter cette incertitude sur l'estimation du critère de performance. Afin d'éviter de faire des expériences inutiles, la stratégie propose de valider expérimentalement les conditions optimales de performance lorsque l'incertitude sur la prédiction de chacun des critères de performance est inférieure à l'incertitude des mesures. Si ce n'est pas le cas,

c'est que le modèle n'est pas assez représentatif de la réalité, il faut donc l'affiner en réalisant une étape d'acquisition de données.

Si, au contraire, la confiance estimée est plus faible que la confiance expérimentale mesurée alors on réalise l'expérience au maximum de performance afin de valider expérimentalement que la performance mesurée est bien conforme aux attentes. Si ce n'est pas le cas, c'est que le modèle n'est pas assez réaliste et la structure est probablement mal adaptée au système réactionnel. Dans ce cas de figure, il est du ressort de l'utilisateur d'apporter une solution car il n'existe pas de méthode systématique pour changer la structure du modèle.

Pour illustrer cette étape, reprenons le cas d'étude liquide-liquide utilisé plus haut. La fonction d'agrégation des critères de performance que nous avons choisie est de la forme de celle décrite au paragraphe 1.1.5. Les deux critères sont le temps de synthèse et les concentrations en C et D dans le produit final. Les valeurs objectives et vetos sont détaillées dans le Tableau 2.8.

Tableau 2.8 : Critères de performances avec leur seuils veto et leur objectifs.

		Critère 1 [s]	Critère 2 [mol / m ³]
Performance	γ_i	t_f	$-(C_C^f + C_D^f)$
Objectif	o_i	0	-240
Veto	v_i	1200	-120

Après optimisation du critère de performance dans chacun des réacteurs, on obtient les résultats estimés suivants :

Tableau 2.9 : Conditions opératoires à l'optimum de performance et valeur des critères de performances.

	t [s]	T [°C]	C_B^0 [mol/m ³]	a [m ² /m ³]	Critère 1 [s]	Critère 2 [mol/m ³]	j_{perf}
Batch	300	90	120	3200	300	-239	0,177
Micro	110	90	120	10000	110	-231	0,0813

Le microréacteur possède un critère de performance plus petit que le réacteur batch (par convention, on cherche toujours à minimiser un critère), on considère donc que c'est la technologie la plus adaptée à ce système réactionnel. Mais comme l'intervalle de confiance du second critère estimé est supérieur à l'intervalle de confiance de la mesure de ce même critère (voir Tableau 2.10), si l'on réalise effectivement l'expérience dans les conditions optimales de performance, le critère mesuré risque d'être assez éloigné du critère prédit. Il faut donc affiner l'estimation des paramètres par une étape supplémentaire d'acquisition de données.

Tableau 2.10 : Critères estimés à l'optimum de performance avec leurs intervalles de confiance des prédictions et des mesures.

	Critère 1 [s]	Critère 2 [mol/m ³]
<i>Estimée</i>	110	-231
<i>Intervalle de confiance du modèle</i>	± 0	± 96
<i>Intervalle de confiance expérimental</i>	± 0	± 49

2.1.2.4 Acquisition de données

Le but de cette étape est de rationaliser les expériences à faire pour améliorer du mieux possible la connaissance des paramètres, c'est-à-dire la précision avec laquelle on les estime. Pour améliorer la connaissance des paramètres, nous utilisons les méthodes décrites au paragraphe 1.3.4 basées sur le calcul de la matrice d'information de Fisher. Notre stratégie propose d'acquérir de l'information grâce à des techniques statistiques, contrairement à l'algorithme de Nancy qui améliore les estimations des paramètres en utilisant les données recueillies à l'optimum de performance du procédé. Cela a pour effet de balayer une plus grande zone du domaine opératoire car l'optimum d'information n'est généralement pas au même endroit que l'optimum de performance. L'objectif est double : ne pas converger vers un optimum local en balayant le mieux possible tout le domaine opératoire et converger plus vite vers l'optimum car, à l'optimum de performance, l'information est souvent pauvre.

Pour savoir quelle expérience apportera le plus d'information, nous utilisons les critères D-, A- et E- optimal (en fonction de la géométrie de l'ellipsoïde de confiance). Si l'on dispose de plusieurs technologies de réacteur, on choisit la technologie qui améliore le plus le critère de design. Rappelons que chaque technologie a une gamme opératoire différente et une flexibilité d'échantillonnage qui lui est propre. Il n'est pas rare que le meilleur équipement pour l'acquisition de données ne soit pas le même à chaque itération.

Reprenons l'exemple de la synthèse liquide-liquide. L'objectif est maintenant de déterminer quelle technologie et quelles conditions expérimentales nous permettront d'améliorer la précision des paramètres. Le Tableau 2.11 montre le domaine opératoire du réacteur batch et du microréacteur. Vu que les temps de passage sont très courts pour le microréacteur, on admet que l'on peut faire varier à chaque échantillon d'une même expérience, la température, la concentration initiale en B et le temps de passage. En revanche, pour le réacteur batch seul un suivi temporel est possible.

Tableau 2.11 : Domaine expérimental des deux réacteurs de l'exemple.

		Batch	Microréacteur
Temps de passage	[s]	300 – 3600	20 – 200
Température	[°C]	10 – 90	10 – 90
Concentration initial B	[mol/m ³]	0 – 120	0 – 120
Aire interfaciale	[m ² /m ³]	200 – 3200	1000 – 10000

Le critère d'optimalité choisi est le critère D-optimal. On recherche l'expérience optimale pour chacun des réacteurs, sachant que l'on prélève à chaque fois 6 échantillons. Les conditions de ces expériences sont reportées dans le Tableau 2.12 ainsi que la valeur du critère D-optimal associée. Le microréacteur ayant un critère D-optimal plus petit que le batch, la prochaine expérience se fera donc en microréacteur selon les conditions d'échantillonnage présentées dans le Tableau 2.12. S'en suivra une nouvelle estimation paramétrique puis l'optimisation des performances et ainsi de suite.

Tableau 2.12 : Expériences proposées par le critère D-optimal en Batch et en Microréacteur.

N° échantillon	Batch $j_{D_{opt}} = 4,33 \times 10^8$				Microréacteur $j_{D_{opt}} = 1,40 \times 10^8$			
	t [s]	T [°C]	C_B^0 [mol/m ³]	a [m ² /m ³]	t [s]	T [°C]	C_B^0 [mol/m ³]	a [m ² /m ³]
1	300	10	120	3200	50	86	110	10000
2	600	10	120	3200	50	88	120	10000
3	2700	10	120	3200	80	70	70	10000
4	3000	10	120	3200	80	70	90	10000
5	3300	10	120	3200	80	84	100	10000
6	3600	10	120	3200	110	80	75	10000

2.1.2.5 En résumé

En résumé, la stratégie cherche à construire un modèle suffisamment prédictif pour pouvoir optimiser le système réactionnel avec une confiance au moins aussi grande que la confiance que l'on peut accorder aux mesures. Si les performances estimées et mesurées divergent significativement d'un point de vue statistique, c'est qu'il faut remettre en doute la structure du modèle. L'utilisateur doit alors apporter son sens critique pour trouver une solution. La Figure 2.4 montre l'ensemble des techniques utilisées pour chaque étape de la stratégie.

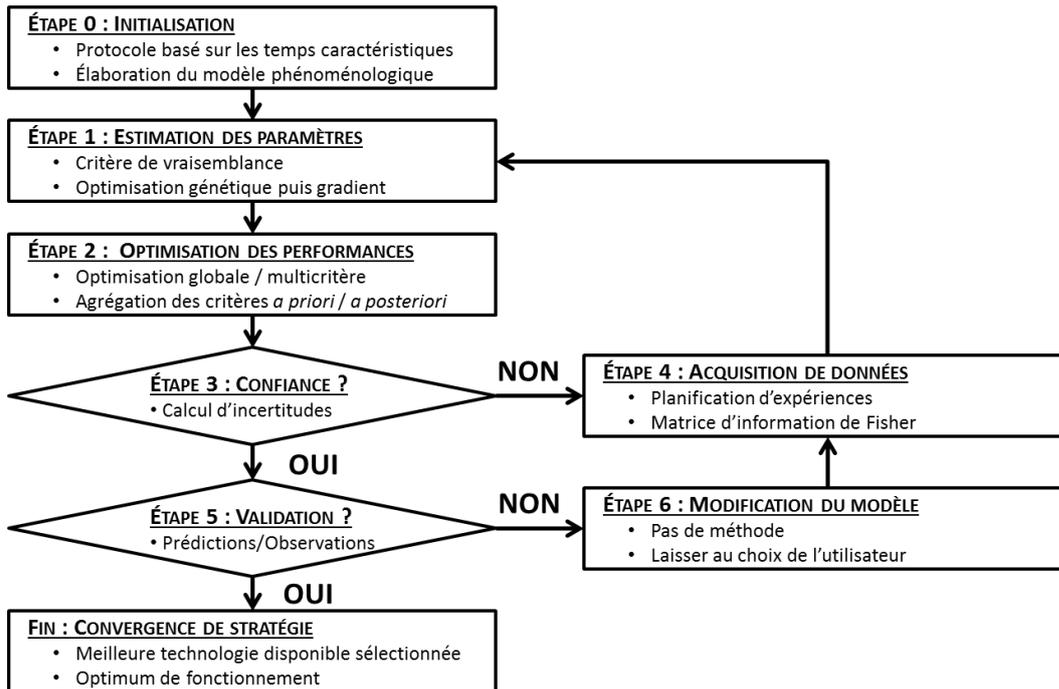


Figure 2.4 : Organigramme de la stratégie avec les outils utilisés.

2.1.3 Choix des outils numériques

La mise en œuvre de la méthodologie fait appel à plusieurs méthodes numériques qui sont détaillées ci-dessous.

2.1.3.1 Calcul des sensibilités

Les outils mathématiques utilisés par notre stratégie, en particulier le calcul de la matrice d'information de Fisher (voir 1.3.4 page 53), nécessitent de calculer les sensibilités du modèle par rapport aux paramètres.

$$s_{iu}(\boldsymbol{\xi}, \boldsymbol{\theta}) = \frac{\partial f_i}{\partial \theta_u}(\boldsymbol{\xi}, \boldsymbol{\theta}) \quad (2.6)$$

Si l'on possède une expression explicite du modèle on peut alors très souvent obtenir de manière analytique l'expression des sensibilités.

Régression polynomiale

Si en revanche le modèle est complexe et nécessite une résolution numérique on peut utiliser la méthode la plus simple, qui consiste à approximer localement la dérivée en faisant une régression polynomiale de degré p sur n points. Dans le cas où $p = 1$ et $n = 3$ cela revient à faire une différence finie centrée :

$$\frac{\partial f_i}{\partial \theta_u}(\boldsymbol{\xi}, \boldsymbol{\theta}) \approx \frac{f_i(\boldsymbol{\xi}, \boldsymbol{\theta} + \delta_u \mathbf{e}_u) - f_i(\boldsymbol{\xi}, \boldsymbol{\theta} - \delta_u \mathbf{e}_u)}{2 \delta_u}$$

$$\text{Avec } \delta_u = \theta_u(1 + h) \quad (2.7)$$

$$\text{et } \mathbf{e}_u = (0, \dots, \underbrace{1}_{u^{\text{ème}}}, \dots, 0) \quad u = 1, \dots, n_p$$

Cette méthode possède deux inconvénients majeurs. Premièrement le choix du pas de dérivation h peut influencer le résultat, il est généralement pris égal à 0,01. Le second point est le temps de calcul. En effet, pour obtenir la matrice jacobienne d'une fonction il faut faire appel à la fonction f deux fois le nombre de paramètres du modèle.

Intégration directe des sensibilités

Si le modèle est décrit par des équations différentielles ordinaires, on peut alors calculer les dérivées du modèle par rapport aux paramètres en utilisant la technique suivante proposée par Rosenbrock et Storey (1966). Un modèle décrit par des EDOs s'écrit sous la forme générale (où \mathbf{x} représente le vecteur des variables d'état) :

$$\dot{\mathbf{x}}(t) = \mathbf{h}(t, \mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\theta}) \quad \text{avec} \quad \dot{\mathbf{x}}(t=0) = \mathbf{c} \quad (2.8)$$

Les grandeurs observées à l'expérience $\boldsymbol{\xi}_k$ sont obtenues par la fonction algébrique \mathbf{g} telle que :

$$\hat{\mathbf{y}}_k(t) = \mathbf{g}(\mathbf{x}(t, \boldsymbol{\xi}_k, \boldsymbol{\theta})) = \mathbf{f}(\boldsymbol{\xi}_k, \boldsymbol{\theta}) \quad (2.9)$$

Ainsi, la matrice jacobienne \mathbf{F} du modèle \mathbf{f} par rapport aux paramètres $\boldsymbol{\theta}$ s'écrit :

i

$$\mathbf{F} = \mathbf{G} \mathbf{D}$$

$$\begin{aligned} \text{Où } \mathbf{F} &= \{F_{ij}\} = \left\{ \frac{\partial}{\partial \theta_j} f_i(\boldsymbol{\xi}_k, \boldsymbol{\theta}) \right\} \\ \mathbf{G} &= \{G_{ij}\} = \left\{ \frac{\partial}{\partial x_j} g_i(\mathbf{x}(t, \boldsymbol{\xi}_k, \boldsymbol{\theta})) \right\} \\ \mathbf{D} &= \{D_{ij}\} = \left\{ \frac{\partial}{\partial \theta_j} x_i(t, \boldsymbol{\xi}_k, \boldsymbol{\theta}) \right\} \end{aligned} \quad (2.10)$$

La matrice \mathbf{D} peut être obtenue en résolvant le système d'équations différentielles suivant, qui provient d'un développement de Taylor de l'équation (2.11) :

$$\begin{aligned} \dot{\mathbf{D}} &= \mathbf{A}(t)\mathbf{D} + \mathbf{B}(t) \quad \text{avec} \quad \mathbf{D}(t=0) = \mathbf{0} \\ \text{où } \mathbf{A}(t) &= \{A_{ij}\} = \left\{ \frac{\partial}{\partial x_j} h_i(t, \mathbf{x}, \boldsymbol{\xi}_k, \boldsymbol{\theta}) \right\} \\ \mathbf{B}(t) &= \{B_{ij}\} = \left\{ \frac{\partial}{\partial \theta_j} h_i(t, \mathbf{x}, \boldsymbol{\xi}_k, \boldsymbol{\theta}) \right\} \end{aligned} \quad (2.11)$$

Ainsi, pour obtenir les sensibilités du modèle par rapport aux paramètres, on résout donc un ensemble de $n_x + n_x \times n_p$ équations différentielles (n_x représentant le

nombre de variables d'état). Cette méthode est bien adaptée lorsque le nombre de paramètres est assez grand.

Approximation itérative de la direction de la dérivée

Cette méthode proposée par Maly et Petzold (1996) s'applique uniquement sur les systèmes décrits par des EDOs. L'idée est de calculer la sensibilité du système directement à chaque itération de la routine d'intégration. Nous ne rentrerons pas dans les détails de cette méthode qui peut être directement utilisée sous Matlab® grâce au paquet fourni par Garcia Molla (2003). La fonction `SENS_SYS` du paquet s'utilise de manière très similaire à n'importe quel solveur d'EDO intégré à Matlab®.

2.1.3.2 Reformulation des problèmes d'optimisation

La stratégie développée nécessite de résoudre des problèmes d'optimisation numérique. L'estimation paramétrique minimise le critère de vraisemblance, la planification d'expériences minimise un critère d'optimalité (ex : D-optimal) et l'amélioration des performances minimise une fonction de coût ou maximise plusieurs critères de performances. La mise en œuvre numérique peut s'avérer difficile pour trois raisons majeures :

- Les fonctions à optimiser possèdent souvent plusieurs optima locaux.
- Les problèmes sont soumis à des contraintes.
- Les nombres manipulés peuvent avoir des ordres de grandeur complètement différents.

Or, les algorithmes d'optimisation sont très efficaces quand il n'y a pas d'optima locaux, quand il n'y a pas de contraintes et quand les nombres manipulés sont proches de l'unité (pas d'un nombre très grand ou très petit). Pour rendre efficaces les algorithmes, il est souvent nécessaire de reformuler le problème pour se placer dans des conditions favorables.

Optimisation globale

Les méthodes basées sur le calcul des dérivées du critère ont l'avantage de converger très vite vers le zéro du gradient. Parmi les plus utilisées on peut citer l'algorithme du gradient, de Gauss-Newton, de Levenberg-Marquardt et de Broyden-Fletcher-Goldfarb-Shanno (BFGS). L'algorithme du simplexe, quant à lui, permet de minimiser la fonction sans pour autant calculer de dérivées.

Malheureusement, selon la complexité du problème traité et le nombre de données observées disponibles, il peut apparaître des optima locaux et les méthodes citées plus haut peuvent ne pas trouver l'optimum global si le point d'amorçage se trouve dans une mauvaise « vallée » (voir Figure 2.5). Pour éviter cela, on peut utiliser plusieurs points d'amorçage et prendre le meilleur d'entre eux ou utiliser des algorithmes d'optimisation globale tels que les algorithmes génétiques et évolutionnaires, le recuit simulé

et l'optimisation par essais particuliers. Ces méthodes sont généralement coûteuses en temps de calcul et convergent relativement lentement.

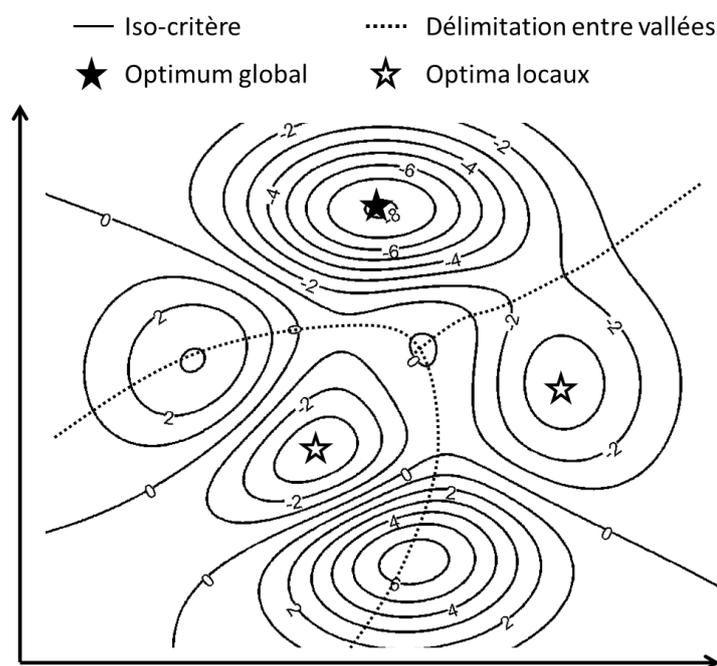


Figure 2.5 : Exemple de fonction possédant plusieurs minima locaux avec leur zone d'attraction (vallée).

L'idéal est d'utiliser des méthodes hybrides comme par exemple l'utilisation d'un algorithme évolutionnaire dont les critères de convergence sont choisis pour que l'algorithme s'arrête rapidement, après quelques itérations. Ceci afin de trouver le meilleur amorçage pour que le second algorithme qui est basé sur le calcul des dérivées puisse converger rapidement vers l'optimum global.

Enfin, lorsque les variables à optimiser ne nécessitent pas une très grande précision, c'est le cas lorsque l'on cherche à optimiser des conditions expérimentales aussi bien pour la planification que pour la recherche de performance, on peut alors discrétiser l'espace de recherche. Le fait de discrétiser l'espace des variables permet aux algorithmes génétiques de converger beaucoup plus vite et permet aussi d'éliminer les minima locaux dont les vallées associées sont plus petites que le pas de discrétisation.

Simplification des contraintes

Il existe quatre sortes de contraintes :

- La plus simple, celle qui pose le moins de problèmes aux algorithmes d'optimisation est la contrainte de bornes, l'espace de recherche est donc un hypercube.
- Ensuite, il existe des contraintes d'inégalités linéaires, ce type de contraintes est assez bien géré par les routines d'optimisation basées sur les dérivées qui utilise pour cela les multiplicateurs de Lagrange (Walter et Pronzato 1994). Les algo-

rithmes génétiques ne fonctionnent pas toujours très bien pour ce type de contraintes, surtout si le domaine autorisé est très petit devant l'hypercube de recherche.

- Les contraintes d'égalités linéaires sont relativement plus difficiles à gérer que les contraintes d'inégalités linéaires pour les algorithmes génétiques. Pour les algorithmes basés sur le calcul des dérivées, l'utilisation des multiplieurs de Lagrange permet de gérer très bien ce type de contraintes.
- Enfin, les contraintes non linéaires sont souvent assez difficiles à résoudre car le domaine autorisé n'est pas estimable de manière analytique. Souvent, la résolution des problèmes sous contraintes d'inégalité passe par l'introduction d'un terme de pénalité (Conn 1991).

L'objectif est donc de formuler le problème d'optimisation pour essayer d'avoir les contraintes les plus faciles à résoudre. Nous ne rentrerons, ici, pas dans le détail des reformulations que nous avons utilisées, mais pour éclairer le lecteur sur ce point, voici un exemple.

Prenons le cas où l'on souhaite optimiser un profil de coulée par rapport à un critère donné (qui n'a pas d'importance ici). On suppose que l'on ne fait qu'un palier (on a donc $V_{ajout} = Q \Delta t$). Dans ce cas, pour des raisons de faisabilité expérimentale, le débit de coulée ne peut être ni trop grand, ni trop petit ($Q_{min} \leq Q \leq Q_{max}$). De même, le temps de coulée doit être supérieur à un temps minimum et inférieur à un temps maximum ($t_{min} \leq \Delta t \leq t_{max}$). Enfin, le volume coulé ne doit pas dépasser le volume du réacteur ($V \leq V_{max}$). La façon la plus intuitive de poser le problème et d'optimiser le débit et le temps de coulée :

$$\begin{cases} Q_{min} \leq Q \leq Q_{max} \\ t_{min} \leq \Delta t \leq t_{max} \\ Q \cdot \Delta t \leq V_{max} \end{cases} \quad (2.12)$$

Cette formulation du problème induit deux contraintes de bornes et une contrainte non-linéaire. On peut s'affranchir de la contrainte non-linéaire, en, par exemple, optimisant le logarithme du débit et du temps de coulé. Les contraintes deviennent alors :

$$\begin{cases} \ln Q_{min} \leq \ln Q \leq \ln Q_{max} \\ \ln t_{min} \leq \ln \Delta t \leq \ln t_{max} \\ \ln Q + \ln \Delta t \leq \ln V_{max} \end{cases} \quad (2.13)$$

La contrainte d'inégalité non-linéaire a été remplacée par une contrainte d'inégalité linéaire.

La transformation d'un problème d'optimisation sous contraintes par un choix judicieux des variables à optimiser est une méthode efficace pour améliorer les performances de la routine d'optimisation. Il n'existe pas de méthodes systématiques permet-

tant de reformuler le problème dans n'importe quelle situation, mais pour certains types de problème, il existe des méthodes efficaces (Gill, Murray, et Wright 1981). Toutefois, il est important de bien vérifier que chaque nouvelle transformation ne complexifie pas d'avantage le problème.

Ordre de grandeur

Les nombres représentant des grandeurs physiques peuvent être très grands (10^{20} pour des facteurs pré-exponentiels) ou très petit (10^{-9} pour des constantes de diffusion). Or, lorsque l'on résout des équations différentielles ou que l'on optimise une fonction, les solveurs numériques ont toujours besoin de seuils de tolérance qui leur permette de savoir si le calcul a convergé ou non. Or, si l'on manipule des nombres dont les ordres de grandeur sont très variés, il est alors difficile de fixer des seuils de tolérances efficaces. De la même manière, lorsque l'on calcule des sensibilités par rapport à des paramètres, les résultats numériques peuvent être d'ordre de grandeur très différent, et du point de vue numérique, cela peut être très gênant.

Pour éviter ce problème, on peut avoir recours, par exemple, à des transformations logarithmiques qui permettent de manipuler des nombres convenables d'un point de vue numérique. Les difficultés numériques liées aux ordres de grandeurs sont souvent rencontrées lors de l'estimation de paramètres (énergie d'activation, constante de diffusion, ...).

Ce qu'il faut retenir de ce paragraphe, c'est que les routines numériques sont très efficaces quand le problème que l'on a à résoudre est facile : pas de nombres et de dérivées trop grandes ou trop petites, pas de domaine d'optimisation complexe.

2.2 Mise en place du système expérimental

Nous avons développé un montage expérimental qui nous a permis de tester notre méthodologie sur des systèmes réactionnels réels. Ce paragraphe s'attache à la description du montage expérimental.

2.2.1 Conception du montage expérimental

Le montage expérimental a été conçu pour valider le concept de notre stratégie d'intensification. Nous avons conçu, en partenariat avec l'atelier mécanique du laboratoire, un montage permettant de répondre aux différentes contraintes de pressions et de température (Figure 2.6).

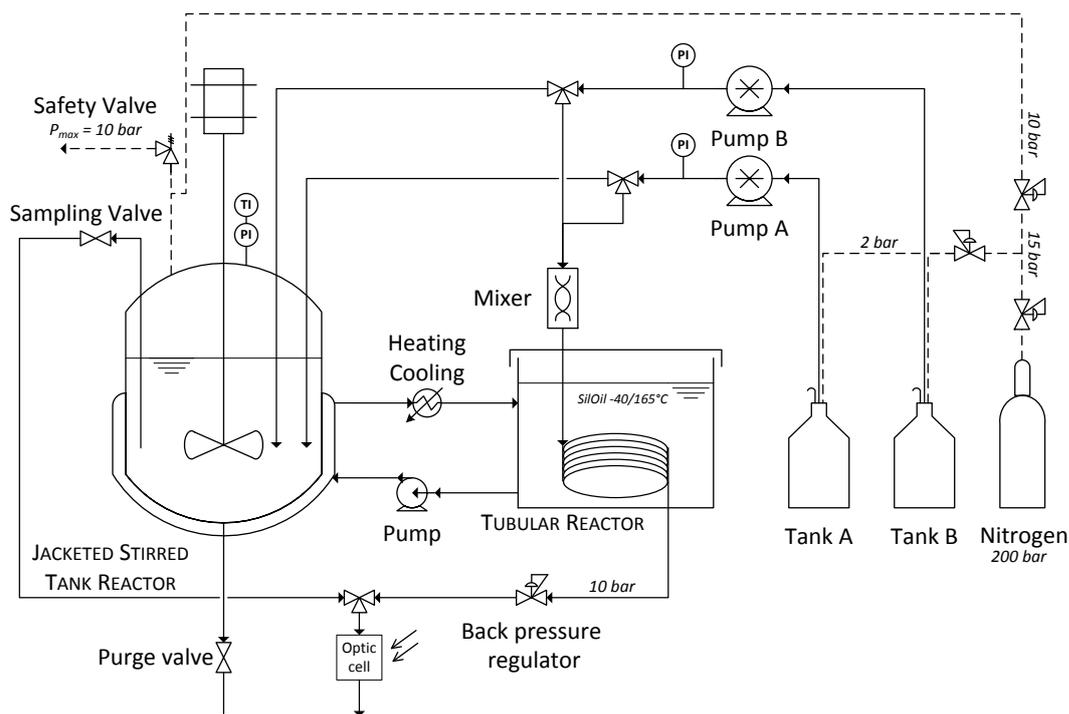


Figure 2.6 : Schéma du montage expérimental

La Figure 2.7 est une vue d'ensemble du montage.



Figure 2.7 : Photo du montage expérimental

Il doit s'adapter au maximum de synthèse chimique possible. Pour cela, le montage a été entièrement réalisé en acier inox 316L, ce qui lui permet de garantir une très bonne résistance chimique. En revanche, de par sa conception, il est plus compliqué de faire des réactions en présence de solide. Pour des synthèses gaz-liquide quelques modifications mineures doivent être apportées.

Le montage doit permettre l'utilisation de différentes technologies de réacteurs de manière très simple. Cela est rendu possible par les vannes trois voies à la sortie des pompes qui permettent d'alimenter en réactifs soit dans un réacteur tubulaire, soit dans un réacteur en cuve agitée. La cuve agitée permet d'exploiter plusieurs régimes : fermé, semi-fermé et ouvert.

Les réacteurs sont régulés en température grâce à un cryostat (Thermo Scientific - Phoenix 2). Le réacteur tubulaire est directement immergé dans le bain d'huile (Huber - SiOil M40). La cuve agitée double enveloppe est calorifugée par des tresses isolantes en laine de verre, de même que les tubes la reliant au bain d'huile.

Le système peut être pressurisé jusqu'à 10 bar afin de pouvoir travailler à des températures supérieures à la température d'ébullition du solvant à pression ambiante. La conception de la cuve agitée pressurisée fait l'objet d'un paragraphe détaillé ci-après (§ 2.2.2.1). Par ailleurs, on dispose de plusieurs réacteurs tubulaires de longueurs différentes permettant ainsi une gamme de temps de passage élargie.

Le but de notre stratégie étant d'acquérir le maximum de données, nous avons prévu une cellule optique en ligne permettant l'utilisation d'une sonde Raman ou UV. Par ailleurs, il est aussi possible d'introduire diverses sondes plongeantes dans la cuve agitée (notamment la mesure du pH).

Les pompes utilisées (Postnova analytics) sont volumétriques, prévues pour la haute pression et programmables. Elles permettent ainsi de faire des profils de coulées complexes.

Certaines réactions, en particulier la réaction utilisée au Chapitre 4, nécessitent d'avoir une atmosphère inerte. Nous avons donc prévu un système de bullage d'azote dans les flacons de réactifs. Dans la cuve agitée, la pression d'azote est utilisée pour inerte, mais aussi pour maintenir une pression supérieure à la pression d'ébullition du solvant.

Enfin, le point le plus important est la sécurisation de l'installation. L'ensemble du montage est fixé sur un châssis. Le châssis est lui-même installé sous une hotte aspirante. Le ciel gazeux de la cuve agitée est relié à une soupape de sécurité tarée à 10 bar. Les pompes sont paramétrées pour se stopper au-delà de 20 bar en cas de bouchage des lignes. L'ensemble du montage a été dimensionné pour résister à 20 bar et 200°C.

2.2.2 Caractérisation

2.2.2.1 *La cuve agitée*

La cuve agitée a été entièrement conçue et réalisée au laboratoire grâce à l'atelier de mécanique. Les dimensions de la cuve respectent les règles de dimensionnement des cuves agitées définie par Rushton, Costich, et Everett (1950). Ainsi, la hauteur du liquide est égale au diamètre de la cuve. Le diamètre de l'agitateur équivaut à un tiers du diamètre de la cuve. L'agitateur est situé par rapport au fond de la cuve à une distance égale au diamètre de l'agitateur. L'épaisseur des contres-pâles vaut $1/10^{\text{ième}}$ du diamètre de la cuve. Le volume utile de la cuve est de 800 mL.

Toute la difficulté est de pressuriser l'ensemble sans avoir recours à un entraînement magnétique de l'agitateur trop volumineux et trop onéreux. L'étanchéité de la partie tournante se fait grâce à un joint à lèvres en PTFE chargé verre et MoS₂. Le palier

de l'axe d'agitation est muni de deux roulements à billes ainsi qu'un accouplement flexible (non visible sur le schéma) qui assure le bon maintien de l'axe malgré les vibrations du moteur d'agitation.

Les raccords de la double enveloppe sont normés au pas M16x100 de manière à ce qu'ils soient compatibles avec les raccords des bains thermostatés classiques. Quant aux raccords du couvercle, ils sont réalisés sous la norme NPT gaz pour s'assurer l'étanchéité.

Les Figure 2.8, Figure 2.9 et Figure 2.10 sont respectivement la vue en coupe 3D, la vue de dessus et la coupe axiale de la cuve agitée.

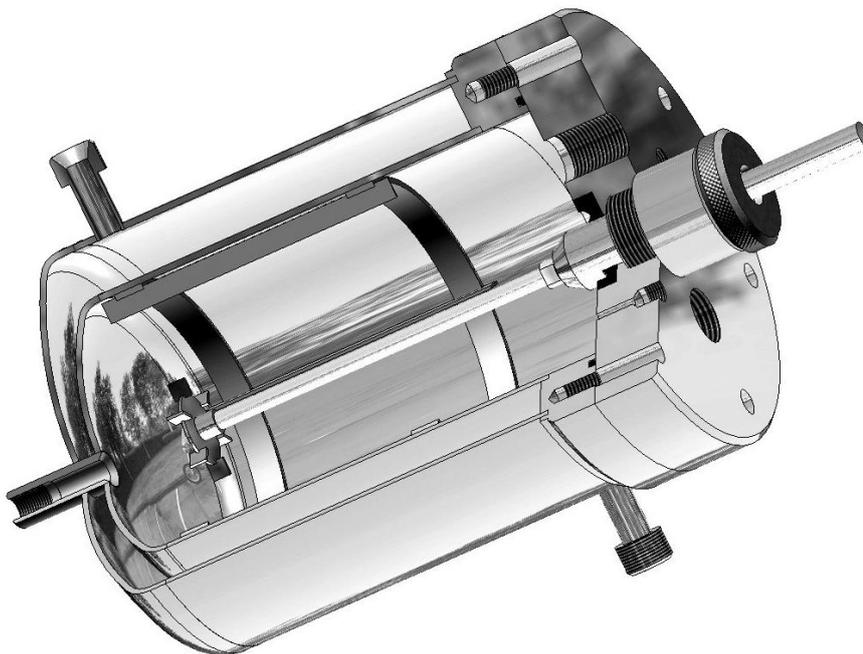


Figure 2.8 : Vue 3D de la cuve agitée

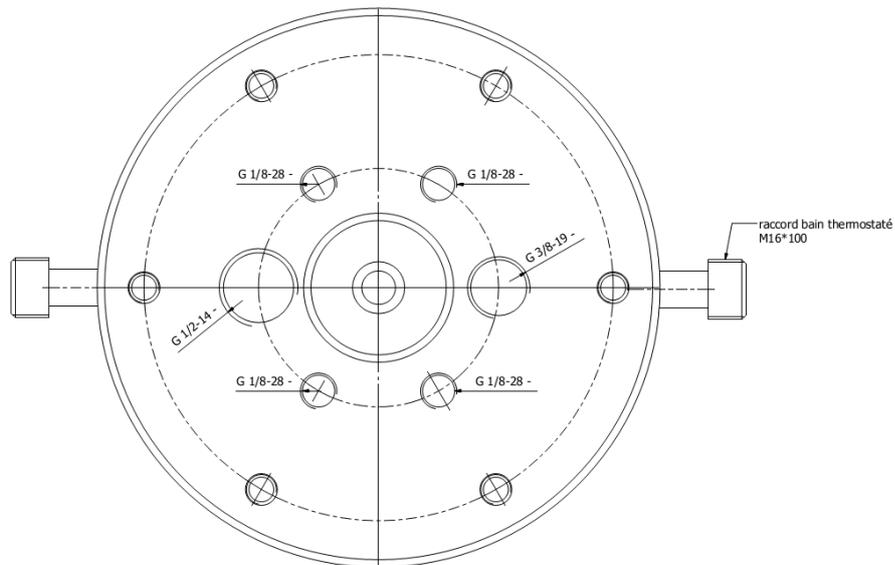


Figure 2.9 : Vue de dessus de la cuve agitée

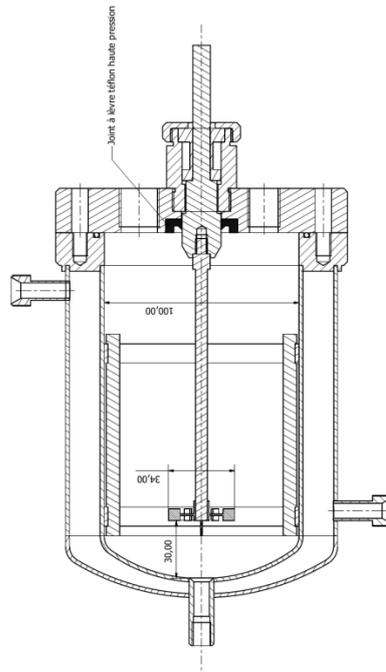


Figure 2.10 : Vue en coupe axiale de la cuve agitée

Aspect thermique

La caractérisation thermique du réacteur est complexe de par sa conception. En effet, plusieurs points difficiles à représenter sont à prendre en compte pour avoir une représentation fine du bilan thermique du réacteur :

- Le transfert thermique entre le milieu réactionnel et la double paroi (noté U_{int}) est le point le plus facile à représenter. Il existe plusieurs corrélations permettant d'évaluer le coefficient

d'échange à la paroi d'une cuve agitée munie des contres-pâles (Chapman, Dallenbach, et Holland 1964).

- L'inertie thermique du réacteur, Γ_r . En effet, le réacteur est relativement lourd (7,8kg d'acier) et sa capacité thermique ($\Gamma_r = 3600 \text{ J/K}$) n'est pas négligeable devant celle du milieu réactionnel ($mCp = 3350 \text{ J/kg}$). Vu sa géométrie, il est fort probable que sa température ne soit pas homogène et modéliser son influence sur la dynamique thermique du système est relativement complexe.
- Les pertes thermiques vers le milieu extérieur, U_{ext} , liées au refroidissement aérien sont elles aussi difficiles à quantifier par calcul.

Il semble donc difficile de faire une analyse fine de la thermique du système. Cependant, nous avons estimé le coefficient de transfert global à la paroi et le coefficient de pertes thermique par corrélation. La capacité calorifique de la cuve a été estimée grâce aux propriétés thermodynamique de l'acier utilisé. Pour faciliter les calculs, nous avons fait les approximations géométriques explicitées par la Figure 2.11, à savoir :

- L'échange thermique à la paroi se fait exclusivement par la surface verticale.
- Le fluide caloporteur se déplace verticalement le long de la double enveloppe.
- Les pertes thermiques peuvent être assimilées aux pertes thermiques d'une surface plane verticale.

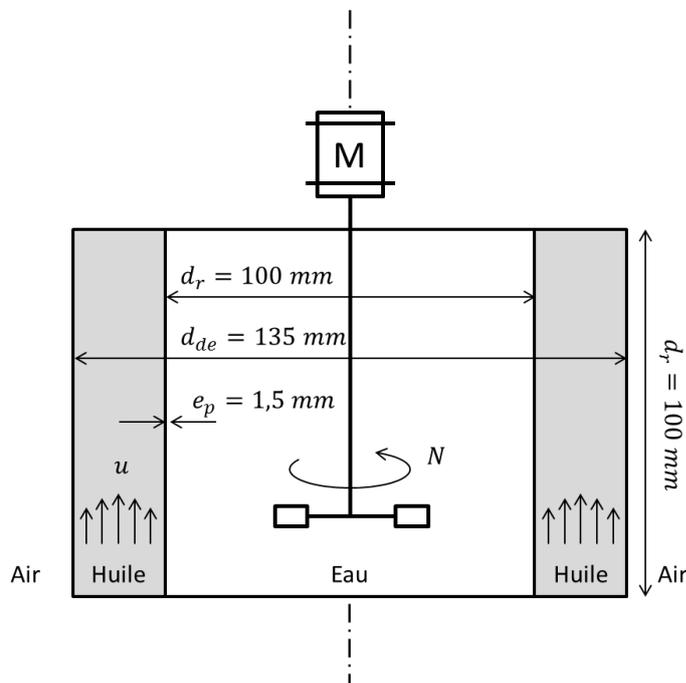


Figure 2.11 : Schéma du modèle thermique de la cuve.

Pour estimer le coefficient de transfert global à la paroi nous utilisons les corrélations proposées dans le livre de D. Schweich (Schweich 2001) :

$$\frac{1}{U} = \frac{1}{h_{int}} + \frac{A_{int} e_p}{A_m \lambda_p} + \frac{A_{int}}{A_{ext}} \frac{1}{h_{ext}} \quad (2.14)$$

Pour obtenir h_{int} on utilise la corrélation de Bourne, Buerli, et Regenass (1981) qui est valable pour les cuves agitées. Les propriétés de l'eau sont évaluées à 80°C :

$$Nu_{int} = \frac{h_{int} d_r}{\lambda_{eau}} = 0,29 Re^{0,678} Pr^{0,33}$$

avec :

$$Re = \frac{\rho_{eau} N d_A}{\mu_{eau}} \quad (2.15)$$

$$Pr = \frac{\mu_{eau} C_{p_{eau}}}{\lambda_{eau}}$$

Le terme h_{ext} s'obtient quant à lui en utilisant une corrélation provenant de Martin et Stephan (1983) s'appliquant en régime laminaire dans une conduite annulaire dont le transfert se fait sur le tube interne seul, les propriétés du caloporteur étant fournies par le fournisseur (ref : Huber SilOil M40.165.10) évalué à 80°C :

$$Nu_{ext} = \frac{h_{ext} d_H}{\lambda_{huile}} = (Nu_{\infty}^3 + Nu_L^3)^{0,33} Vis^{0,14}$$

avec :

$$Nu_{\infty} = 7,54 (1 + 0,328 K^{-0,8})$$

$$Nu_L = 2,935 Gz^{0,33} (1 + 0,14 K^{-0,5})$$

$$K = d_r / d_{de} \quad (2.16)$$

$$d_H = d_{de} - d_r$$

$$Gz = \frac{\rho_{huile} u C_{p_{huile}} d_H^2}{\lambda d_r}$$

$$Vis = \frac{\mu_{huile}}{\mu_{eau}}$$

L'estimation des pertes thermiques par convection d'une surface plane verticale est donnée par la corrélation de Churchill et Chu (1975) :

$$Nu = \left(0,6 + \frac{0387 Ra_{\delta}^{\frac{1}{4}}}{[1 + (0,559/Pr)^{9/16}]^{8/27}} \right)^2 = \frac{h_{conv} \lambda_{air}}{d_r} \quad (2.17)$$

Auxquelles il faut ajouter les pertes thermiques par rayonnement qui s'obtiennent par la relation :

$$h_{ray} = \varepsilon_{acier} \sigma (T_r^2 + T_{air}^2) (T_r + T_{air}) \quad (2.18)$$

Les propriétés thermodynamiques utilisées pour les calculs des corrélations proviennent du fabricant pour l'huile et pour l'air et l'eau du Perry's handbook (Perry et Green 2008) sont détaillées dans le tableau ci-dessous :

Tableau 2.13 : Propriétés thermodynamique des fluides.

		Air	Huile	Eau	Acier
T	[°C]	20	80	80	80
ρ	[kg/m ³]	1,184	888	9710	–
μ	[Pa.s]	–	$4,17 \times 10^{-3}$	$0,36 \times 10^{-3}$	–
λ	[W/m/K]	0,0263	0,125	0,660	14,5
Cp	[J/K/kg]	–	1680	4180	–
ε	[–]	–	–	–	0,55

Les résultats obtenus sont donnés dans le Tableau 2.14 page 93. Par ailleurs, nous avons tout de même tenté d'estimer le coefficient de transfert global à la paroi, le coefficient de pertes thermique et la capacité calorifique en réalisant trois expériences.

- La première a pour but d'estimer les pertes thermiques du réacteur liées à la convection forcée de la hotte aspirante. Pour cela, on chauffe le réacteur rempli d'eau à la température de consigne $T_c = 80^\circ\text{C}$ puis on laisse refroidir l'ensemble en coupant la régulation de température du bain thermostaté. L'équation différentielle régissant cette expérience est donc :

$$(mCp + \Gamma_r) \frac{dT}{dt} = \bar{h}_{ext} A_{ext} (T_{air} - T) \quad (2.19)$$

- La deuxième consiste à préchauffer le bain d'huile ainsi que le réacteur à la température de consigne ($T_c = 80^\circ\text{C}$). Lorsque le système est à l'équilibre thermique, on introduit de l'eau froide dans le réacteur. En procédant ainsi, on peut s'affranchir de la capacité thermique du réacteur. On obtient alors :

$$(mCp) \frac{dT}{dt} = \bar{h}_{int} A_{int} (T_c - T) + \bar{h}_{ext} A_{ext} (T_{air} - T) \quad (2.20)$$

- Enfin, la dernière expérience prend en compte tous les termes simultanément. On préchauffe le bain d'huile à la température de consigne sans pour autant chauffer le réacteur qui doit être à l'équilibre thermique avec l'eau qu'il contient. A $t = 0$, on fait passer le fluide caloporteur dans la double enveloppe. Le bain thermostaté n'est pas assez puissant pour maintenir la température du fluide caloporteur constante, il faut donc en tenir compte dans l'équation différentielle :

$$(mCp + \Gamma_r) \frac{dT}{dt} = \bar{h}_{int} A_{int} (T_c(t) - T) + \bar{h}_{ext} A_{ext} (T_{air} - T) \quad (2.21)$$

L'estimation paramétrique réalisée grâce aux relevés de la température de l'eau contenue dans le réacteur et de la température du bain d'huile sont représentés sur la Figure 2.12 et donne les résultats reportés dans le Tableau 2.14.

Tableau 2.14 : Caractéristiques thermiques de cuve agitée obtenues par corrélation et par expérimentation

		Corrélation	Expérimentation
Γ_r	[J/K]	3600	660 ± 404
\bar{h}_{int}	[W/m ² /K]	323	302 ± 23
\bar{h}_{ext}	[W/m ² /K]	11,1	$10,4 \pm 1,3$

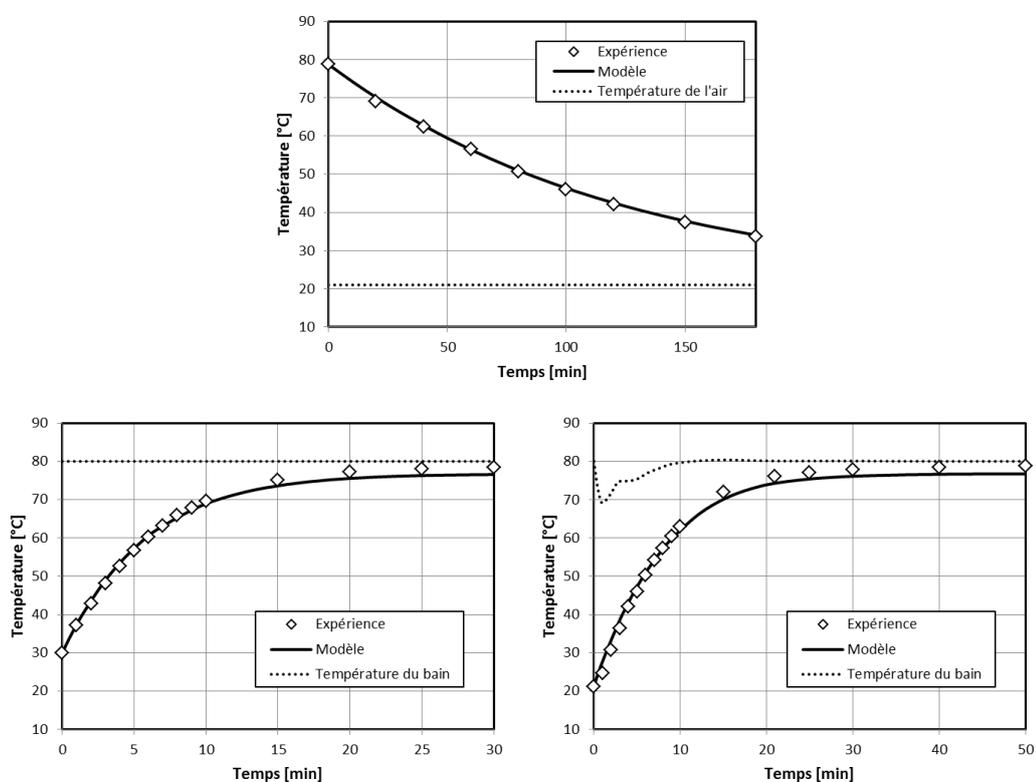


Figure 2.12 : Regressions paramétriques réalisées dans le but de caractériser thermiquement la cuve agitée.

Le Tableau 2.14 donne les résultats obtenus par corrélation et par estimation paramétrique. L'estimation de la capacité thermique du réacteur est peu précise et est très en deçà de la valeur attendue ($m_r C p_{acier} = 3600$ J/K), ce qui prouve que le réacteur n'est pas homogène en température. En revanche, les coefficients de transfert de la double enveloppe global et du refroidissement aérien sont quant à eux estimés de manière assez précise et correspondent assez bien aux valeurs obtenues par corrélations.

Bien que le réacteur ait été caractérisé thermiquement, il semble périlleux d'essayer d'obtenir des informations cinétiques en utilisant la mesure de la température du réacteur. Un réacteur calorimétrique serait beaucoup plus adapté. Ainsi, les mesures de

température du milieu réactionnel et du bain d'huile n'apporteront que des informations qualitatives sur l'exothermicité des réactions et ne pourront pas être utilisées à des fins cinétiques.

2.2.2.2 Le réacteur tubulaire

Le montage dispose de deux réacteurs tubulaires de longueurs différentes (1m et 6m) et de diamètre externe de 1/8 pouce. Les deux réacteurs tubulaires, constitués par deux bobines, sont directement plongés dans le bain thermostaté (voir Figure 2.7). Il est possible d'utiliser soit une seule bobine, soit les deux montées en séries. Entre la fin du réacteur et la cellule optique, il y a 55 cm de tube 1/8 po.

Aspect mélange

À l'entrée de la bobine, les deux flux sont mélangés dans un Té de 1/16^e de pouce dont le diamètre interne est de 0,05 po soit 1,27 mm suivi d'un petit tube en 1/16^e de pouce long de 4 centimètres. Ce petit tube se justifie par le calcul de la longueur caractéristique de mélange. En effet, on peut calculer le temps de mélange dans un tel dispositif grâce à l'équation suivante (Falk 2010) :

$$t_{mix} \approx 0,15 \varepsilon^{-0,45}$$

$$\text{Avec } \varepsilon = \frac{32 \mu u^2}{\rho D_h^2} \quad \text{et} \quad u = \frac{4Q}{\pi D_h^2} = \frac{l_{mix}}{t_{mix}} \quad (2.22)$$

Ainsi, on peut calculer la longueur de tube 1/16^e de pouce nécessaire pour atteindre le mélange en fonction du débit d'entrée.

$$l_{mix} = \frac{4Q}{\pi D_h^2} t_{mix} \approx 0,036 \frac{Q^{0,1} D_h^{0,7} \rho^{0,45}}{\mu^{0,45} \pi^{0,1}} \quad (2.23)$$

La Figure 2.13 montre l'évolution de la longueur caractéristique de mélange en fonction du débit.

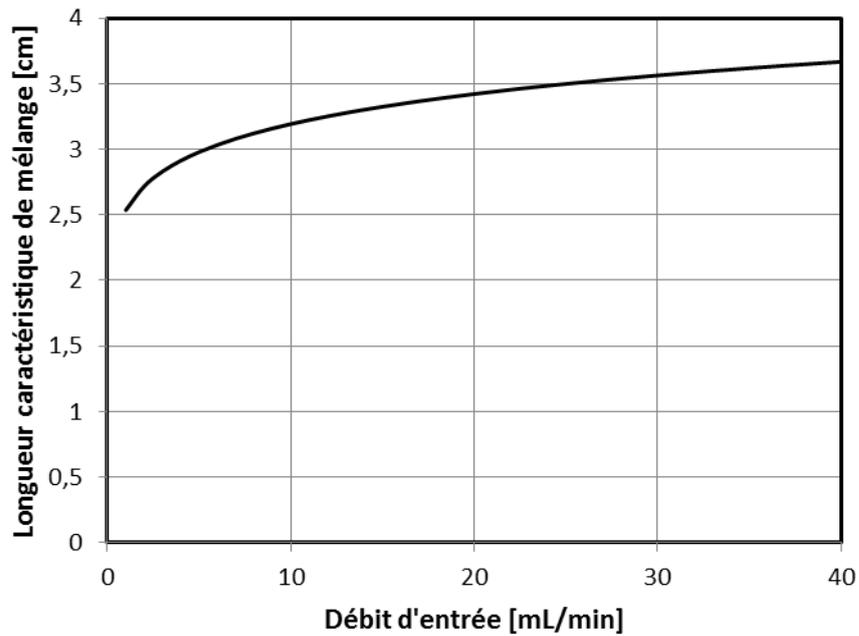


Figure 2.13 : Évolution de la longueur caractéristique de mélange en fonction du débit d'entrée dans le mélangeur en Té de 1/16^e de pouce.

Ainsi, dans la plage de débit admissible par les pompes il suffit d'un peu moins de 4 centimètres de tube en 1/16^e de pouce pour avoir un mélange correct.

Distribution des temps de séjour du réacteur tubulaire

La caractérisation de la distribution des temps de séjour (DTS) du réacteur tubulaire a été réalisée par un signal d'entrée de type échelon. Pour ce faire, nous avons établi dans le réacteur un écoulement avec de l'eau pure, puis nous avons basculé l'entrée par de l'eau colorée à la nigrosine. L'absorption de la nigrosine à 300nm est le signal mesuré $F(t)$, signal que nous avons pris à l'entrée et à la sortie du réacteur voir Figure 2.14.

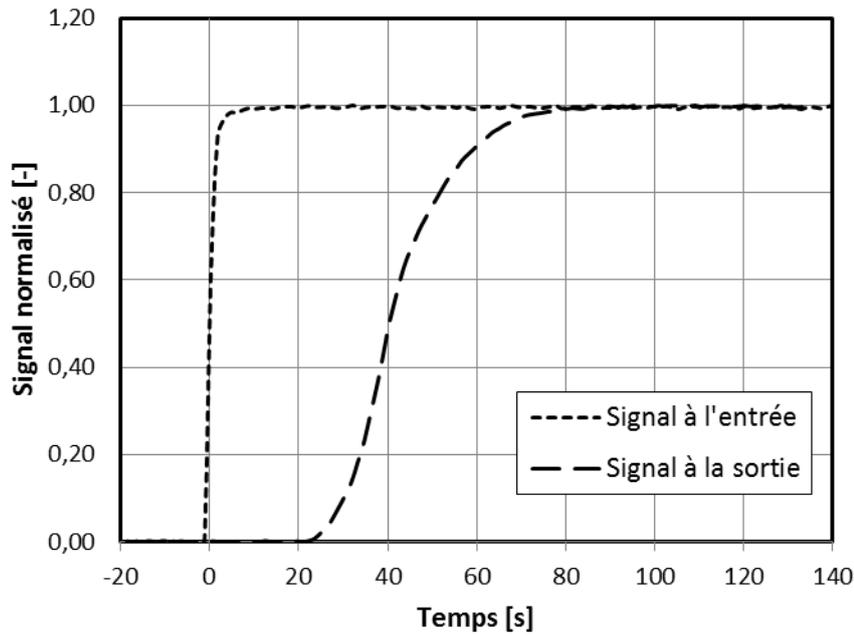


Figure 2.14 : Signal à l'entrée et à la sortie du réacteur de 6 mètres de long avec un débit de 15 mL/min

On peut calculer la valeur de la moyenne \bar{t} et de la variance σ^2 pour le signal d'entrée et le signal de sortie grâce aux expressions (2.24) et (2.25) :

$$\bar{t} = \int_0^{\infty} (1 - F(t)) dt \tag{2.24}$$

$$\sigma^2 = 2 \int_0^{\infty} t(1 - F(t)) dt - \bar{t}^2 \tag{2.25}$$

On obtient alors la valeur moyenne et la variance de la distribution de temps de séjour grâce à la relation de Van der Laan :

$$\begin{cases} \bar{t}_{DTS} = \bar{t}_{sortie} - \bar{t}_{entree} \\ \sigma_{DTS}^2 = \sigma_{sortie}^2 - \sigma_{entree}^2 \end{cases} \tag{2.26}$$

Le Tableau 2.15 reporte les résultats :

Tableau 2.15 : Résultat de la DTS du réacteur de 6 mètres à 15mL/min

		Entrée	Sortie	DTS
\bar{t}	[s]	0	61	61
σ	[s]	1	11	10

En modélisant la distribution de temps de séjour obtenue en supposant que la DTS provient d'une dispersion axiale, on obtient l'équation suivante :

$$E(t) = \frac{1}{2} \left(\frac{P}{\pi \tau t} \right)^{1/2} \exp \left(-\frac{P (\tau - t)^2}{4 \tau t} \right) \quad (2.27)$$

où le paramètre P , qui caractérise la DTS, est obtenu par les moindres carrés. On obtient alors la distribution de temps de séjours théorique ($\hat{P} = 109$, ce qui en pratique est assimilé à un réacteur piston). La Figure 2.15 la compare à la courbe expérimentale obtenue par dérivation du signal $F(t)$.

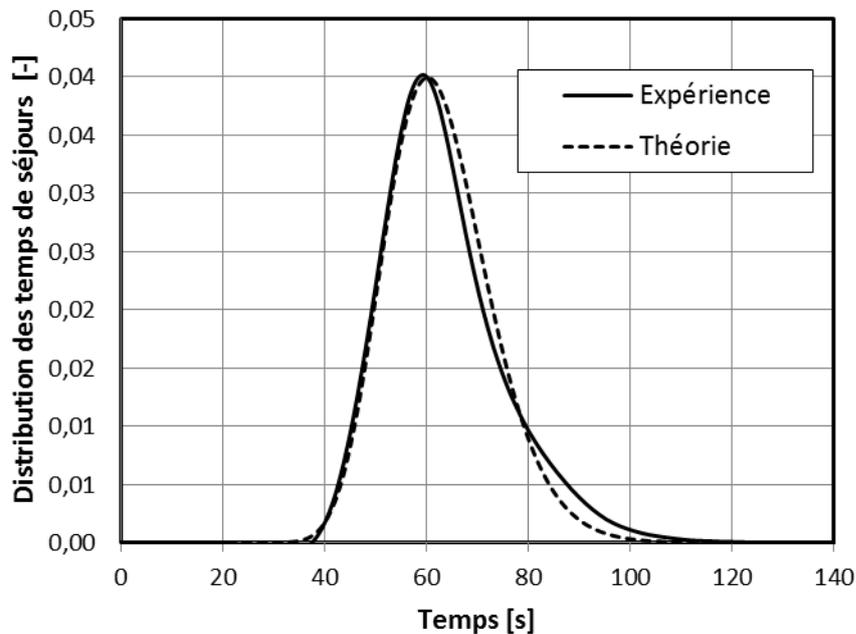


Figure 2.15 : Comparaison entre la DTS expérimentale et la théorique dans le réacteur tubulaire.

La courbe théorique est en bonne adéquation avec la courbe mesurée. L'estimation du paramètre P s'exprime en fonction du paramètre de diffusion axiale D , de la vitesse du fluide u et de la longueur du tube l . Cela permettra d'estimer la DTS dans d'autres conditions lorsque cela sera nécessaire.

$$P = \frac{u l}{D} \quad \text{d'où} \quad D = \frac{u l}{P} = \frac{0,104 \times 6,55}{109} = 0,0062 \text{ m}^2/\text{s} \quad (2.28)$$

Aspect thermique

Nous avons vérifié que le fluide soit mis en température le plus rapidement possible dans le réacteur, pour que l'on puisse admettre que la température est homogène le long du réacteur. Nous n'avons pas les moyens de mesurer la température du fluide T le long du réacteur, nous l'avons donc estimé par le calcul. Pour cela nous avons établi le bilan thermique du tube plongé dans le bain d'huile, ce dernier étant maintenu à la température T_b .

$$\dot{m} C_p \frac{dT}{dz} = \pi D_h h (T_b - T) \quad (2.29)$$

En intégrant l'équation (2.29) on obtient :

$$\frac{T(z) - T_b}{T(0) - T_b} = \exp\left(-\frac{h \pi D_h z}{\dot{m} C_p}\right) \quad (2.30)$$

On peut donc calculer la longueur de tube nécessaire l_{min} pour atteindre la température du bain d'huile à une tolérance $\pm \Delta T_{tol}$:

$$l_{min} = -\frac{\dot{m} C_p}{h \pi D_h} \ln\left(\frac{\Delta T_{tol}}{T(0) - T_b}\right) \quad (2.31)$$

La relation (2.31) est linéaire au débit de fluide (dans ce cas de l'eau). La Figure 2.16 montre l'évolution de l_{min} en fonction de la différence de température entre le fluide entrant et la température du bain, en se plaçant au débit maximum $Q = 40 \text{ mL/min} \Leftrightarrow \dot{m}_{max} = 6,66 \times 10^{-4} \text{ kg/s}$:

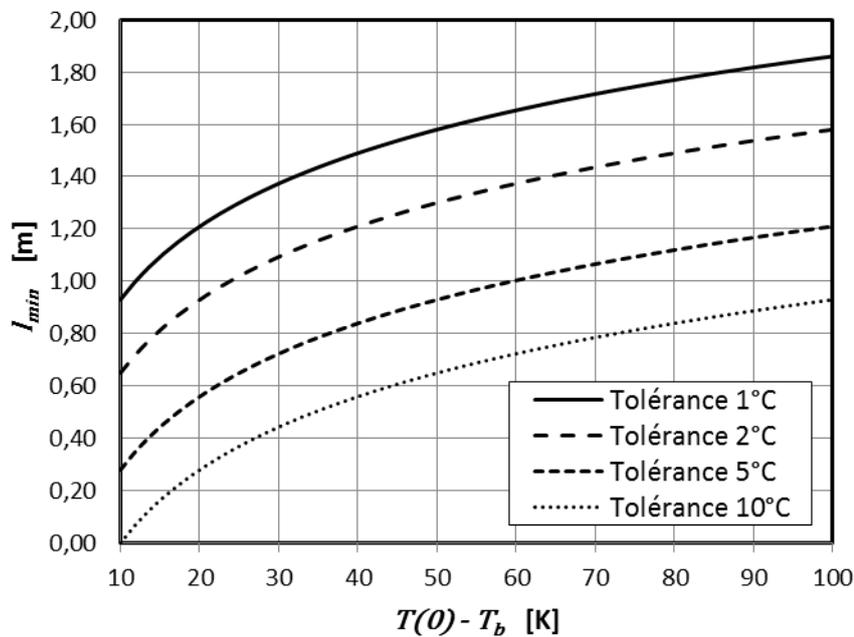


Figure 2.16 : Évolution de la longueur minimum de chauffe pour atteindre la tolérance en fonction de l'écart entre la température du bain et du fluide à l'entrée du réacteur pour un débit de 40mL/min.

La Figure 2.17 montre l'évolution de la différence de température entre fluide et le bain d'huile le long du réacteur au débit de fluide maximal ($Q = 40 \text{ mL/min}$). Ainsi, dans les conditions les plus défavorables (fort débit, grand écart de température), la température n'est pas homogène le long du réacteur, et il faudra en tenir compte dans la modélisation.

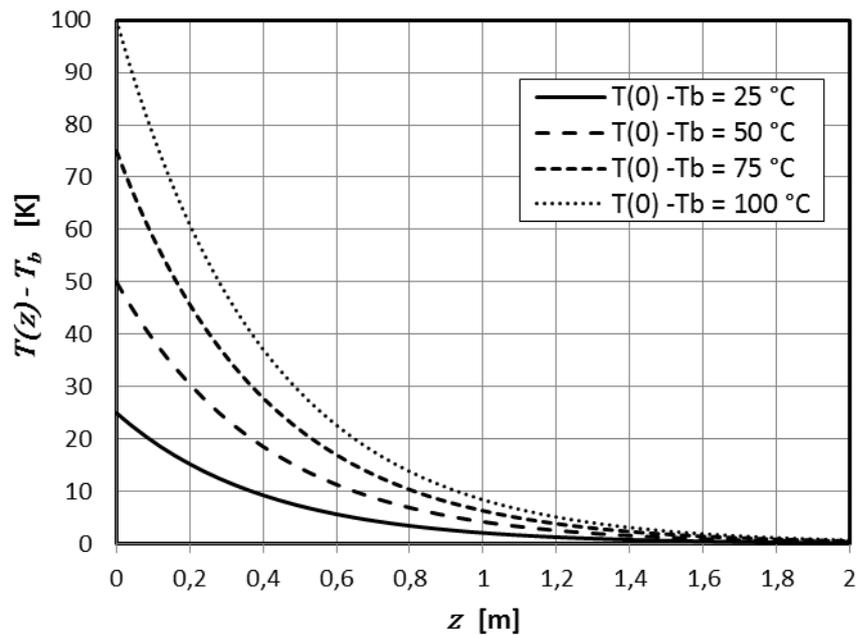


Figure 2.17 : Évolution de la différence de température entre fluide et le bain d'huile le long de réacteur au débit de fluide maximal.

2.3 Conclusion

La prise en compte de plusieurs technologies a de fortes implications sur les techniques mathématiques à utiliser. Nous avons opté pour l'utilisation de l'approche phénoménologique car elle est la seule à gérer l'utilisation de différentes technologies. Notre stratégie, alterne successivement des phases de calculs et des phases expérimentales. Le fait qu'elle soit itérative permet la création d'un modèle juste suffisant. C'est à dire que la précision de ses prédictions est de l'ordre de grandeur de la précision de la mesure expérimentale, et ce, à l'optimum de performance. La gestion des différentes technologies, qui a première vue peut être considérée comme une contrainte, se révèle, dans les chapitres qui suivent, un atout non négligeable pour la phase d'acquisition de données.

Dans les prochains chapitres, nous testerons la stratégie de manière virtuelle afin de s'assurer de sa robustesse, puis nous l'appliquerons à un cas expérimental relativement simple avant de l'appliquer à un cas industriel réel.

Chapitre 3

Application virtuelle de la stratégie

L'intérêt d'appliquer virtuellement la stratégie est de pouvoir, dans un intervalle de temps raisonnable, simuler une très grande quantité d'expériences. Ainsi, nous avons pu mettre à l'épreuve la stratégie et valider son bien-fondé, sur quelques exemples. Pour cela nous avons cherché à répondre aux quatre questions suivantes :

- La stratégie converge-t-elle toujours de la même façon (répétabilité) ?
- la stratégie converge-t-elle encore en faisant varier le nombre d'espèces mesurées, l'erreur expérimentale et la rapidité des phénomènes (robustesse) ?
- Est-il vraiment intéressant de prendre en compte les différentes technologies pour l'acquisition de données ?
- Aurait-on pu faire aussi bien avec des stratégies expérimentales existantes ?

Les quatre dernières sections de ce chapitre résument les réponses à ces questions tandis que le premier s'attache à décrire les systèmes expérimentaux virtuels utilisés dans ce chapitre. Bien sûr, les expériences virtuelles ne sont pas suffisantes et ce ne sont qu'un premier pas vers la démonstration de l'efficacité de la stratégie.

Les expériences virtuelles

Pour simuler des points expérimentaux dans les conditions expérimentales ξ données, nous calculons les grandeurs vraies tel que $\mathbf{y}^* = f(\xi, \theta^*)$, où f et θ^* représentent respectivement la structure du modèle et les paramètres théoriques fixés de manière arbitraire au paragraphe 3.1. Puis, pour créer les données expérimentales virtuelles $\tilde{\mathbf{y}}$, nous ajoutons un bruit aléatoire gaussien multivarié de covariance Σ_y tel que :

$$\tilde{\mathbf{y}} = f(\xi, \theta^*) + \mathcal{N}(\mathbf{0}, \Sigma_y) \quad (3.1)$$

La matrice de covariance est elle aussi définie arbitrairement au paragraphe 3.1. Dans ce chapitre, le modèle utilisé lors de l'application de la stratégie Stanislas est le même que celui utilisé pour la simulation des points expérimentaux. Lorsque les expériences sont bien réelles (Chapitre 4 et 5), le modèle que l'on utilise avec la stratégie n'est pas connu, ce qui rend les choses plus ardues. Ainsi, dans ce chapitre, nous étudions spécifiquement l'influence du caractère aléatoire des mesures sur le comportement de la stratégie.

Comment juger de la pertinence de la stratégie ?

Pour bien comprendre sur quels critères est jugée la stratégie il est nécessaire de bien définir les grandeurs utilisées. La Figure 3.1 résume ces différentes notions. Le premier point important est le fait que l'on connaisse les vraies conditions optimales de performances, puisque le modèle qui sert à générer les points expérimentaux simulés a été défini par nous-même. Nous noterons les performances optimales vraies γ^* . Ces performances optimales vraies sont obtenues dans les conditions expérimentales optimales vraies, notées ξ^* . Les critères de performance estimés par la stratégie sont notés $\hat{\gamma}$ et sont obtenus dans les conditions expérimentales optimales estimées $\hat{\xi}$. Dans les conditions $\hat{\xi}$, si l'on répétait indéfiniment la même mesure, on obtiendrait ce que l'on appelle les critères de performance observés réels $\hat{\gamma}^*$ (que nous appellerons simplement « observés »).

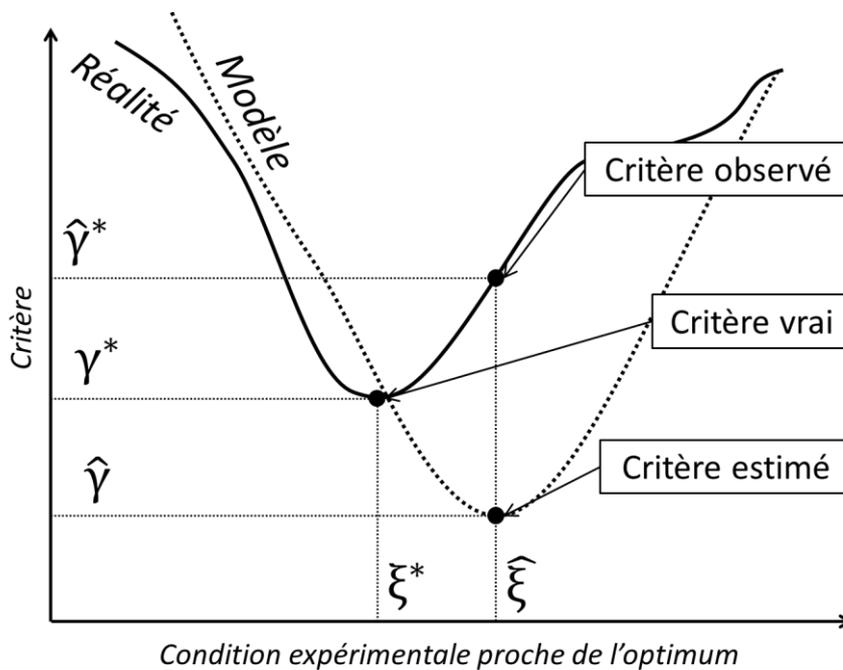


Figure 3.1 : Comparaison entre critères estimé, observé et vrai.

Afin de pouvoir juger de la réussite de la stratégie, nous comparons les critères de performance estimés $\hat{\gamma}$ et les critères observés $\hat{\gamma}^*$ aux critères vrais γ^* . Il est important de faire les deux comparaisons (estimés/vrais et observés/vrais) car certains critères, comme le temps de synthèse, ne dépendent pas directement de grandeurs mesurées. Il n'est donc pas possible d'évaluer un intervalle de confiance pour les critères indépendants des grandeurs mesurées. Cela implique que la stratégie peut avoir convergé sans pour autant avoir atteint l'optimum vrai. Dans l'exemple présenté dans la Figure 3.2, le rendement observé est, pour les deux cas, dans l'intervalle de confiance du rendement estimé. Dans les deux cas, le critère de convergence de la stratégie est atteint. Mais, nous ne sommes pas supposés connaître la courbe « réalité ». Ceci a pour conséquence que, dans le cas de la figure de droite, les conditions optimales vraies sont assez éloignées des conditions observées, bien que la stratégie ait convergé. Ainsi, lorsque la

stratégie a convergé, cela assure un certain niveau de performance, mais n'assure en rien le fait d'avoir effectivement trouvé le meilleur optimum de performance possible.

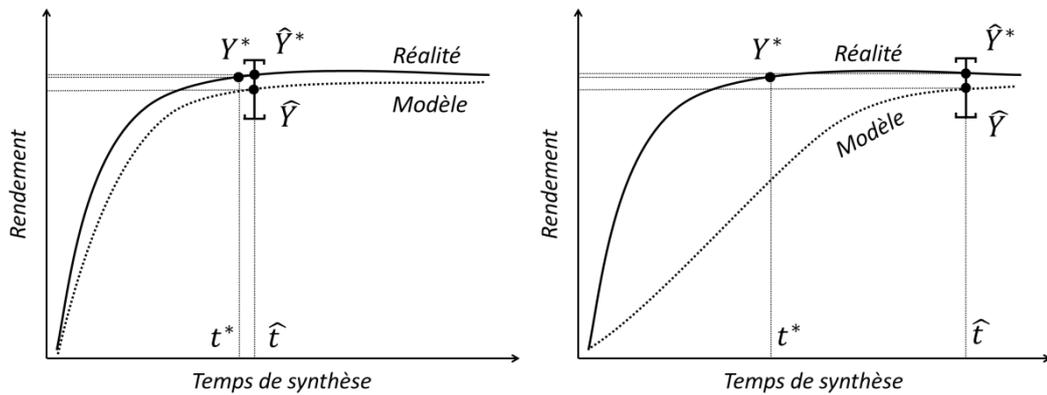


Figure 3.2 : Comparaison de deux cas où la stratégie converge mais n'atteint pas forcément l'optimum vrai.

Comme le nombre de critères est souvent supérieur à deux, il est assez complexe de visualiser graphiquement les différents critères. Nous avons donc défini un écart relatif aux seuils vetos et objectifs. Cela nous permet d'avoir une compréhension plus aisée :

$$\text{err}(\hat{\mathcal{Y}}) = 100 \times \frac{1}{n_j} \sum_{i=1}^{n_j} \left| \frac{\hat{\mathcal{Y}}_i - \mathcal{Y}_i^*}{v_i - o_i} \right| \quad (3.2)$$

et

$$\text{err}(\hat{\mathcal{Y}}^*) = 100 \times \frac{1}{n_j} \sum_{i=1}^{n_j} \left| \frac{\hat{\mathcal{Y}}_i^* - \mathcal{Y}_i^*}{v_i - o_i} \right| \quad (3.3)$$

Enfin, nous portons attention au nombre d'itérations nécessaires à la convergence de la stratégie.

3.1 Description des cas d'études

Pour valider notre stratégie de manière virtuelle nous avons créé les quatre systèmes réactionnels présentés ci-dessous. Le choix des systèmes réactionnels a été fait dans le but d'utiliser les principaux schémas réactionnels en phase liquide. Ils font intervenir, des constantes pré-exponentielles, des énergies d'activation, des ordres de réaction et des coefficients de transfert liquide-liquide. L'étude de répétabilité, le test de robustesse et la comparaison à d'autres stratégies ont été effectués sur chaque système réactionnel.

3.1.1.1 Système n°1 : consécutif-parallèle simplifié

Ce système simplifié est un système consécutif-parallèle qui ne dépend pas de la température :



Ce système est un cas d'école pour l'optimisation de la sélectivité. Dans cette version simplifiée, l'objectif est donc de trouver un compromis entre le temps de synthèse et le rendement en R. On suppose que seul le composé B est mesuré avec un écart-type correspondant à 5% de la concentration initiale en réactif C_{TOT}^0 , qui dans ce cas est fixée à $0,1 \text{ mol/m}^3$ (soit $\sigma_B = 0,005 \text{ mol/m}^3$). Pour réaliser cette synthèse, nous supposons que trois réacteurs sont disponibles, un réacteur batch, un piston et un semi-batch. Le détail des conditions accessibles se trouve dans le Tableau 3.1. À chaque expérience, cinq échantillons sont prélevés. En réacteur batch et semi-batch, les échantillons sont prélevés à différents temps de séjour. En réacteur piston, les cinq échantillons sont prélevés à différents temps de séjour et différent ratio initial de réactif.

Tableau 3.1 : Gamme opératoire du système n°1.

		Batch	Piston	Semi-Batch
t	[s]	300 – 3600	10 – 200	300 – 3600
$Ratio = [A]^0 / C_{TOT}^0$	[%]	5 – 95	5 – 95	5 – 95
Coulée B	[–]	aucune	aucune	4 paliers
Volume du réacteur	[mL]	800	15	800
Échantillonnage		t	t et $Ratio$	t

Le modèle représentant l'évolution des concentrations est paramétré par le logarithme base 10 des constantes de vitesse (pour faciliter la variation ultérieure, § 3.3.2, du temps caractéristique de chaque réaction de plusieurs ordres de grandeur). Les valeurs des paramètres ont été fixées arbitrairement et sont détaillées dans le Tableau 3.2.

Tableau 3.2 : Valeur théorique des paramètres du système n°1.

		k_1	k_2
Valeur théorique	[m ³ /mol/s]	0,316	0,0316

Deux critères de performance (le temps de synthèse et le rendement global en R) ont été définis et sont agglomérés par la technique décrite au paragraphe 1.1.5 équation (1.26). Le Tableau 3.3 donne le détail des valeurs des seuils objectifs et véto.

Tableau 3.3 : Définition des critères de performances du système n°1.

		Critère 1 [s]	Critère 2 [%]
Performance	γ_i	t_f	$-100 \times \frac{2n_R(t_f)}{n_A^{in} + n_B^{in}}$
Objectif	o_i	0	- 95
Veto	v_i	900	- 5

3.1.1.2 Système n°2 : compétitif parallèle

Ce système est constitué des deux réactions suivantes sensibles à la température :



L'objectif est donc de trouver un compromis entre accélération de la réaction par action de la température et maintien de la sélectivité au-dessus d'un seuil critique. On suppose que seuls les réactifs sont mesurés avec un écart-type qui correspond à 5% de la concentration totale initiale la plus grande ($C_{TOT}^0 = 0,1 \text{ mol/m}^3$) soit $\sigma_B = 0,005 \text{ mol/m}^3$. Pour réaliser cette synthèse, nous supposons que trois réacteurs sont disponibles, un réacteur batch, un piston et un semi-batch. Le détail des conditions accessibles se trouve dans le Tableau 3.4.

Tableau 3.4 : Gamme opératoire du système n°2.

		Batch	Piston	Semi-Batch
t	[s]	300 – 3600	10 – 200	300 – 3600
T	[°C]	5 – 95	5 – 150	5 – 95
$Ratio = [A]^0 / C_{TOT}^0$	[%]	5 – 95	5 – 95	5 – 95
Coulée B	[-]	aucune	aucune	4 paliers
Volume du réacteur	[mL]	800	15	800
Échantillonnage		t	t et $Ratio$	t

Le modèle représentant l'évolution des concentrations est paramétré par les énergies d'activation et les facteurs pré-exponentiels. Pour des questions numériques, nous les avons reparamétrés par la méthode décrite dans l'équation **Erreur ! Source du renvoi introuvable.** Les températures de référence T_{min} et T_{max} ont été prises 5°C au-dessus et en dessous du domaine de température accessible. Les valeurs des paramètres ont été fixées arbitrairement (Tableau 3.5).

Tableau 3.5 : Valeur théorique des paramètres du système n°2.

		k_1^{min}	k_1^{max}	k_2^{min}	k_2^{max}
Valeur théorique	[m ³ /mol/s]	0,001	1	0,0001	10

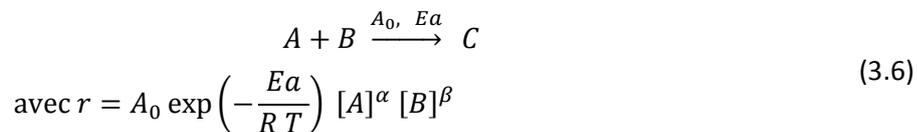
Le Tableau 3.6 donne le détail des valeurs des seuils objectifs et vétos.

Tableau 3.6 : Définition des critères de performance du système n°2.

		Critère 1 [s]	Critère 2 [%]
Performance	γ_i	t_f	$-100 \times \frac{2n_R(t_f)}{n_A^{in} + n_B^{in}}$
Objectif	o_i	0	- 95
Veto	v_i	1200	- 40

3.1.1.3 Système n° 3 : ordre de réaction fractionnaire

La particularité de cette synthèse homogène est d'avoir des ordres partiels fractionnaires.



Dans cette réaction, seuls les deux réactifs sont mesurés avec un écart-type de $\sigma_A = \sigma_B = 100 \text{ mol/m}^3$, (la concentration totale initiale étant fixée à 2000 mol/m^3).

De la même manière que dans le premier système, nous disposons de trois réacteurs dont les gammes opératoires sont données dans le Tableau 3.7.

Tableau 3.7 : Gamme opératoire du système n°3.

		Batch	Piston	Semi-Batch
t	[s]	300 – 3600	10 – 200	300 – 3600
T	[°C]	5 – 95	5 – 150	5 – 95
Ratio = $[A]^0 / C_{TOT}^0$	[%]	5 – 95	5 – 95	5 – 95
Coulée de B	[-]	aucune	aucune	4 paliers
Volume du réacteur	[mL]	800	15	800
Échantillonnage		t	t et Ratio	t

Là encore, nous effectuons la reparamétrisation de l'énergie d'activation à l'aide des vitesses de réaction aux deux températures limites $T_{min} = 0^{\circ}\text{C}$ et $T_{max} = 155^{\circ}\text{C}$. Les valeurs des paramètres par défaut sont données dans le Tableau 3.8.

Tableau 3.8 : Valeur théorique des paramètres du système n°3.

	k^{min} [m ³ /mol/s]	k^{max} [m ³ /mol/s]	α [-]	β [-]
Valeur théorique	$1,0 \times 10^{-11}$	$1,0 \times 10^{-5}$	1,5	0,5

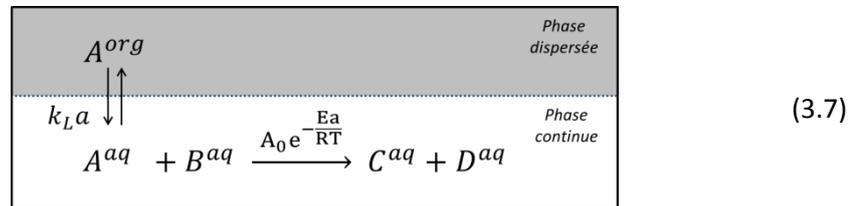
La performance optimale est définie par l'agrégation de deux critères : le temps de synthèse et le rendement global (voir Tableau 3.9).

Tableau 3.9 : Définition des critères de performances du système n°3.

		Critère 1 [s]	Critère 2 [%]
Performance	γ_i	t_f	$-100 \times \frac{2n_R(t_f)}{n_A^{in} + n_B^{in}}$
Objectif	o_i	0	-100
Veto	v_i	1200	-70

3.1.1.4 Système n°4 : liquide-liquide

Ce dernier système est très similaire à celui proposé au paragraphe 2.1.2. C'est un système liquide-liquide où le réactif A est un composé organique insoluble dans la phase aqueuse et doit réagir avec un second réactif présent dans la phase aqueuse. La réaction est donc ralentie par le transfert de matière à l'interface. Le transfert de matière est décrit par le coefficient de transfert k_L à l'interface. Ce système réactionnel a été inspiré par une réaction de saponification.



Seules quelques différences mineures ont été apportées, en particulier la reparamétrisation de l'énergie d'activation et les critères de coût (voir Tableau 3.10, Tableau 3.11 et Tableau 3-12). En effet, nous avons ajouté un critère caractérisant la puissance dissipée nécessaire à l'obtention de l'aire interfaciale désirée. Le calcul de la puissance dissipée se fait au moyen des équations présentées au paragraphe 2.1.2.1. Seules les concentrations des produits sont mesurées et leur écart-type est de $\sigma_A = \sigma_B = 6 \text{ mol/m}^3$.

Tableau 3.10 : Gamme opératoire du système n°4.

		Batch	Piston	Semi-Batch
t	[s]	300 – 3600	10 – 200	300 – 3600
T	[°C]	5 – 95	5 – 150	5 – 95
$[B]^0$	[mol/m ³]	0 – 120	0 – 120	0 – 120
a	[m ² /m ³]	200 – 3200	1000 – 10000	200 – 3200
Coulée de B	[mL/min]	<i>aucune</i>	<i>aucune</i>	4 paliers
Volume	[mL]	800	15	800
Échantillonnage		t	t et $[B]^0$	t

Tableau 3.11 : Valeur théorique des paramètres du système n°4.

	k_{min} [m ³ /mol/s]	k_{max} [m ³ /mol/s]	k_L [m/s]	m [%]
Valeur théorique	$1,0 \times 10^{-5}$	$1,0 \times 10^{-3}$	3×10^{-5}	300

Tableau 3-12 : Définition des critères de performances du système n°4.

		Critère 1 [s]	Critère 2 [W/kg]	Critère 3 [mol]
Performance	γ_i	t_f	ε	$- ([B]^0 - [C]^f)$
Objectif	o_i	0	0	-120
Veto	v_i	1200	1	-80

3.2 Répétabilité

Pour tester la répétabilité de la stratégie nous avons appliqué 10 fois de suite la stratégie à chacun des quatre systèmes les valeurs des paramètres et des seuils objectifs et veto définies au paragraphe 3.1. Nous vérifions que la stratégie converge toujours vers le même optimum en nous basant sur les 10 répétitions.

3.2.1 Détail des résultats pour le système n°1

L'enjeu de ce système est d'estimer le mieux possible les constantes de réaction afin d'atteindre un rendement maximal en un minimum de temps. L'optimum vrai se situe dans le domaine du micro-réacteur (temps de passage de 160 secondes pour un ratio en A de 50%). À cet optimum, le rendement vrai est de 80,7%.

La Figure 3.3 montre les erreurs relatives des critères estimés $\text{err}(\hat{\gamma})$ et des critères réels $\text{err}(\hat{\gamma}^*)$ par rapport au vrai. Le trait en pointillés épais représente la moyenne des erreurs observés/vrai sur les dix répétitions $\overline{\text{err}(\hat{\gamma}^*)}$. En pointillé fin, l'écart-type associé $\sigma_{\text{err}(\hat{\gamma}^*)}$. On remarque que 3 répétitions sur 10 sont assez éloignées de l'optimum vrai. Par exemple, la répétition n°2 prévoit que l'on puisse atteindre 89% de rendement en seulement 100 secondes alors que la répétition n°3 prévoit seulement 48% de rendement en 300 secondes. Pour autant, lorsque l'on compare l'optimum observé $\hat{\gamma}^*$ à l'optimum vrai γ^* , l'amplitude de l'erreur est plus faible.

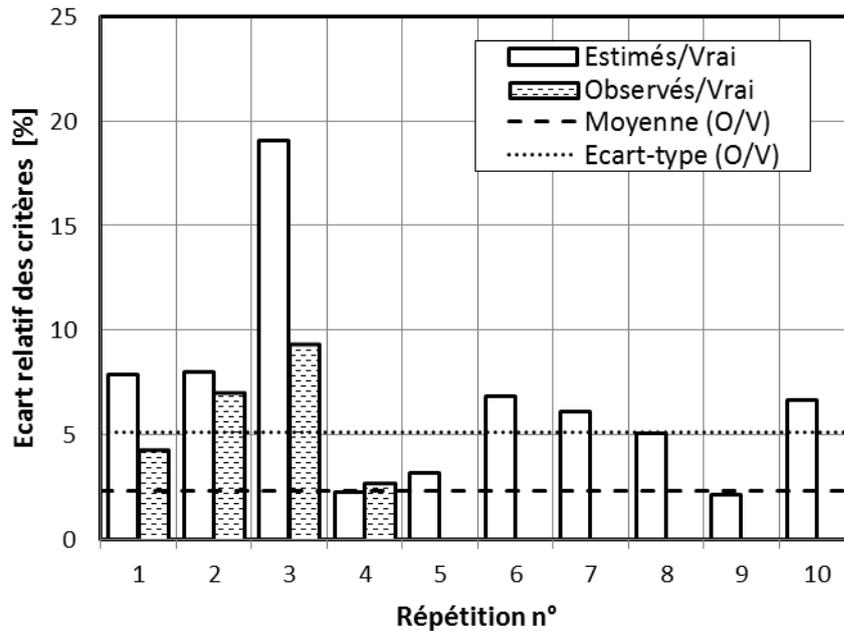


Figure 3.3 : Écart relatif entre critères estimés/observés et vrai (système n°1, test de répétabilité).

Enfin, La Figure 3.4 montre le nombre d'expériences nécessaires pour que la stratégie converge à chaque répétition. On observe qu'en moyenne il faut un peu plus de trois expériences pour converger, au minimum 2 et au maximum 5.

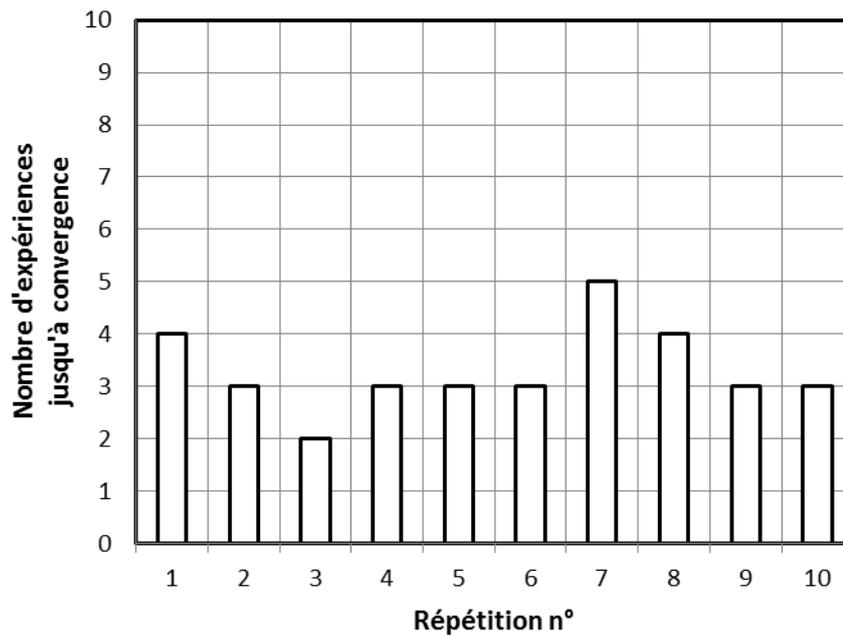


Figure 3.4 : Nombre d'expériences nécessaires jusqu'à convergence (système n°1, test de répétabilité).

On peut conclure que sur ce système, la stratégie est assez reproductible et atteint l'optimum théorique 7 fois sur 10 et s'en approche à 10% près le reste des répétitions.

3.2.2 Synthèse des systèmes n°2, 3 et 4

Dans le cas du système réactionnel consécutif-parallèle avec prise en compte de l'effet de la température, l'optimum de performance est un compromis entre temps de réaction et température. Plus on chauffe, plus le système est rapide mais plus on perd en sélectivité. La précision avec laquelle sont estimées les énergies d'activation (par l'intermédiaire de la reparamétrisation) est donc primordiale pour que la méthode converge correctement.

La Figure 3.5 montre que la stratégie est moins performante sur ce système que sur le système n°1. Cependant, sur les dix répétitions, trois ont atteint l'optimum vrai, six l'ont atteint avec environ 15% de précision et une avec plus de 40%.

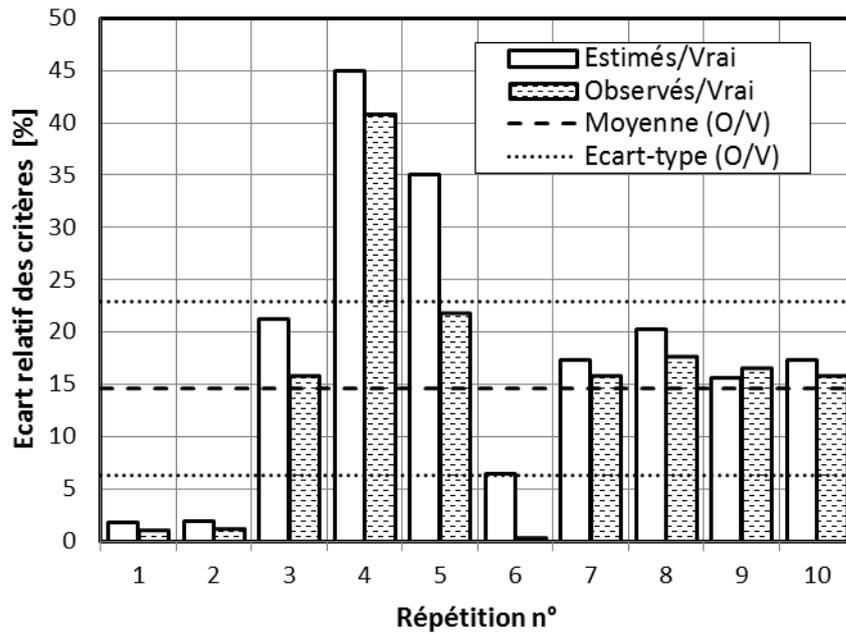


Figure 3.5 : Écart relatif entre critères estimés/observés et vrai (système n°2, test de répétabilité).

D'après la Figure 3.6, le nombre moyen d'expériences nécessaires à la convergence de la stratégie est proche de 3. Sur les dix répétitions, seule la répétition n°4 a échoué, c'est-à-dire que le critère de convergence n'a pas été atteint au bout de 10 expériences.

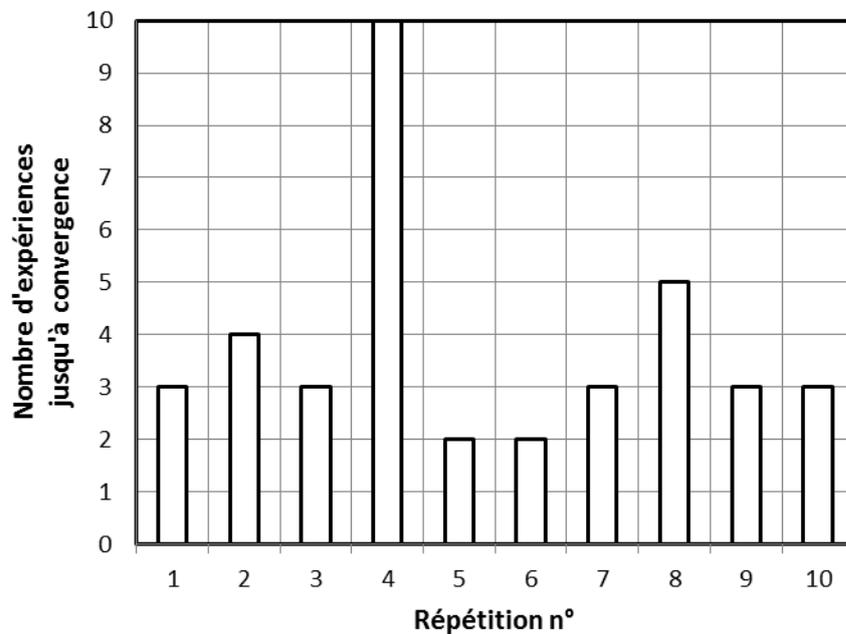


Figure 3.6 : Nombre d'expériences nécessaires jusqu'à convergence (système n°2, test de répétabilité).

Dans le système réactionnel n°3 (ordre de réaction fractionnaire), il n'y a pas de problème de sélectivité, l'optimum de performance est donc le compromis entre le

temps de synthèse et la conversion. En utilisant les conditions par défaut décrites au paragraphe 3.1.1.3, on trouve que l'optimum vrai est atteint grâce au microréacteur en 200 secondes à 150°C avec un léger excès de A, $Ratio = 0,55$. La conversion optimum vraie est de 78,5%.

Avec ce système, toutes les répétitions ont convergé avec moins de 3% d'erreur relative (voir Figure 3.7).

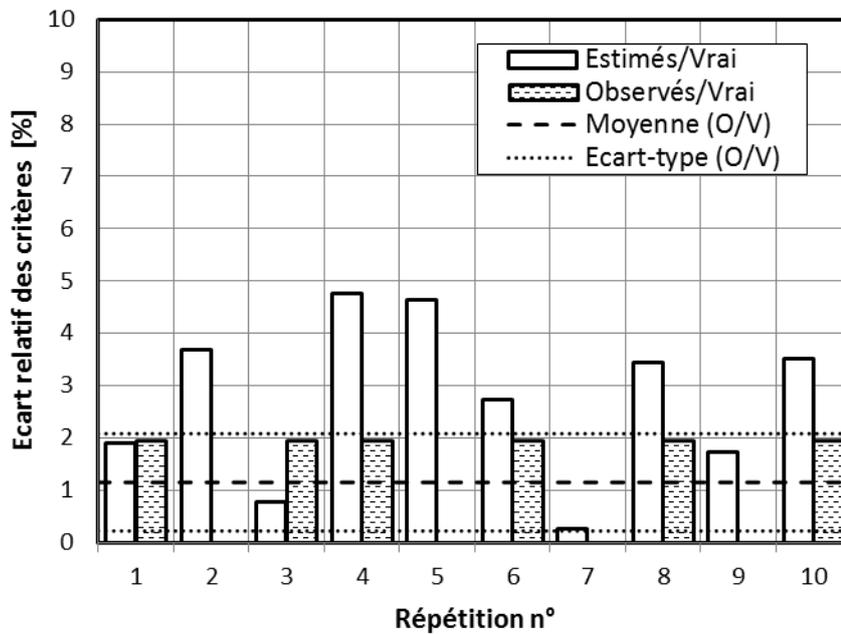


Figure 3.7 : Écart relatif entre critères estimés/observés et vrai (système n°3, test de répétabilité).

Comme le montre la Figure 3.8, en moyenne, la stratégie a convergé en un peu moins de 3 expériences.

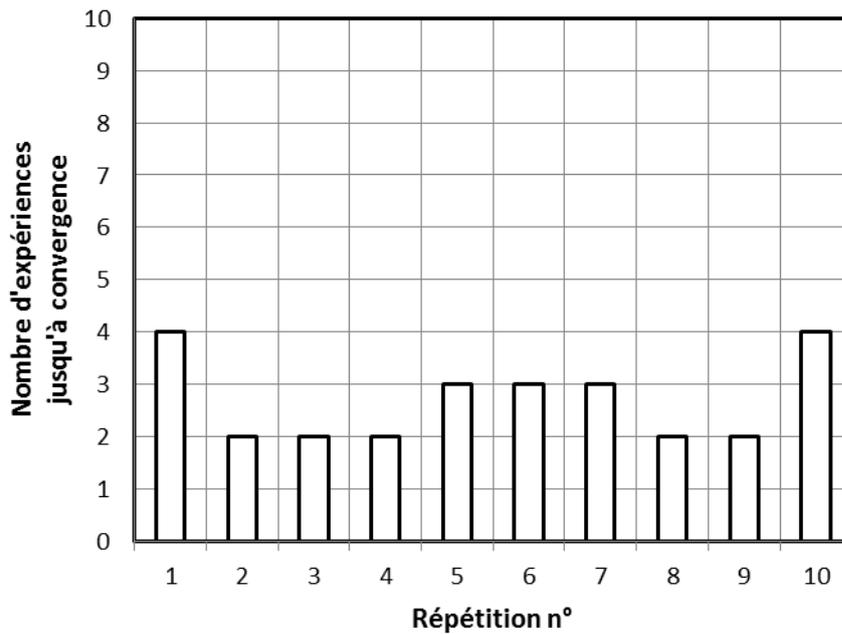


Figure 3.8 : Nombre d'expériences nécessaires jusqu'à convergence (système n°3, test de répétabilité).

Les conclusions pour le système n°4 (liquide-liquide) sont très comparables à celles du système n°3, il n'y a pas de problème de sélectivité, l'optimum est un compromis entre conversion, énergie dissipée et temps de synthèse. Toutes les répétitions convergent dans une zone proche de l'optimum vrai (moins de 10% d'erreur, voir Figure 3.9) qui se situe en microréacteur à 150°C, une aire interfaciale de 2000 m²/m³ pour un temps de synthèse de 200 secondes.

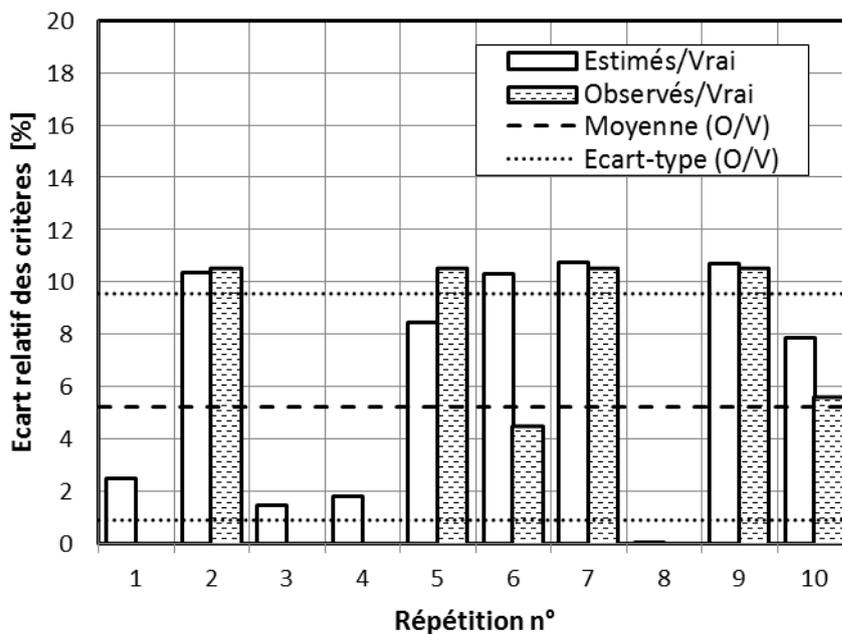


Figure 3.9 : Écart relatif entre critères estimés/observés et vrai (système n°4, test de répétabilité).

Le nombre moyen d'expériences nécessaires à la convergence de la stratégie pour ce système réactionnel est très légèrement supérieur à 2 comme l'atteste la Figure 3.10.

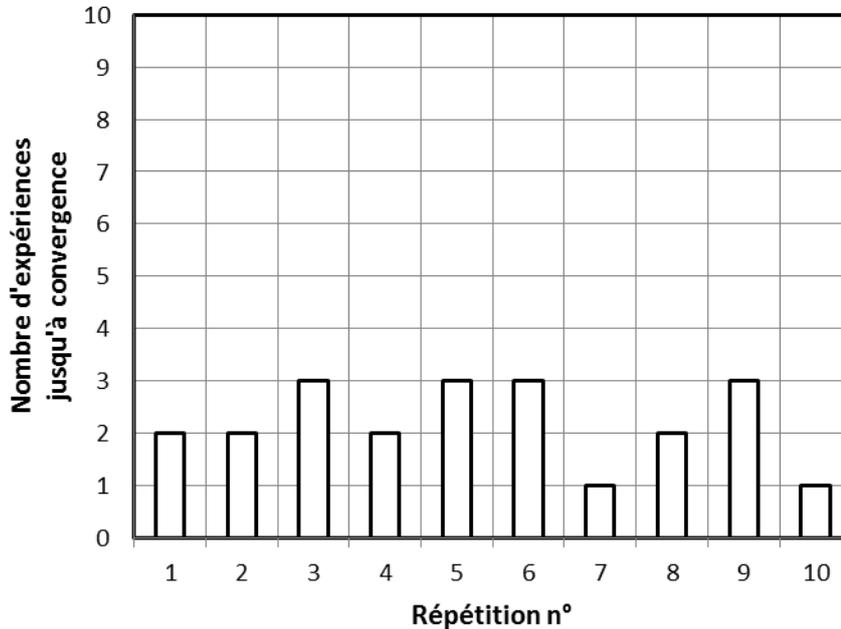


Figure 3.10 : Nombre d'expériences nécessaires jusqu'à convergence (système n°4, test de répétabilité).

3.2.3 Conclusion sur la répétabilité

Ce paragraphe a permis de mettre en lumière le fait que la stratégie converge avec une précision dépendante du système réactionnel étudié (voir récapitulatif Tableau 3.13). Le nombre moyen d'expériences requises pour converger est relativement faible quel que soit le système (entre 2,2 et 3,8 expériences de validation comprises).

Toutefois, il paraît difficile d'expliquer, à ce niveau, les différences obtenues pour chacun de ces systèmes. Pour cela, il faudrait faire une analyse détaillée de la sensibilité des grandeurs observées et des critères de performances de chaque système par rapport aux paramètres et aux conditions expérimentales. Malgré cela, étant donné le nombre de dimension qu'il faudrait analyser, une interprétation des résultats resterait tout de même difficile.

Par ailleurs, dans la colonne « Observés/Vrai » du Tableau 3.13, nous donnons l'écart-type de la précision de la convergence observée de la stratégie (noté $\sigma_{err(\hat{Y}^*)}$). Cette grandeur sera utilisée dans le paragraphe suivant 3.3.

Tableau 3.13 : Moyenne pour chaque système de l'écart relatif entre critères estimés/observés et vrai et du nombre d'expériences nécessaires à la convergence de la stratégie.

	Estimés/Vrai $\overline{\text{err}(\hat{Y})}$	Observés/Vrai $\overline{\text{err}(\hat{Y}^*)} \pm \sigma_{\text{err}(\hat{Y}^*)}$	Nombre d'expériences
Système n°1	6,7	2,3 ± 2,8	3,3
Système n°2	14,5	14,6 ± 8,3	3,8
Système n°3	2,8	1,2 ± 0,9	2,7
Système n°4	2,7	5,23 ± 4,3	2,2

Ainsi, pour les tests de robustesse de la stratégie, il faudra tenir compte du « bruit de non-répétabilité » de la stratégie, savoir si le nombre de grandeurs mesurées, la vitesse des phénomènes et les erreurs de mesures ont une influence sur la précision avec laquelle converge la stratégie.

3.3 Test de robustesse de la stratégie

Nous avons testé la robustesse de la stratégie en examinant la convergence suivant différentes conditions telles que :

- l'influence des grandeurs mesurées : comment réagit la stratégie si l'on mesure une seule ou plusieurs espèces chimiques ?
- l'influence de l'erreur expérimentale : quel est l'impact de l'erreur expérimentale sur le comportement de la stratégie ?
- l'influence des paramètres : comment se comporte la stratégie lorsque les paramètres vrais sont plus ou moins grands ?

On considère que la stratégie échoue lorsque l'écart relatif entre les critères observés et les critères vrais est significativement supérieur au « bruit de non-répétabilité » détaillé au paragraphe 3.2.3. Ce qui donne en équation :

$$\text{err}(\hat{Y}^*) \geq \overline{\text{err}(\hat{Y}^*)} + \sigma \quad (3.8)$$

Afin, de minimiser le plus possible l'impact du « bruit de non-répétabilité » de la stratégie, tout en essayant d'avoir un temps de calcul raisonnable, nous avons répété trois fois la stratégie dans chacune des conditions de convergence présentées ci-dessous où chaque mesure de l'écart entre critères estimés/observés et le critère vrai est donc la moyenne de trois répétitions.

3.3.1 Influence des grandeurs mesurées

Dans ce paragraphe, nous nous intéressons à l'influence du nombre et au type d'espèces chimiques mesurées sur la vitesse de convergence de la stratégie proposée.

Dans le cas du système n°1 (consécutif-parallèle simplifié), la Figure 3.11 montre comment évolue l'écart relatif des critères estimés/observés en fonction du nombre d'espèces mesurée. Il apparaît clairement que, si seul le composé A est mesuré, la stratégie échoue (l'écart entre l'optimum trouvé par la stratégie et le vrai optimum est significativement grand au regard du test de répétabilité du paragraphe 3.2). Ceci est facilement compréhensible car l'estimation de la concentration en A est structurellement très peu sensible à la valeur de la constante de la seconde réaction, or c'est le ratio entre k_1 et k_2 qui détermine le rendement maximal de la synthèse. Dans tous les autres cas, la stratégie converge au moins aussi bien. Pour rappel, lors du test de répétabilité (§3.2.1), seul B était mesuré.

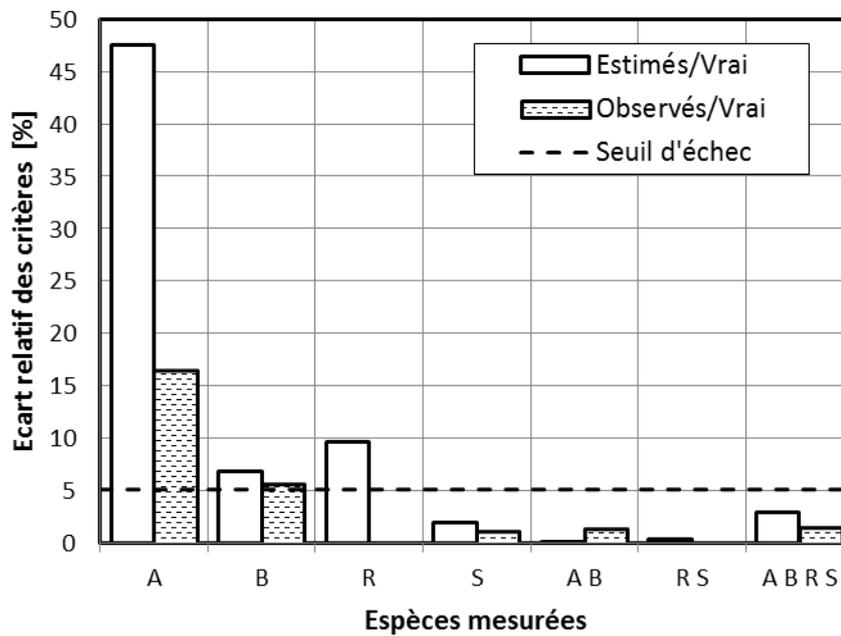


Figure 3.11 : Écart relatif entre critères estimés/observés et vrai (système n°1, test de robustesse, espèces mesurées).

On peut faire la même analyse pour le système n°2 qui structurellement est assez similaire (voir Figure 3.12) : le cas où seul A est mesuré ne converge pas l'optimum vrai.

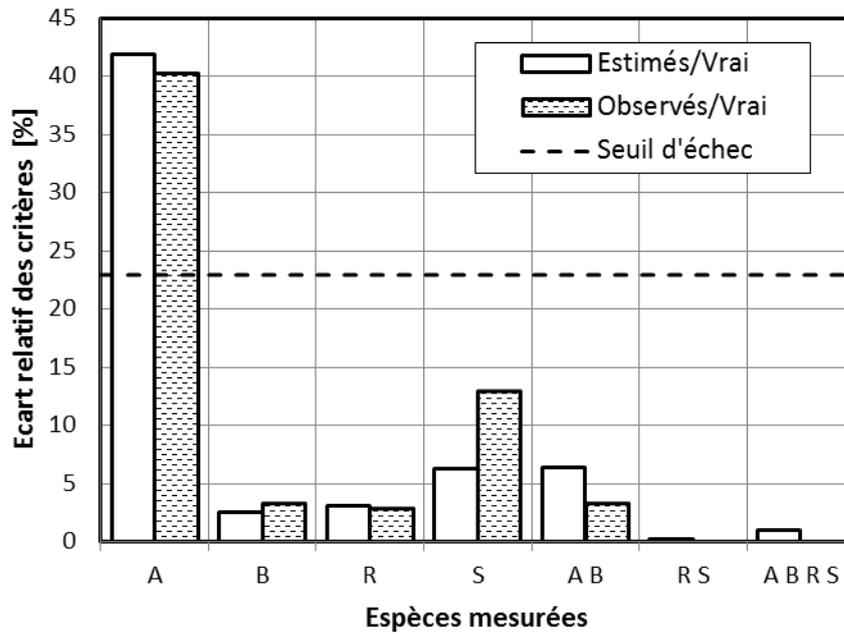


Figure 3.12 : Écart relatif entre critères estimés/observés et vrai (système n°2, test de robustesse, espèces mesurées).

Pour le système n°3 (ordre fractionnaire), quelles que soient les espèces mesurées, la stratégie converge vers l'optimum à moins de 2% près, comme le montre la Figure 3.13.

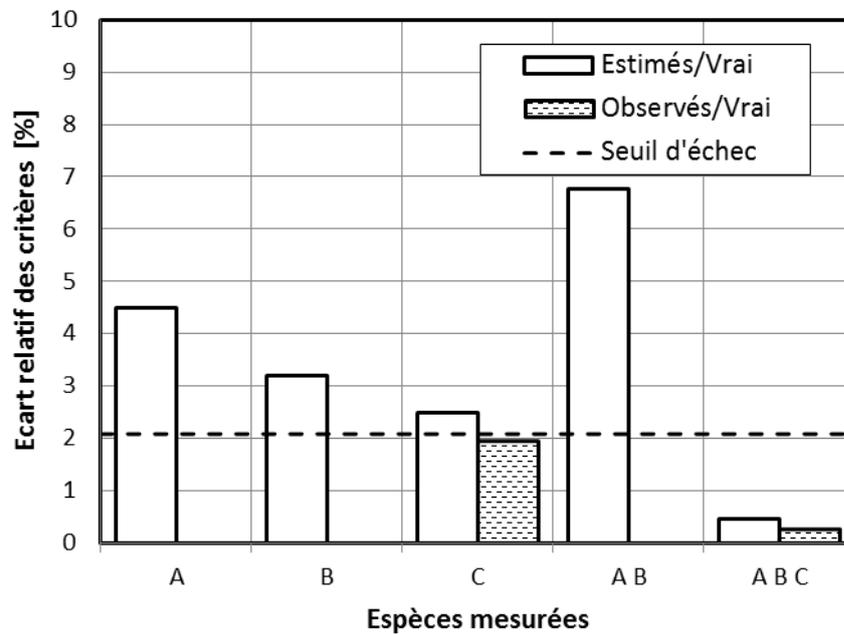


Figure 3.13 : Écart relatif entre critères estimés/observés et vrai (système n°3, test de robustesse, espèces mesurées).

Enfin, la Figure 3.14, montre que pour le système n°4 (liquide-liquide) il n'y a pas non plus de cas où la stratégie fonctionne moins bien que d'autres. Comme l'on pouvait s'y attendre, lorsque toutes les espèces sont mesurées, la stratégie converge avec une

excellente précision. À noter que les espèces C et D jouant le même rôle, nous n'avons pas envisagé le cas où seul D était mesuré.

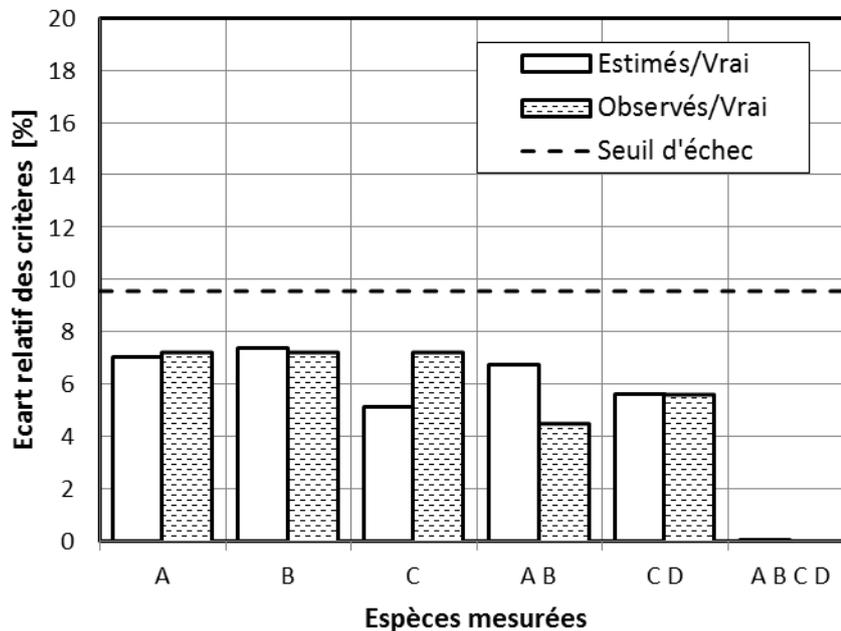


Figure 3.14 : Écart relatif entre critères estimés/observés et vrai (système n°4, test de robustesse, espèces mesurées).

3.3.2 Influence de la valeur des paramètres

Dans ce paragraphe, nous avons voulu évaluer comment la stratégie se comporte si les paramètres changent d'ordre de grandeur. Physiquement, cela revient à accélérer ou à ralentir les phénomènes. Ainsi, si le phénomène est trop rapide ou trop lent, les conditions expérimentales riches en information risquent d'être en dehors de la fenêtre opératoire des équipements.

Pour étudier cela, nous avons fait varier les paramètres des quatre systèmes de manière à faire varier le temps caractéristique de chaque phénomène d'un facteur 0,1 et d'un facteur 10. Chaque système est paramétré par 4 paramètres (excepté le n°1 où les deux paramètres pouvaient varier d'un facteur 0,01 à un facteur 100), nous avons donc fait un plan factoriel complet pour tester toutes les combinaisons possibles. À chaque jeu de paramètres, correspond un optimum vrai qui lui est propre.

3.3.2.1 Système n°1

Avec le système n°1, consécutif-parallèle simplifié, nous avons observé qu'une seule situation où la stratégie ne converge pas (voir Figure 3.15 et Tableau 3.14). C'est-à-dire, qu'une seule valeur est au-dessus du seuil d'échec défini au paragraphe 3.3 (le seuil d'échec du système n°1 est égal à 5,2 %). La raison de cet échec peut s'expliquer dans ce cas par le fait que l'optimum de performance vrai se situe dans une zone qui n'est couverte ni par le domaine du batch, ni par le domaine du piston (voir Figure 3.16).

L'estimation de l'optimum est donc instable, car une petite variation dans l'estimation des paramètres fait que la stratégie choisie est, soit un optimum à la frontière haute du piston, soit à la frontière basse du batch. Si la zone de temps comprise entre 200 et 300 secondes avait été accessible, même si la prédiction est erronée, la stratégie aurait trouvé des conditions optimales ($\hat{t} = 250$ s) très proches des conditions optimales vraies ($t^* = 240$ s).

Tableau 3.14 : Valeurs testées des paramètres (système n°1, test de robustesse, paramètres).

N°	k_1 [m ³ /mol/s]	k_2 [m ³ /mol/s]
1	0,01	0,001
2	0,1	0,001
3	1	0,001
4	10	0,001
5	0,01	0,01
6	0,1	0,01
7	1	0,01
8	10	0,01
9	0,01	0,1
10	0,1	0,1
11	1	0,1
12	10	0,1
13	0,01	1
14	0,1	1
15	1	1
16	10	1

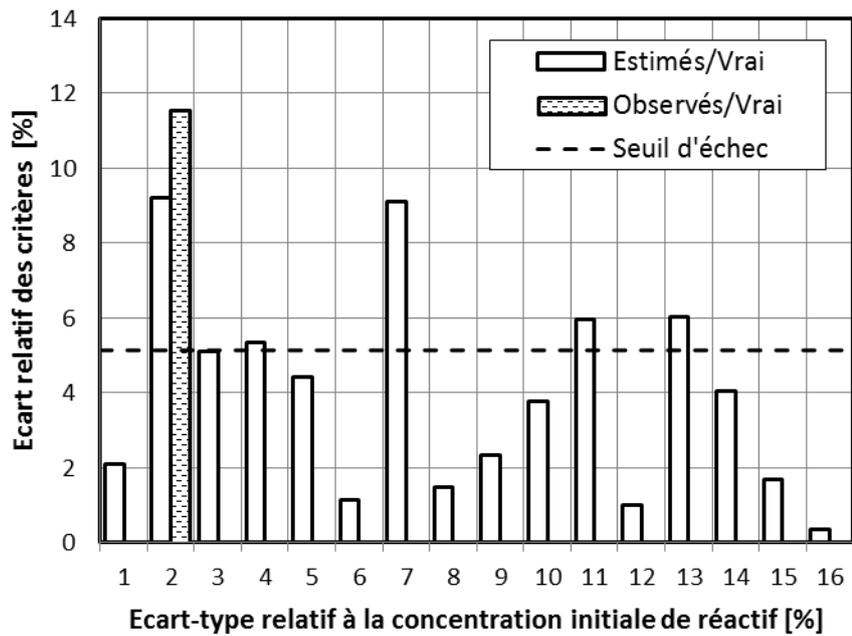


Figure 3.15 : Écart relatif entre critères estimés/observés et vrai (système n°1, test de robustesse, paramètres).

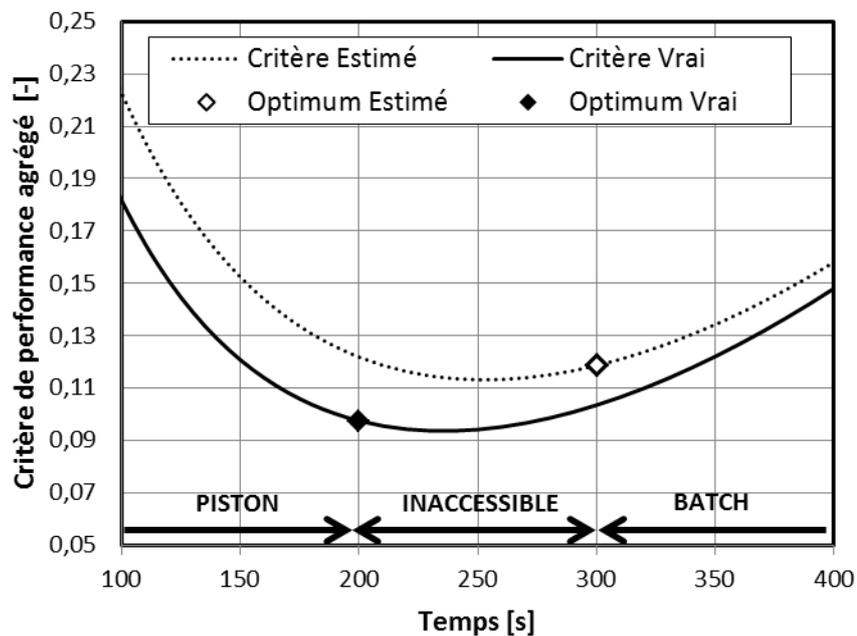


Figure 3.16 : Évolution du critère de performance agrégé (estimé et vrai) en fonction du temps (système n°1, test de robustesse, paramètres, test N°2).

3.3.2.2 Système n°2

Les résultats obtenus pour le système n°2 (consécutif-parallèle) sont détaillés dans la Figure 3.17. Les valeurs des paramètres testés sont données dans le Tableau 3.15. Le seuil d'échec vaut pour ce système 22,9 %. On observe que la stratégie con-

3.3. TEST DE ROBUSTESSE DE LA STRATÉGIE

verge quelles que soient les valeurs des paramètres à moins de 6% près, sauf pour le jeu n°6 où les constantes à basse température sont minimales et maximales à haute température, ce qui correspond à des énergies d'activation élevées (70 kJ/mol pour la 1^{ière} réaction et 101 kJ/mol pour la seconde), et donc au fait que le rendement est très sensible à la température.

Tableau 3.15 : Valeurs testées des paramètres (système n°2, test de robustesse, paramètres).

N°	k_1^{min} [m ³ /mol/s]	k_1^{max} [m ³ /mol/s]	k_2^{min} [m ³ /mol/s]	k_2^{max} [m ³ /mol/s]
1	0,0001	0,1	0,00001	1
2	0,0001	0,1	0,00001	100
3	0,0001	0,1	0,001	1
4	0,0001	0,1	0,001	100
5	0,0001	10	0,00001	1
6	0,0001	10	0,00001	100
7	0,0001	10	0,001	1
8	0,0001	10	0,001	100
9	0,01	0,1	0,00001	1
10	0,01	0,1	0,00001	100
11	0,01	0,1	0,001	1
12	0,01	0,1	0,001	100
13	0,01	10	0,00001	1
14	0,01	10	0,00001	100
15	0,01	10	0,001	1
16	0,01	10	0,001	100

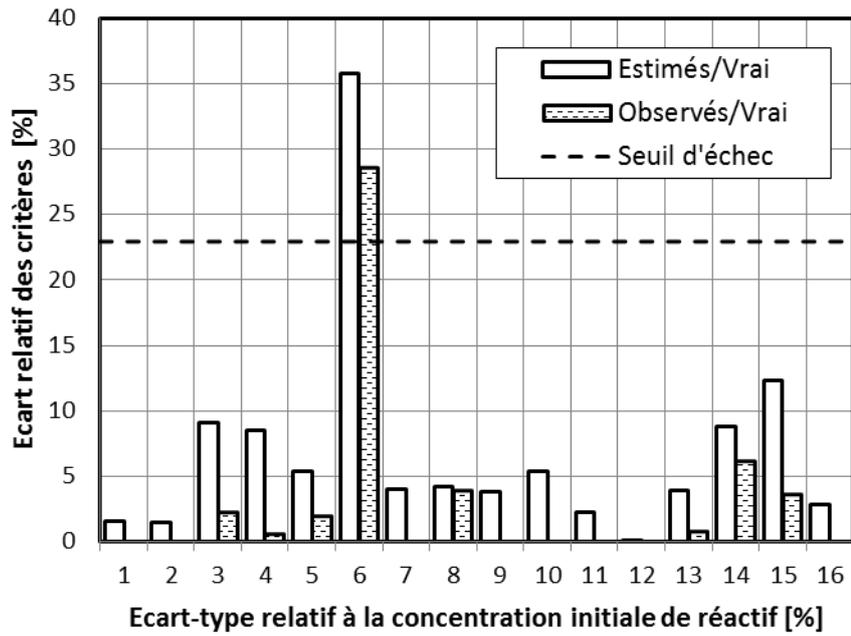


Figure 3.17 Écart relatif entre critères estimés/observés et vrai (système n°2, test de robustesse, paramètres).

3.3.2.3 Système n°3

Dans le cas du système n°3 avec les ordres fractionnaires, les ordres α et β ont été choisis de telle sorte que l'ordre global soit constant et égal à 2 (voir Tableau 3.16),. Nous n'avons pas observé de cas où la stratégie ne converge pas à moins de 2% près soit une valeur inférieure au « seuil d'échec » obtenue dans la phase de répétabilité (le seuil d'échec du système n°3 est égal à 2,8 %).

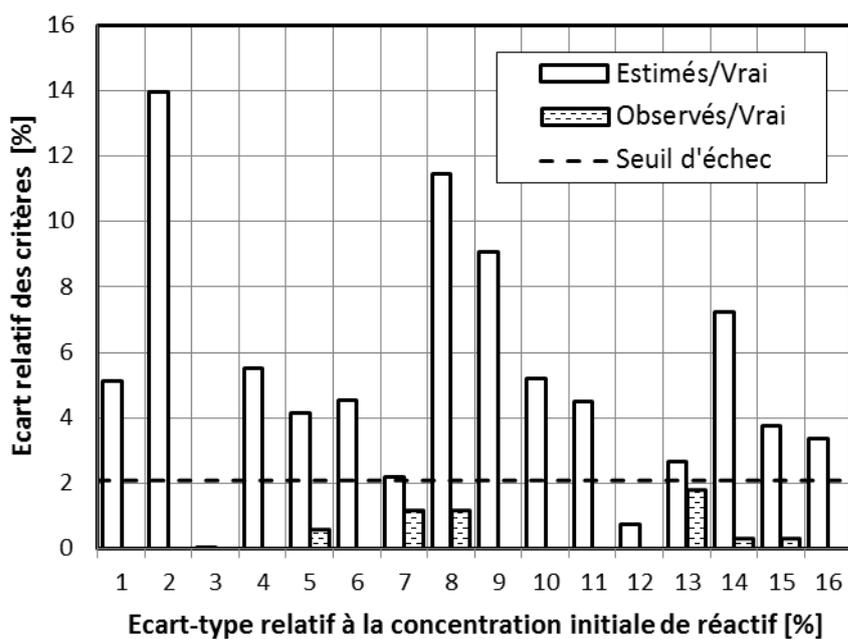


Figure 3.18 : Écart relatif entre critères estimés/observés et vrai (système n°3, test de robustesse, paramètres).

Tableau 3.16 : Valeurs testées des paramètres (système n°3, test de robustesse, paramètres).

N°	k^{min} [m ³ /mol/s]	k^{max} [m ³ /mol/s]	α [-]	β [-]
1	1×10^{-12}	1×10^{-6}	1	1
2	1×10^{-12}	1×10^{-6}	1,33	0,67
3	1×10^{-12}	1×10^{-6}	1,5	0,5
4	1×10^{-12}	1×10^{-6}	1,6	0,4
5	1×10^{-12}	1×10^{-4}	1	1
6	1×10^{-12}	1×10^{-4}	1,33	0,67
7	1×10^{-12}	1×10^{-4}	1,5	0,5
8	1×10^{-12}	1×10^{-4}	1,6	0,4
9	1×10^{-10}	1×10^{-6}	1	1
10	1×10^{-10}	1×10^{-6}	1,33	0,67
11	1×10^{-10}	1×10^{-6}	1,5	0,5
12	1×10^{-10}	1×10^{-6}	1,6	0,4
13	1×10^{-10}	1×10^{-4}	1	1
14	1×10^{-10}	1×10^{-4}	1,33	0,67
15	1×10^{-10}	1×10^{-4}	1,5	0,5
16	1×10^{-10}	1×10^{-4}	1,6	0,4

3.3.2.4 Système n°4

Pour le système n°4, d'après la Figure 3.19, l'erreur entre les critères estimés et le critère vrai est inférieure ou égale au seuil d'échec observée lors du test de répétabilité (égal à 9,6%), pour tous les jeux de paramètres (voir Tableau 3.17).

Tableau 3.17 : Valeurs testées des paramètres (système n°4, test de robustesse, paramètres).

N°	k^{min} [m ³ /mol/s]	k^{max} [m ³ /mol/s]	k_L [m/s]	m [-]
1	1×10^{-6}	0,0001	1×10^{-5}	100
2	1×10^{-6}	0,0001	1×10^{-5}	1000
3	1×10^{-6}	0,0001	1×10^{-4}	100
4	1×10^{-6}	0,0001	1×10^{-4}	1000
5	1×10^{-6}	0,01	1×10^{-5}	100
6	1×10^{-6}	0,01	1×10^{-5}	1000
7	1×10^{-6}	0,01	1×10^{-4}	100
8	1×10^{-6}	0,01	1×10^{-4}	1000
9	1×10^{-4}	0,0001	1×10^{-5}	100
10	1×10^{-4}	0,0001	1×10^{-5}	1000
11	1×10^{-4}	0,0001	1×10^{-4}	100
12	1×10^{-4}	0,0001	1×10^{-4}	1000
13	1×10^{-4}	0,01	1×10^{-5}	100
14	1×10^{-4}	0,01	1×10^{-5}	1000
15	1×10^{-4}	0,01	1×10^{-4}	100
16	1×10^{-4}	0,01	1×10^{-4}	1000

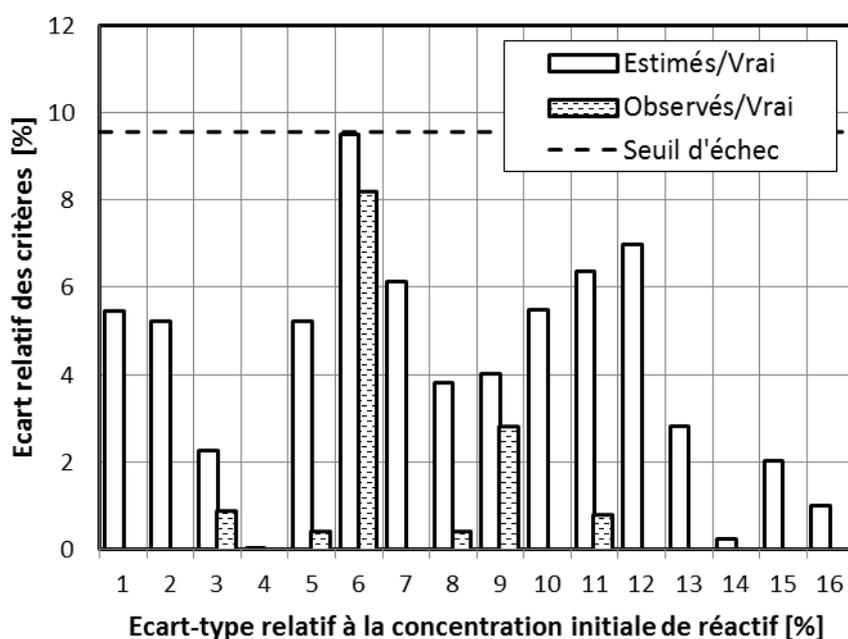


Figure 3.19 : Écart relatif entre critères estimés/observés et vrai (système n°4, test de robustesse, paramètres).

3.3.3 Influence de l'erreur de mesure

Pour chacun des systèmes réactionnels, nous avons fait varier l'erreur de mesure afin de voir quelle était l'erreur admissible pour que la stratégie continue de converger. Ainsi nous avons fait varier l'écart-type de l'erreur de mesure de 0,1% à 50% de la concentration initiale en réactif. Sans surprise, dans tous les systèmes, plus l'erreur de mesure est grande, plus l'écart relatif entre les critères estimés (respectivement réels) augmente. À part pour le système n°1, pour les valeurs de 20% et 50% d'erreur de mesure, la stratégie ne trouve l'optimum que de façon approximative comme le montrent les figures 3.20 à 3.23 :

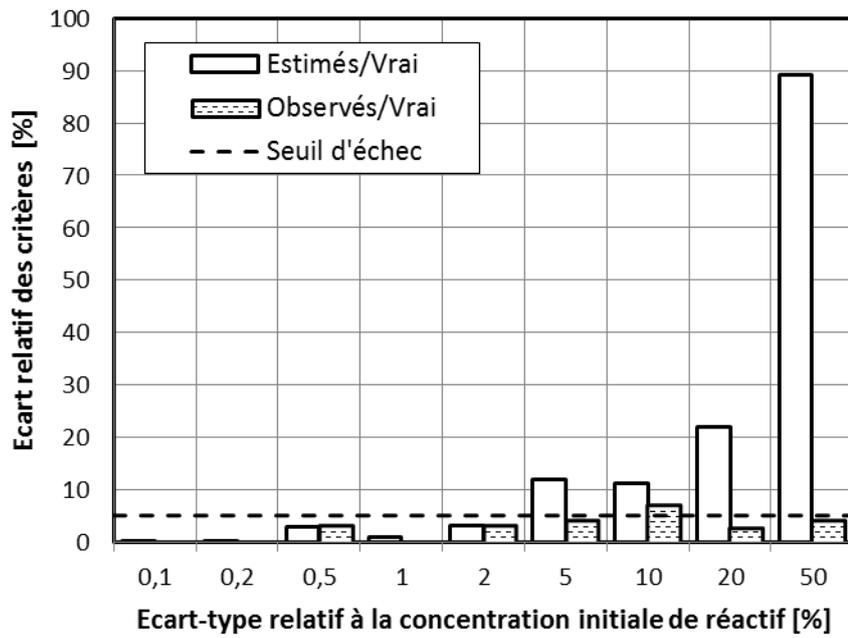


Figure 3.20: Écart relatif entre critères réels et théoriques (système n°1, test de robustesse, erreur de mesure).

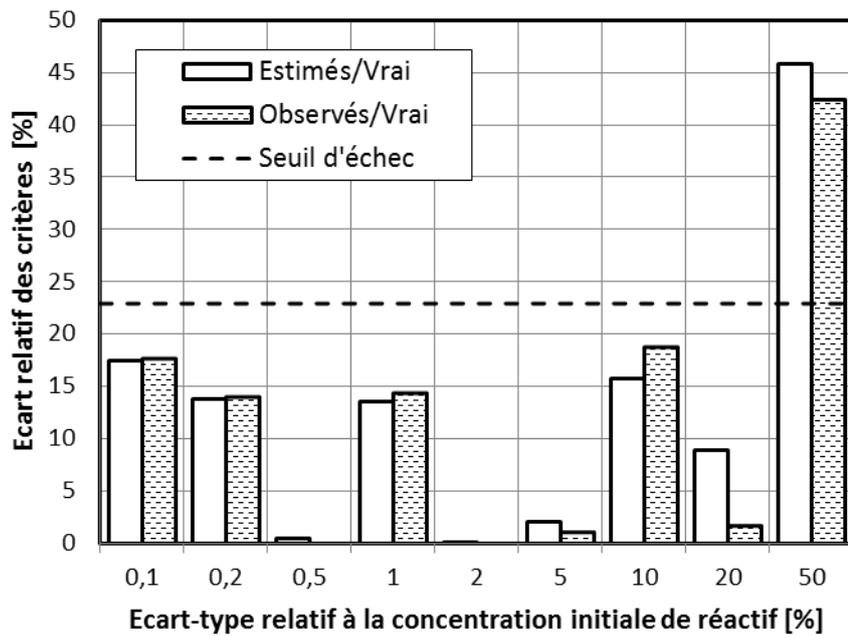


Figure 3.21 : Écart relatif entre critères réels et théoriques (système n°2, test de robustesse, erreur de mesure).

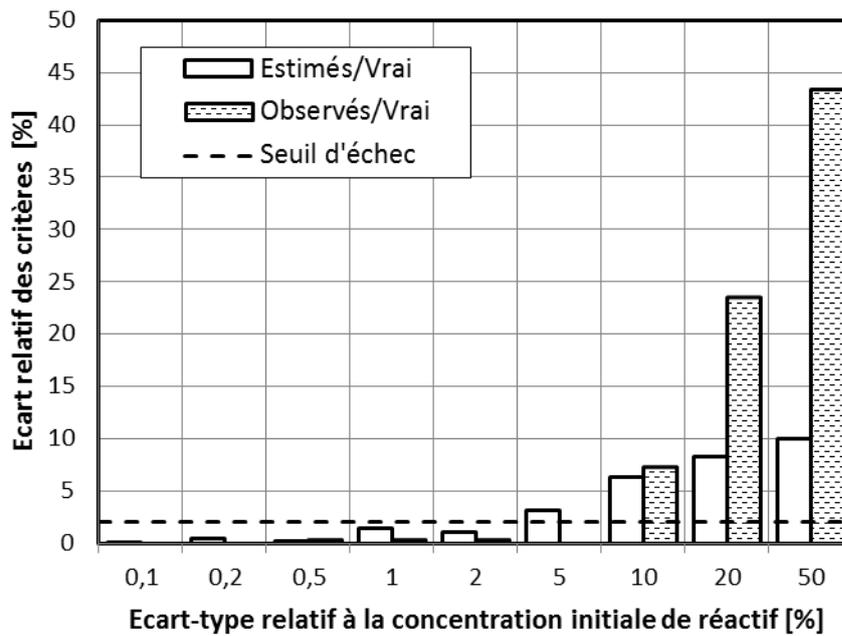


Figure 3.22 : Écart relatif entre critères réels et théoriques (système n°3, test de robustesse, erreur de mesure).

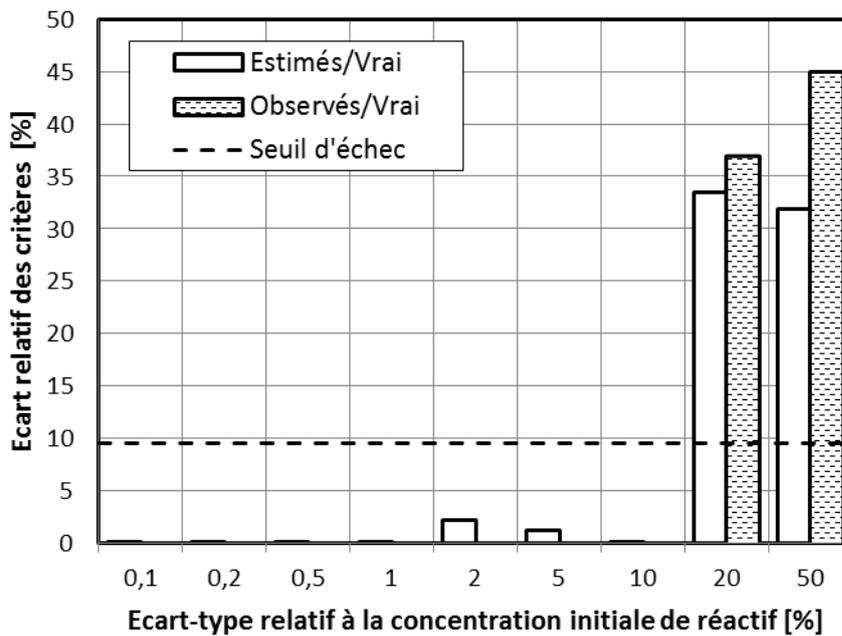


Figure 3.23: Écart relatif entre critères réels et théoriques (système n°4, test de robustesse, erreur de mesure).

3.3.4 Synthèse sur le test de robustesse

Nous avons observé que, dans la plupart des cas de ce test de robustesse, la stratégie converge au moins aussi bien vers l'optimum réel que dans la phase de répétabilité (§ 3.2). Le Tableau 3.18 synthétise les cas où la stratégie échoue à trouver l'optimum

vrai. Nous n'avons pas testé l'influence de la taille du domaine expérimental de chaque équipement, car cela revient, à peu de choses près, à faire varier les paramètres et donc le temps caractéristique des phénomènes qu'ils représentent.

Tableau 3.18 : Synthèse de la robustesse de la stratégie vis-à-vis des espèces mesurées.

Systeme n°	Schéma réactionnel	Échec grandeur mesurée	Échec valeurs des paramètres	Échecs erreur de mesure
1	$A + B \xrightarrow{k_1} R$ $R + B \xrightarrow{k_2} S$	A	N°2 : Optimum vrai inaccessible	—
2	$A + B \xrightarrow{A_1, Ea_1} R$ $R + B \xrightarrow{A_2, Ea_2} S$	A	N°6 : Sensibilité aux paramètres très élevées	50%
3	$A + B \xrightarrow{A_0, Ea, \alpha, \beta} C$	—	—	10, 20 et 50%
4	$A^{org} \xrightleftharpoons[k_L, m]{} A^{aq}$ $A^{aq} + B \xrightarrow{A_0, Ea} C + D$	—	—	20 et 50 %

3.4 L'avantage de la prise en compte de plusieurs technologies

Dans cette section, nous avons appliqué la stratégie en restreignant le domaine expérimental à chacune des trois technologies (batch, piston et semi-batch) lors des phases d'acquisition de données. Il faut noter que l'optimum de performance n'est pas forcément dans le domaine de l'équipement utilisé pour l'acquisition de données ce qui implique de devoir faire une extrapolation du modèle. Puis, nous avons comparé les résultats obtenus à la stratégie classique qui consiste à laisser libre le choix du réacteur pour l'acquisition de données. Cette étude a été menée sur les quatre systèmes réactionnels. Afin de minimiser l'influence de l'imparfaite répétabilité de la convergence de la stratégie, nous avons lancé la stratégie cinq fois. Les résultats présentés ci-dessous sont la moyenne des cinq répétitions.

Comme l'on pouvait s'y attendre, l'utilisation de toutes les technologies disponibles permet d'atteindre l'optimum de performance théorique plus rapidement qu'en utilisant seulement une technologie (voir Figure 3.24).

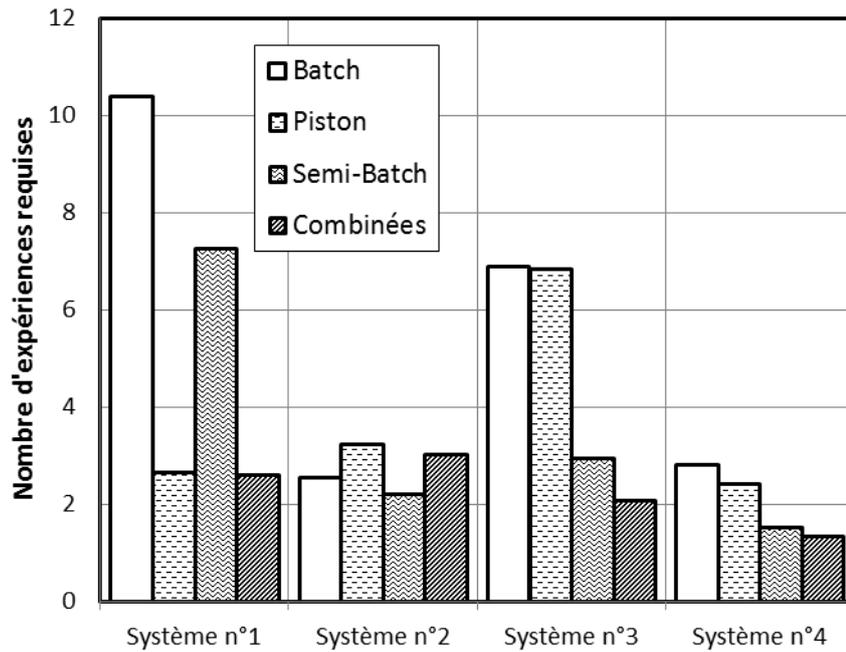


Figure 3.24 : Nombre d'expériences nécessaires jusqu'à convergence (pour chaque système, influence de la technologie).

Par ailleurs, lorsque l'on compare la précision avec laquelle la stratégie converge vers l'optimum vrai, on s'aperçoit que le fait d'utiliser plusieurs technologies pour l'acquisition de données permet, non seulement de minimiser le nombre d'expériences, mais aussi d'améliorer la précision des prédictions.

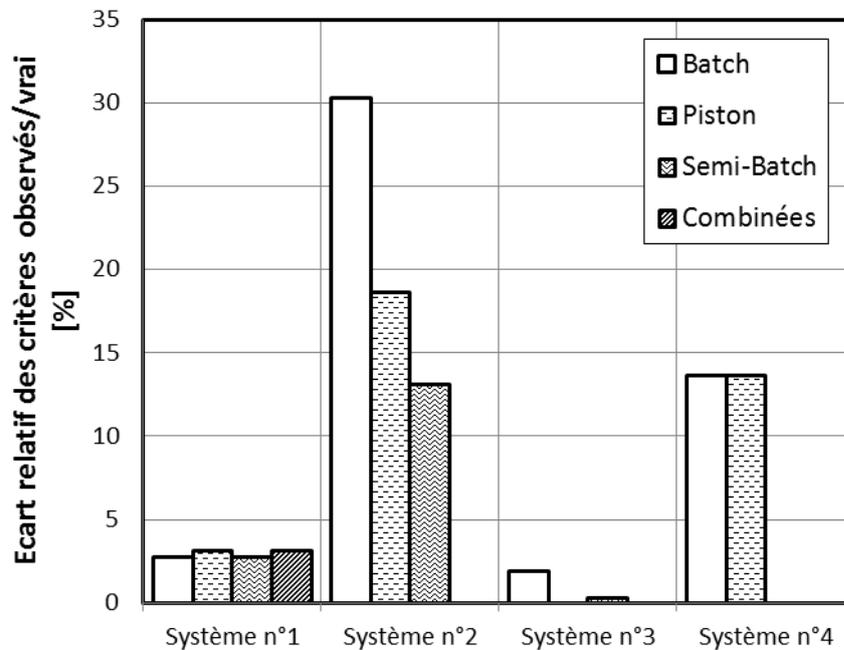


Figure 3.25 : Écart relatif entre critères réels et théoriques (pour chaque système, influence de la technologie).

On note aussi que le semi-batch est à chaque fois plus performant que le batch. Cela met bien en lumière le fait que la qualité de l'acquisition des données cinétiques est plus importante dans les phases transitoires que dans les phases stationnaires, et couler un ou plusieurs réactifs permet de maintenir plus longtemps le régime transitoire.

De la même manière, le piston est systématiquement meilleur que le batch. Ceci montre que la flexibilité d'échantillonnage couplée au fait que son domaine opératoire permet d'atteindre des zones différentes peut se révéler très intéressant pour l'acquisition de données.

3.5 Comparaison à d'autres stratégies

Nous avons comparé notre stratégie à des approches classiques (plan d'expériences et Algorithme de Nancy) sur chacun des quatre systèmes réactionnels utilisés dans le paragraphe 3.1. Dans les paragraphes 3.5.1 et 3.5.2 nous détaillons la mise en œuvre des deux approches respectives. Les résultats sont regroupés dans le paragraphe 3.5.3. Les valeurs des paramètres réels, de l'erreur expérimentale, des domaines opératoires et des espèces mesurées sont les valeurs définies par défaut au paragraphe §3.1 (les mêmes que celles utilisées pour tester la répétabilité de la stratégie). Afin de réduire l'influence du caractère aléatoire des expériences, les trois stratégies expérimentales (approche tendancielle par les plans d'expériences, Algorithme de Nancy et stratégie développée) sont lancées dix fois dans les mêmes conditions. Les résultats présentés dans ce paragraphe sont moyennés sur ces dix répétitions.

De la même manière que précédemment, nous évaluons l'efficacité avec laquelle l'approche tendancielle trouve l'optimum de performance sur deux points :

- La comparaison des critères estimés et vrai.
- Le nombre d'expériences nécessaires à la convergence de la stratégie.

3.5.1 Approche tendancielle

Pour les raisons détaillées au §2.1, nous devons écrire un modèle pour chaque technologie. Parmi les plus importantes, nous pouvons rappeler que l'approche par plan d'expérience ne s'adapte pas très bien lorsque le domaine opératoire disponible n'est pas un hypercube (ce qui est le cas lorsque plusieurs technologies sont utilisées). Les modèles linéaires utilisés comportent les termes suivants :

- Un terme constant,
- Un terme linéaire pour chaque facteur,
- Un terme quadratique pour les facteurs autres que les profils de coulée,
- Un terme d'interaction pour chaque couple de facteurs autres que les profils de coulée.

Ainsi, pour le système n°1 en semi-batch, nous avons :

- Un terme constant,
- Dix termes linéaires (temps, ratio molaire et 8 pour le profil de coulée de 4 paliers),
- Deux termes quadratiques (temps et ratio molaire),
- Un terme d'interaction entre le temps et le ratio molaire.

$$\begin{aligned}
 f(\xi, \theta) = & \theta_0 + \theta_1 t + \theta_2 R_A \\
 & + \theta_3 Q_1 + \theta_4 Q_2 + \theta_5 Q_3 + \theta_6 Q_4 \\
 & + \theta_7 \Delta t_1 + \theta_8 \Delta t_2 + \theta_9 \Delta t_3 + \theta_{10} \Delta t_4 \\
 & + \theta_{11} t^2 + \theta_{12} R_A^2 \\
 & + \theta_{13} t R_A
 \end{aligned} \tag{3.9}$$

Soit un total de 14 termes. Le nombre de paramètres du modèle linéaire est égal au nombre de termes. Afin d'avoir un nombre minimum d'expériences à faire, le nombre de points de prélèvement est égal au nombre de termes plus un. Pour déterminer les expériences à faire nous utilisons l'algorithme de Fedorov (Fedorov 1972) qui permet de générer un plan d'expériences D-optimal. Pour le réacteur batch du système n°1 (un terme constant, deux linéaires, deux quadratiques et un d'interaction, soit six au total) nous obtenons le plan d'expériences suivant :

Tableau 3.19 : Plan d'expérience D-optimal pour un modèle linéaire dans le cas du réacteur batch (système n°1).

N° échantillon	t [s]	Ratio = $[A]^0 / C_{Tor}^0$ [%]	N° expérience
1	300	5	1
2	1950	5	1
3	3600	5	1
4	1950	5	2
5	3600	5	2
6	300	95	3
7	3600	95	3

Lorsque cela est possible, on génère plusieurs points de prélèvements en une seule expérience (en limitant le nombre d'échantillons à huit par expérience). Dans le cas du Tableau 3.20, seules trois expériences sont nécessaires à la génération des 7 points de prélèvement (dernière colonne du Tableau 3.19).

Tableau 3.20 : Nombre d'échantillon / d'expériences requis pour l'approche tendancielle.

	Système n°1	Système n°2	Système n°3	Système n°4
<i>Batch</i>	7 / 3	11 / 7	11 / 7	16 / 15
<i>Piston</i>	7 / 1	11 / 2	11 / 2	16 / 2
<i>Semi-Batch</i>	15 / 15	19 / 19	19 / 19	24 / 24
<i>Total</i>	29 / 19	41 / 28	41 / 28	56 / 41

3.5.2 Algorithme de Nancy

Enfin, nous avons comparé notre stratégie avec l'algorithme de Nancy (paragraphe 1.4.2). Ce dernier est sensiblement basé sur le même principe que la stratégie développée à ceci près que la planification d'expériences pour améliorer la connaissance du modèle est inexistante. Ainsi, seules les expériences faites à l'optimum de performance sont utilisées pour l'estimation paramétrique. Afin de ne pas fausser les résultats par rapport à notre stratégie, cinq échantillons sont prélevés.

3.5.3 Résultats

La Figure 3.26 montre le nombre d'expériences nécessaires pour chacune des trois approches appliquées à chacun des quatre systèmes réactionnels. Il apparaît clairement que les stratégies itératives nécessitent beaucoup moins d'expériences que l'approche tendancielle (de 3 à 20 au moins). On peut trouver trois explications à cela :

- L'utilisation de modèles phénoménologiques pour les stratégies itératives, ce qui est en soit un apport important d'information,
- La non-gestion du domaine opératoire de l'approche tendancielle, ce qui multiplie directement le nombre d'expériences à faire par le nombre de technologies disponibles.
- Le fait que notre stratégie et que l'algorithme de Nancy soient itératifs permet de ne pas faire plus d'expériences que nécessaire.

L'algorithme de Nancy, quant à lui, nécessite un peu plus d'expériences que la stratégie Stanislas pour converger.

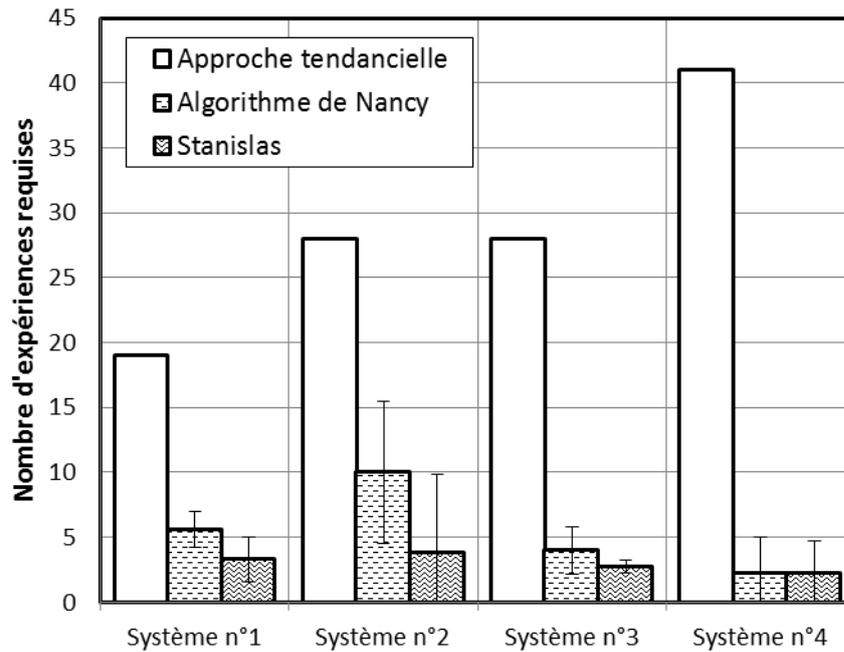


Figure 3.26 : Nombre d'expériences nécessaires jusqu'à convergence (pour chaque système, comparaison des stratégies).

La Figure 3.4 montre l'écart relatif entre critères observés et critère vrai pour chaque système et chaque stratégie. Les résultats obtenus vont dans le même sens que ceux détaillés plus haut. La stratégie Stanislas, en plus de nécessiter moins d'expériences que les autres stratégies, apporte une précision supérieure en moyenne à celle des deux autres stratégies.

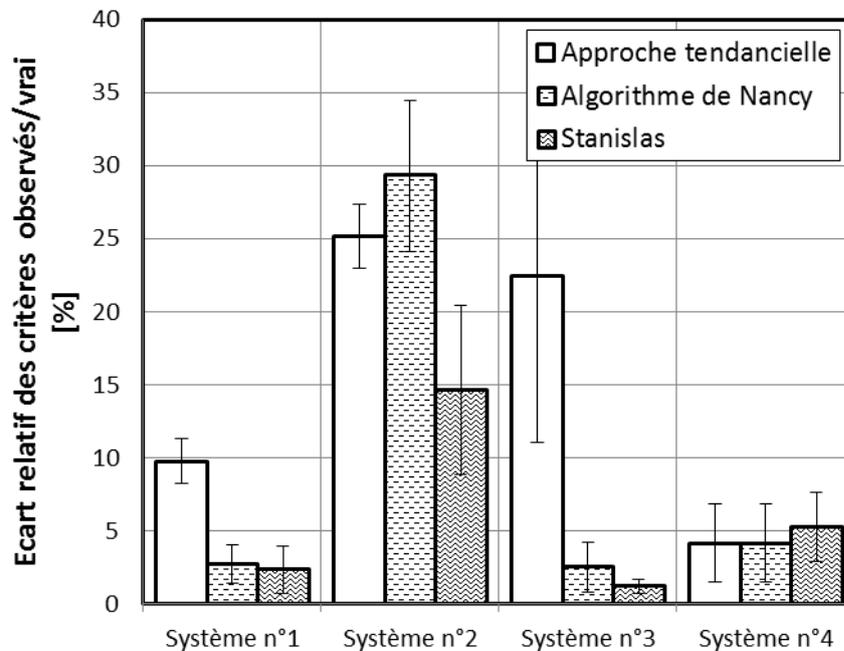


Figure 3.27 : Écart relatif entre critères réels et théoriques (pour chaque système, comparaison des stratégies).

Le fait que notre stratégie est plus efficace que l'Algorithme de Nancy montre le réel gain apporté par l'étape d'acquisition de données. C'est précisément cette étape qui différencie l'Algorithme de Nancy de notre stratégie.

Il faut cependant relativiser les mauvaises performances de l'approche tendancielle. En effet, ces approches sont plutôt destinées à être utilisées dans les cas où il n'est pas possible d'écrire un modèle phénoménologique et que le domaine expérimental est un hypercube. Or, dans notre cas, le domaine expérimental n'est pas un hypercube et n'est pas connexe.

3.6 Conclusion

Appliquer la stratégie Stanislas de manière virtuelle nous a permis d'étudier le comportement de la stratégie Stanislas proposée sur un grand nombre de systèmes. En effet, la génération d'expériences virtuelles permet de réduire considérablement le temps de génération des points expérimentaux. Nous avons donc pu utiliser plusieurs types de schémas réactionnels classiques (compétitif/parallèle, ordre fractionnaire et liquide/liquide).

Pour valider le bienfondé de la stratégie Stanislas, nous avons vérifié, dans un premier temps, que la stratégie converge toujours vers l'optimum vrai. Ce test de répétabilité nous a montré, qu'en général, la stratégie converge (voir Tableau 3.13) et ce pour les différents systèmes réactionnels étudiés.

Dans un deuxième temps, nous avons testé la robustesse de la stratégie vis-à-vis du nombre d'espèces mesurées, de la vitesse des phénomènes et de l'erreur de mesure. Cela nous a montré que dans certains cas la stratégie ne converge pas vers l'optimum. En particulier, lorsque la seule espèce mesurée ne dépend quasiment pas d'un des paramètres du modèle, lorsque l'optimum vrai est situé hors du domaine expérimental et lorsque l'erreur de mesure est trop importante.

Après avoir montré que la stratégie Stanislas converge dans la plupart des cas, nous avons, dans un troisième temps, montré l'importance de considérer plusieurs technologies. Non seulement pour améliorer les performances, mais surtout pour rendre plus efficace l'étape d'acquisition de données. Un résultat connexe a été de montrer la supériorité du réacteur semi-batch sur le réacteur batch liée à sa capacité à maintenir un état instationnaire propice à l'acquisition de données cinétiques, par des profils de coulées adaptés. En outre, le réacteur piston est lui aussi systématiquement supérieur au réacteur batch grâce à sa flexibilité d'échantillonnage.

Enfin, dans un dernier temps, nous avons comparé la stratégie Stanislas avec une stratégie plus conventionnelle par plan d'expérience et avec l'Algorithme de Nancy, stratégie itérative basée sur un modèle phénoménologique. Nous avons mis en évidence que la stratégie par plan d'expériences était nettement moins efficace que les deux autres stratégies. En effet, l'approche par plan d'expériences n'est pas très bien

adaptée à l'utilisation de multiples réacteurs (ce qui rend le domaine des conditions expérimentales accessibles pas forcément connexe et de géométrie plus complexe qu'un simple hypercube). Par ailleurs, l'étape d'acquisition, qui est la principale différence entre l'Algorithme de Nancy et la stratégie Stanislas, permet de réduire le nombre d'expériences et d'améliorer la précision de la convergence surtout dans les cas où le système étudié est très sensible aux paramètres (consécutif/parallèle avec effet de température).

L'étude présentée dans ce chapitre, nous conforte dans l'idée que la stratégie Stanislas est efficace pour trouver la technologie la plus performante et les conditions expérimentales associées. Dans les chapitres suivants, nous confrontons la stratégie à des systèmes expérimentaux réels ; le chapitre 4 concerne l'étude d'une synthèse très simple et le chapitre 5 présente les résultats d'une synthèse industrielle multiphasique complexe.

Chapitre 4

Application à un système réactionnel compétitif en phase homogène

Nous avons utilisé un cas d'étude expérimental dans le but de valider notre stratégie. Nous avons choisi l'iodation de la L-tyrosine, dont le système réactionnel est décrit au paragraphe 4.1.1. C'est un cas d'étude simple, facile à modéliser et facile à mettre en œuvre. Pour montrer l'intérêt d'utiliser plusieurs technologies pour l'acquisition de données, nous avons comparé les précisions de l'estimation des constantes de vitesses obtenues par différentes technologies temps (§ 4.2). Enfin, nous avons appliqué la stratégie dans son ensemble (voir §4.3).

4.1 Description du cas d'étude

4.1.1 La chimie

La synthèse que nous avons choisie pour servir de cas d'étude est l'iodation de la L-tyrosine. C'est une réaction consécutive-parallèle (voir Figure 4.1) car l'iode réagit une première fois avec la tyrosine (abr. T) pour former de la 3-monoiodotyrosine (abr. TI), puis peut réagir une seconde fois avec cette dernière pour former la 3,5-diiodotyrosine (abr. TI₂) :

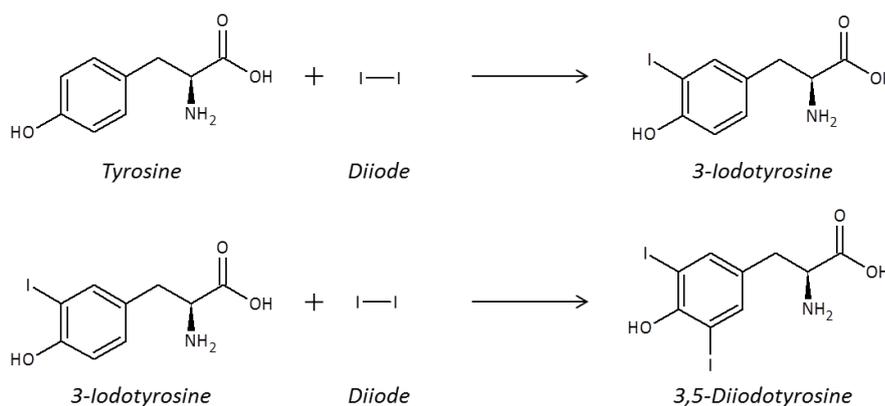
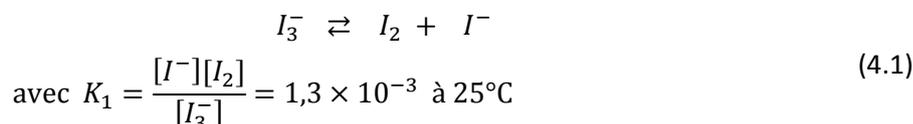


Figure 4.1 : schéma réactionnel de l'iodation de la tyrosine.

Un des intérêts de cette réaction réside dans le fait que les vitesses des réactions peuvent être réglées sur plusieurs ordres de grandeur en jouant sur plusieurs paramètres : la concentration en iodure de potassium, le pH et la température. Cela nous a permis d'ajuster les vitesses de réaction pour qu'elles ne soient ni trop lentes (temps de

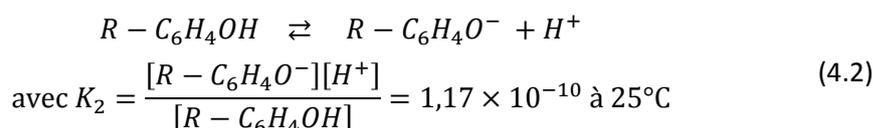
synthèse inférieure à une heure) ni trop rapides (temps de synthèse supérieur à quelques minutes).

Cette réaction a lieu dans l'eau, or le diiode est insoluble dans l'eau. Il faut donc le solubiliser par ajout d'iodure de potassium (au moins 100 fois plus de $[I^-]$ que de $[I_2]$ en moles) afin de former l'ion triiodure par l'équilibre suivant :



La disponibilité en diiode dans le milieu dépend donc de la concentration en ions iodure.

Le pH joue également un rôle important sur la cinétique du système réactionnel. En effet, comme l'a suggéré Mayberry (1964), le diiode réagit avec le cycle aromatique lorsque le groupement alcool est sous forme basique. Or la disponibilité de la forme déprotonée est liée à l'équilibre acido-basique de l'alcool dont le pKa vaut 10,07.



En considérant que chacune des deux réactions est du premier ordre par rapport à chacun des réactifs (diiode et forme basique de l'alcool) on a :

$$r_i = k_i [R - C_6H_4O^-][I_2] \quad (4.3)$$

En utilisant les constantes d'équilibre K_1 et K_2 on peut faire apparaître une constante de vitesse apparente k'_i telle que :

$$r_i = k'_i [T][I_3^-]$$

$$= k_i \frac{K_1 K_2}{[I^-](K_2 + [H^+])} [T][I_3^-] \quad (4.4)$$

$$\text{en posant } [T] = [R - C_6H_4O^-] + [R - C_6H_4OH]$$

Ainsi, en faisant varier la concentration en iodure de potassium de 0,01 à 0,1 mol/L, on peut accélérer la réaction d'un facteur 10, et en faisant varier le pH de 6 à 11 d'un facteur 10000¹⁰. Le pH est donc un facteur très sensible de la réaction. Pour le maintenir constant nous utilisons des solutions tampon avec différentes espèces et ratios acide-base selon le pH voulu.

¹⁰ Ceci est vrai en maintenant la température à 25°C

Tableau 4.1 : Couple acide-base utilisé pour les tampons de pH

Gamme de pH	Acide	Base
6 - 8,5	NaH_2PO_4	Na_2HPO_4
8,5 - 12	H_2CO_3	$NaHCO_3$
11 - 13	Na_2HPO_4	Na_3PO_4

4.1.2 Méthodes d'analyse

Pour ce système réactionnel, nous avons utilisé deux méthodes d'analyse : une méthode permettant de suivre la concentration en ion triiodure servant à faire l'acquisition des données cinétiques, et une seconde méthode permettant de doser la concentration des autres espèces. Cette dernière méthode est destructrice, et sera utilisée uniquement pour la mesure du rendement de la réaction lors de l'étape de validation des prédictions du modèle.

L'ion triiodure, qui est directement lié au diiode par l'équilibre (5.1), possède l'intéressante propriété d'absorber très fortement la lumière dans l'ultraviolet. Or, il est très commode d'un point de vue expérimental de mesurer en ligne l'absorbance UV car la mesure est directe (pas de préparation d'échantillon), instantanée, non destructive. La Figure 4.2 montre le spectre d'absorbance de l'ion triiodure :

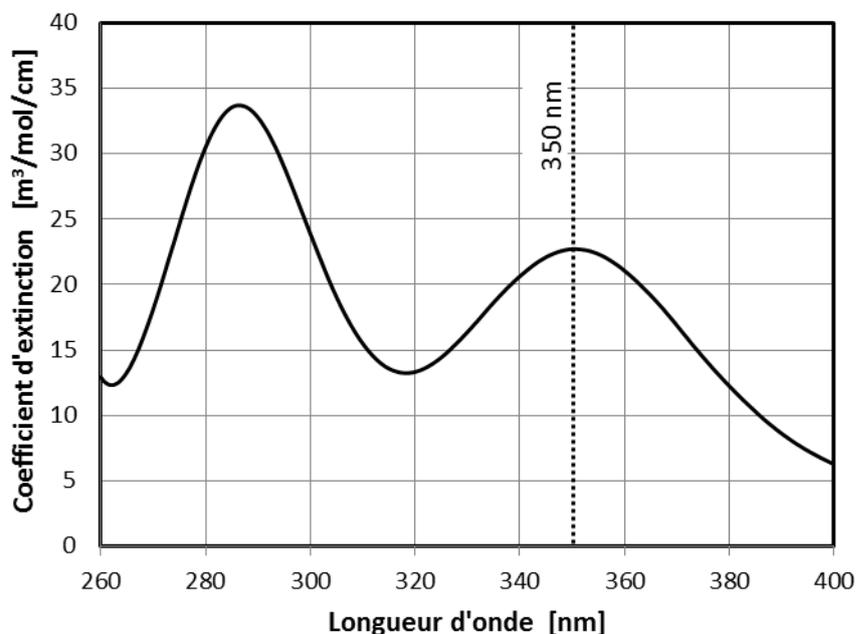
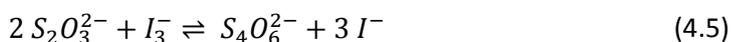


Figure 4.2 : spectre d'adsorbance de l'ion triiodure
avec $[I_3^-] = 5 \times 10^{-5} \text{ mol/L}$

La longueur d'onde que nous retiendrons pour le suivi de la quantité d'iode en solution est 350 nm. À cette longueur d'onde, les autres espèces en solution absorbent

très faiblement (voir Figure 4.3) comparativement à l'ion triiodure. On considère donc, qu'à cette longueur d'onde seule cette dernière espèce est mesurée. La mesure dans la cuve agitée se fait en ouvrant la vanne d'échantillonnage pour que le milieu réactionnel passe à travers la cellule UV. Avec le réacteur tubulaire, l'acquisition se fait en continu, et la mesure retenue est prise lorsque la stabilité du système est atteinte (soit environ deux temps de passage, soit encore à six écart-type de la DTS).

On peut quantifier la tyrosine, la 3-iodo-tyrosine et la 3,5-diiodo-tyrosine par décomposition du spectre UV après avoir éliminé l'ion triiodure par ajout de thiosulfate de sodium. Lors de l'ajout de thiosulfate, on note une décoloration totale et instantanée de la solution liée à la disparition des ions triiodure. Comme nous n'avons accès qu'à une mesure in-situ, le fait d'éliminer l'ion triiodure implique que l'on ne puisse faire qu'une seule mesure des différentes espèces.



Chacune de ces trois espèces possède un spectre UV différent pour la forme acide et pour la forme basique. La Figure 4.3 présente les spectres d'absorbance de chaque molécule obtenus expérimentalement par régression linéaire sur la gamme de concentration suivante (0,04 ; 0,08 ; 0,12 ; 0,14 ; 0,16 ; 0,18 ; 0,20 mol/m³). Pour doser ces trois espèces à n'importe quel pH, on dose en réalité les six espèces (les trois molécules sous leurs deux formes acide et basique), puis on somme la forme basique et la forme acide de chaque espèce.

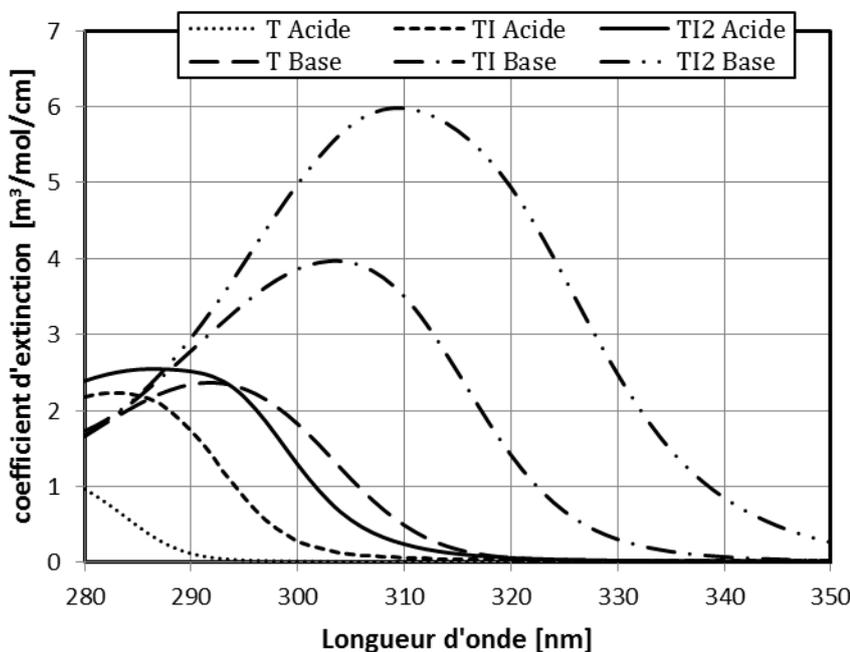


Figure 4.3 : Spectre UV des formes acide-base de chaque composé.

Pour vérifier la validité de la mesure, nous avons créé un échantillon de validation ayant des concentrations connues ($[T] = 0,075 \text{ mol/m}^3$; $[TI] = 0,050 \text{ mol/m}^3$; $[TI_2] = 0,025 \text{ mol/m}^3$). Une fois le spectre de la solution mesurée (voir Figure 4.4), il a

été décomposé comme la somme des six spectres présentés dans la Figure 4.3. Pour ce faire, nous avons utilisé une régression linéaire multivariée. Soit A le vecteur colonne contenant l'absorbance de la solution pour chaque longueur d'onde, l la longueur de la cuve, ε la matrice des coefficients d'extinction de chaque espèce et C le vecteur de concentrations inconnues. Nous avons l'expression suivante :

$$A = l \varepsilon C \quad (4.6)$$

L'estimation de \hat{C} est alors obtenue par :

$$\hat{C} = \frac{1}{l} (\varepsilon^T \varepsilon)^{-1} \varepsilon^T A \quad (4.7)$$

On peut aussi estimer la matrice \hat{A}_c de l'absorbance de chaque espèce contenue dans la solution par :

$$\hat{A}_c = l \varepsilon \text{diag}(\hat{C}) \quad (4.8)$$

On remarque que les deux formes (acide et basique) de chaque espèce sont bien présentes.

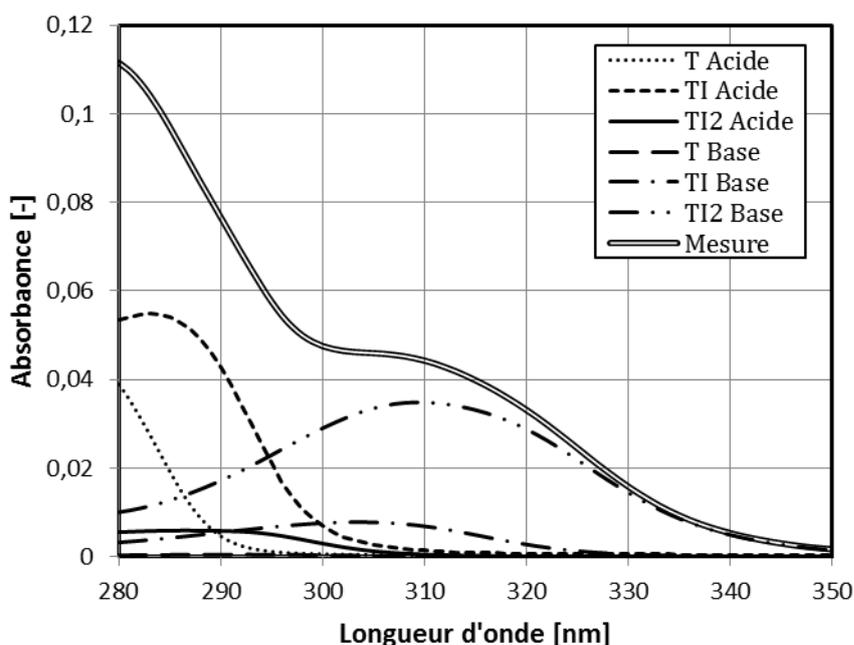


Figure 4.4 : Décomposition du spectre de l'échantillon de validation.

Après avoir sommé les concentrations acide-base de chaque espèce on obtient les résultats donnés par la Figure 4.5. Les valeurs estimées des concentrations sont très proches des concentrations connues, la méthode d'analyse des espèces fournit des résultats cohérents et des intervalles de confiance à 95% d'environ $0,005 \text{ mol/m}^3$.

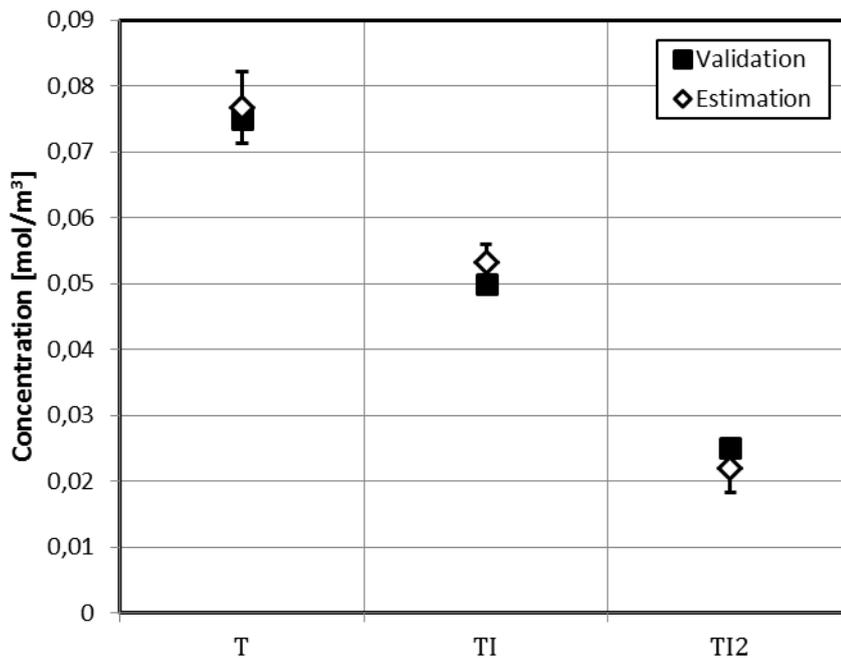


Figure 4.5 : Validation de la méthode de dosage par décomposition du spectre UV.

4.1.3 Simplification du problème

Ce cas d'étude doit être relativement simple pour mieux comprendre le comportement de la stratégie. Ainsi, nous avons délibérément choisi de ne pas tenir compte de l'effet de température : la température a été fixée à 25°C. De plus, nous avons aussi fixé la concentration en iodure ($[I^-] = 0,015 \text{ mol/L}$) ainsi que le pH ($\text{pH} = 7,40$). Le système se résume donc à deux réactions consécutive-parallèle. Le système réactionnel est détaillé ci-dessous.

- La seule grandeur mesurée pour l'étape d'acquisition de données est la concentration en ions triiodure. Et cela nous permettra de montrer que même en ne mesurant qu'une seule espèce, la stratégie Stanislas converge vers un optimum de performance.
- Les paramètres du modèle sont les deux constantes de vitesse apparente k'_1 et k'_2 présentée dans l'équation (4.4) telle que $r_1 = k'_1 [T][I_3^-]$ et $r_2 = k'_2 [TI][I_3^-]$. Pour simplifier les notations, nous les noterons k_1 et k_2 par la suite.
- Les gammes de conditions expérimentales disponibles pour chacun des réacteurs sont présentées dans le Tableau 4.2.

Tableau 4.2 : Gammes des conditions expérimentales pour la tyrosine.

		Batch	Piston	Semi-Batch
t	[min]	5 – 60	0,5 – 7,5	5 – 60
$Ratio = [T]^0 / C_{TOT}^0$	[%]	5 – 95	5 – 95	5 – 95
$C_{TOT}^0 = [T]^0 + [I_3^-]^0$	[mol/m ³]	0,01 – 0,1	0,01 – 0,1	0,01 – 0,1
Coulée I_3^-	[mL/min]	aucune	aucune	4 paliers
Échantillonnage		t	t et $Ratio$	t

4.1.4 Estimation indépendante

Afin de pouvoir juger de la validité des résultats obtenus dans les sections 4.2 et 4.3, nous avons estimé les constantes de vitesse k_1 et k_2 de manière indépendante. Pour réaliser l'estimation indépendante des deux constantes de vitesses, nous avons procédé en deux temps :

- Premièrement, nous avons estimé la constante de la seconde réaction en utilisant de la 3-iodo-tyrosine pure en quantité équimolaire par rapport au diiode. Il est important de noter que dans les paragraphes 4.2 et 4.3 nous avons supposé qu'il n'était pas possible d'utiliser le produit intermédiaire (la 3-iodo-tyrosine) comme réactif de départ.
- Puis, dans un second temps, nous avons utilisé un large excès de tyrosine (rapport molaire de 4 pour 1) pour minimiser l'impact de la seconde réaction.

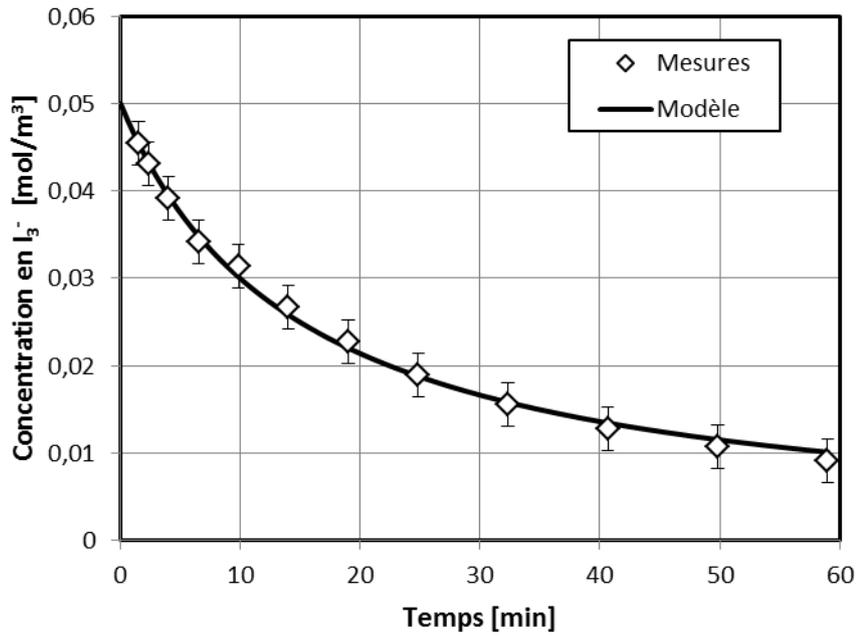


Figure 4.6 : Estimation indépendante k_1 .
 $([T]^0 = [I_3^-]^0 = 0,05 \text{ mol/m}^3)$

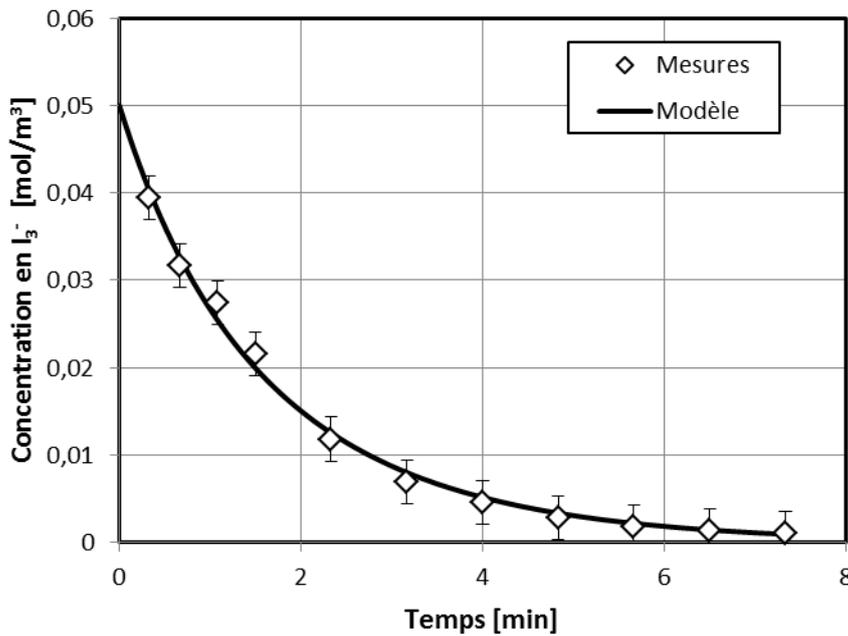


Figure 4.7 : Estimation indépendante k_2 .
 $([T]^0 = 0,20 \text{ mol/m}^3 \text{ et } [I_3^-]^0 = 0,05 \text{ mol/m}^3)$

Nous avons obtenu les valeurs données dans le Tableau 4.3. Vu la précision de l'estimation indépendante, nous considérerons par la suite comme étant la valeur vraie des paramètres ($\theta_{ind} = \theta^*$).

Tableau 4.3 : Estimation indépendante des paramètres.

		k_1	k_2
Estimation	[m ³ /mol/s]	0,0562 ± 0,0020	0,0193 ± 0,0021

4.2 Validation expérimentale de l'acquisition de données

Pour l'application expérimentale de la stratégie, nous avons voulu avancer pas à pas. Ainsi, dans un premier temps, nous avons voulu valider l'intuition et corroborer les résultats obtenus au paragraphe 3.4, à savoir qu'utiliser plusieurs technologies pour l'acquisition de données est un avantage non négligeable. L'objectif a donc été de quantifier la précision avec laquelle les paramètres k_1 et k_2 sont estimés en fonction de la technologie utilisée sans prendre en compte l'aspect « optimisation des performances ». Les trois technologies (batch, piston et semi-batch) ont d'abord été utilisées séparément et nous avons laissé dans un dernier temps la liberté de choisir la technologie la plus adaptée à l'acquisition de données lors de chaque itération.

Ainsi, l'objectif est d'estimer du mieux possible les deux constantes cinétiques k_1 et k_2 en faisant 4 expériences D-optimales où 8 échantillons sont prélevés. L'idée est donc au bout du même nombre d'expériences (4 est un compromis entre évolution de la confiance des paramètres et temps passé au laboratoire) de comparer la performance de chaque réacteur du point de vue de l'acquisition de données. La Figure 4.8 représente le protocole utilisé pour cette section.

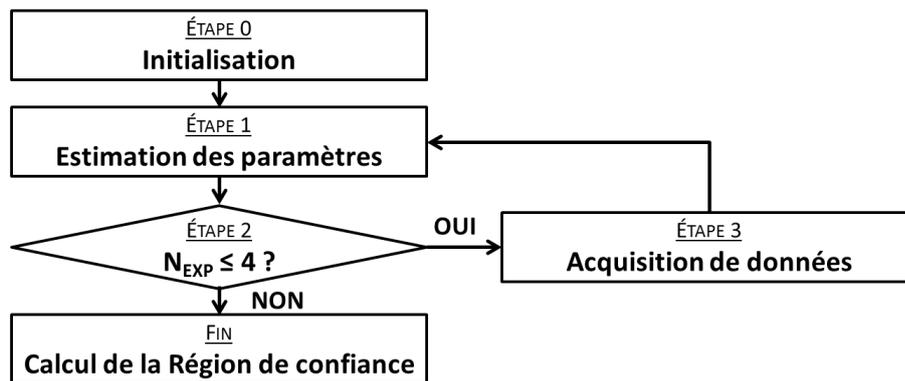


Figure 4.8 : Protocole utilisé pour la section 4.2.

4.2.1 Réacteur batch

Les conditions de l'expérience préliminaire sont les suivantes :

- Temps d'échantillonnage toutes les 5 minutes, de 5 à 40 minutes.
- Ratio en tyrosine égal à 33%, de manière à avoir deux fois plus d'iode que de tyrosine.
- La concentration totale est de $0,075 \text{ mol/m}^3$.

La Figure 4.9 montre la régression non-linéaire de l'expérience préliminaire. À l'issue de cette première estimation nous avons obtenu des valeurs des paramètres préliminaires qui nous ont servi à planifier l'expérience D-optimale suivante.

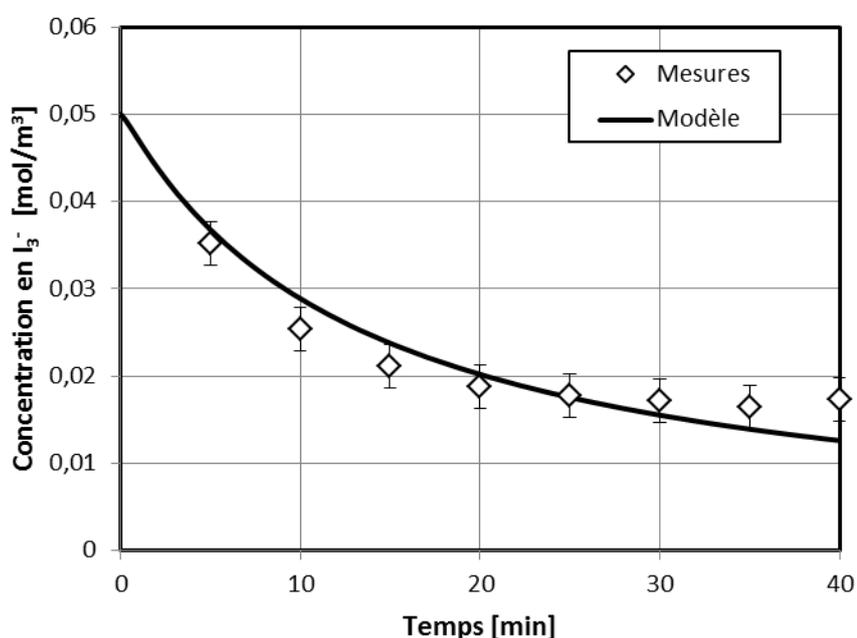


Figure 4.9: Régression paramétrique de l'expérience préliminaire.

Une fois cette nouvelle expérience réalisée, une nouvelle estimation des paramètres a été obtenue et ainsi de suite 4 fois. La Figure 4.10 montre comment sont répartis les 8 échantillons de chacune des 4 expériences D-optimales. Il s'agit d'une projection : l'axe de la concentration totale en réactif n'est pas représenté. Cependant, la concentration totale ne diffère que très peu d'une expérience à l'autre (de $0,075$ à $0,095 \text{ mol/m}^3$). On peut d'ores et déjà remarquer que la stratégie ne balaye pas l'ensemble du domaine expérimental de manière homogène.

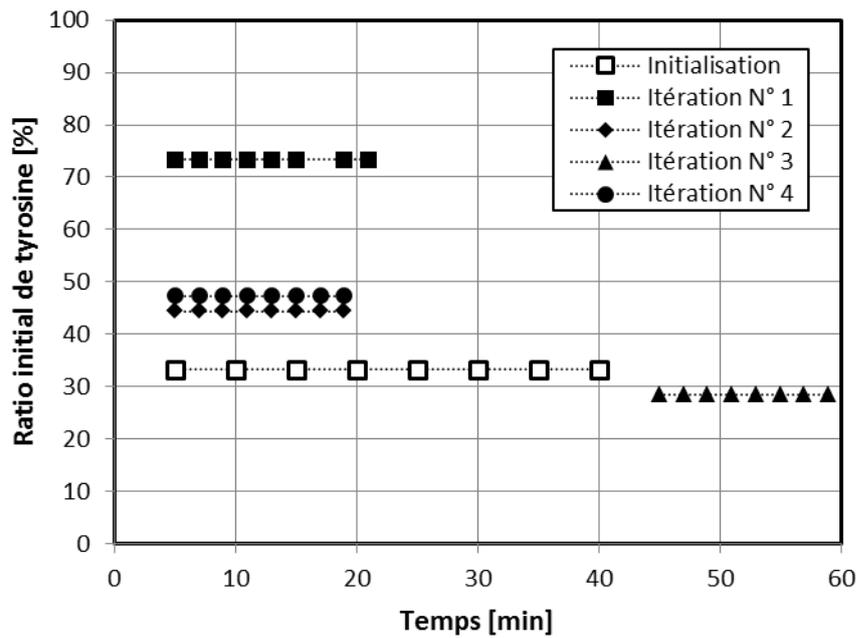


Figure 4.10 : Répartition des échantillons après quatre expériences *D*-optimales en réacteur batch (la concentration totale n'est pas représentée).

La Figure 4.11 montre l'estimation des paramètres à chaque itération et l'ellipsoïde associé (indice de confiance $1 - \alpha = 95\%$). Le premier constat est qu'après chaque itération, la taille de l'ellipsoïde ne diminue pas de manière significative. Le second point important est que l'ellipsoïde de confiance est très fortement allongé. Cela veut dire que les paramètres k_1 et k_2 sont fortement corrélés ; il n'est pas possible de les estimer séparément.

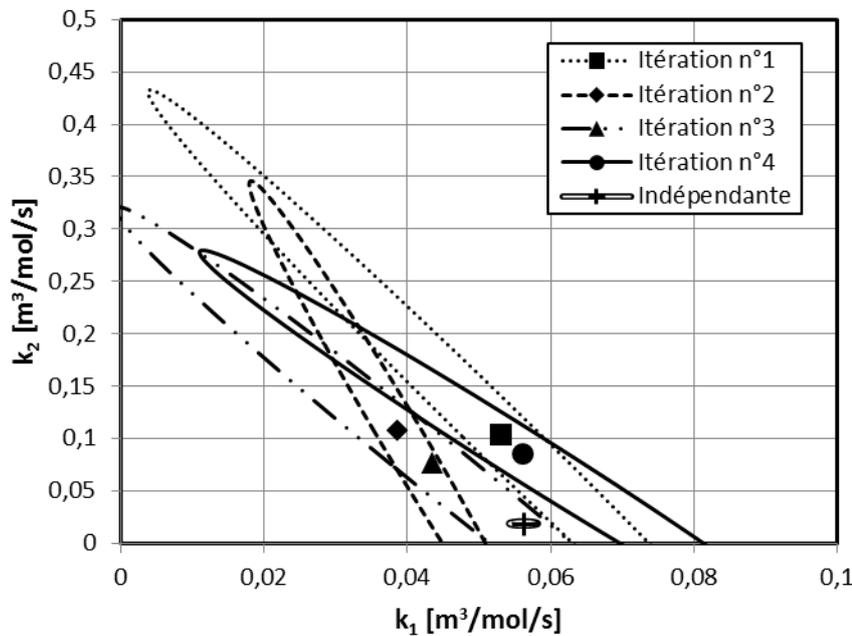


Figure 4.11: Évolution de l'ellipsoïde de confiance en réacteur batch au cours des quatre itérations.

4.2.2 Réacteur piston

Nous avons eu la même démarche pour le réacteur piston que pour le réacteur batch : nous regardons comment évolue la confiance dans les paramètres au bout de quatre expérience D-optimale. Ce qui diffère entre le piston et le batch c'est, d'une part, le domaine opératoire (temps de séjour plus court), et d'autre part, la flexibilité d'échantillonnage (possibilité d'échantillonner sur le temps et le ratio initial de réactif en même temps). Pour l'expérience préliminaire, il a été arbitrairement choisi de fixer le temps de passage à 2 minutes et de faire varier le ratio molaire de 20% à 55%, la concentration totale en réactif étant fixée à $0,1 \text{ mol/m}^3$. On obtient la Figure 4.12, qui représente la régression paramétrique de l'expérience préliminaire. L'axe des abscisses n'est plus le temps comme précédemment, mais le ratio en tyrosine.

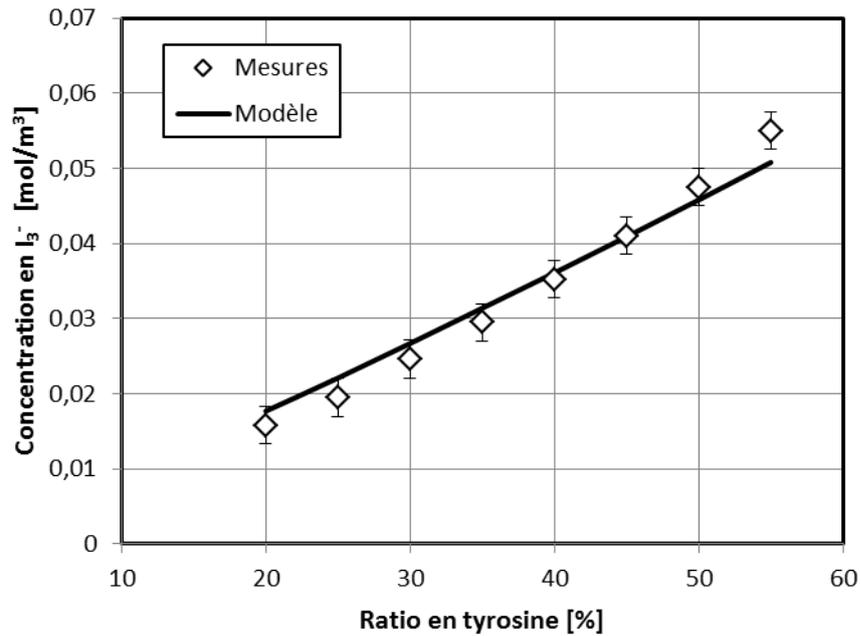


Figure 4.12 : Régression paramétrique pour l'expérience préliminaire en réacteur piston (l'axe des abscisses représente le ratio en tyrosine).

Une fois que l'estimation préliminaire des paramètres a été faite, nous avons fait chercher les expériences D-optimales. Les conditions de ces expériences sont présentées dans la Figure 4.13. Dans ce cas, la concentration totale en réactif ne varie pas du tout, pour toute les expériences, $C_{TOT}^0 = 0,1 \text{ mol/m}^3$.

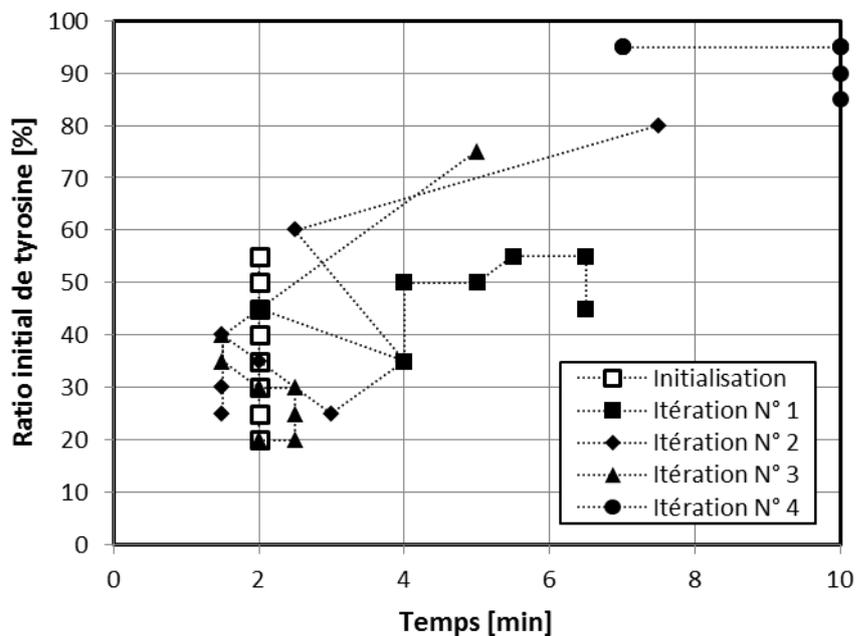


Figure 4.13 : Répartition des échantillons après quatre expériences D-optimales en réacteur piston (la concentration totale n'est pas représentée).

Comme pour le réacteur batch, les points de prélèvement ne sont pas répartis de façon intuitive. Cependant, la zone des temps de passage court et forte concentration en tyrosine (partie supérieure gauche) et la zone des temps de passage long et excès d'iode (partie inférieure droite) ne sont pas du tout explorées. Qualitativement cela est bien compréhensible, car dans ces zones, le système ne varie quasiment pas. Ce qui montre qu'il y a peu d'information sur la cinétique. De plus, il faut remarquer que certains points peuvent être répétés plusieurs fois.

La Figure 4.14 montre l'évolution de l'estimation des paramètres et leur ellipsoïde de confiance au fur et à mesure des itérations. Il est important de noter que l'échelle de l'axe des ordonnées (k_2) de cette figure est dix fois plus resserrée que pour la Figure 4.11. Ceci implique que l'estimation que l'on obtient de k_1 et k_2 est beaucoup plus précise. En outre, les ellipses sont peu allongées, les paramètres ont donc été décorrélés. On remarque aussi que l'estimation indépendante est incluse dans les ellipses des itérations 3 et 4.

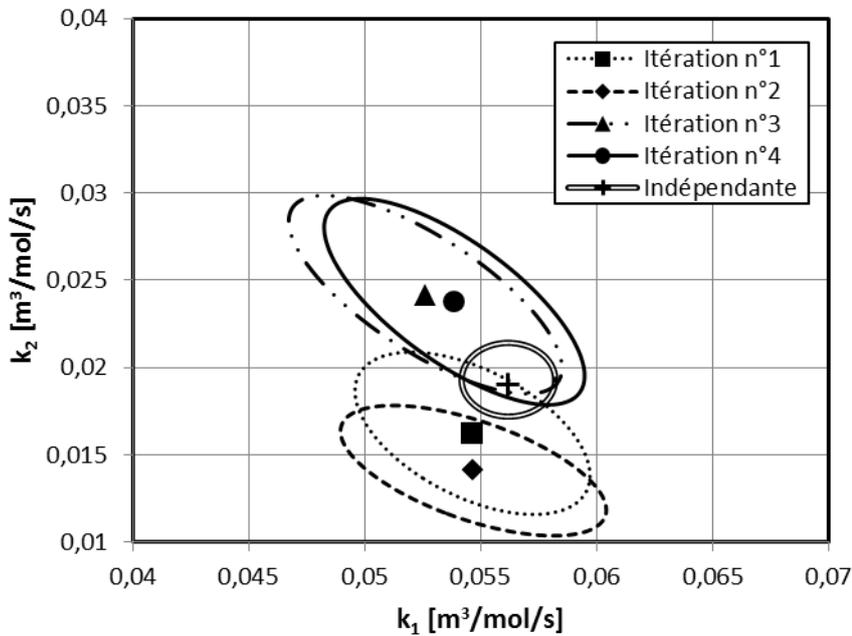


Figure 4.14: Évolution de l'ellipsoïde de confiance en réacteur piston au cours des quatre itérations.

La différence entre le réacteur piston et le réacteur batch provient d'une part de la fenêtre expérimentale (le réacteur piston permet d'accéder à des temps de passage plus courts) et, d'autre part, de sa flexibilité d'échantillonnage (en une seule expérience, le réacteur piston permet de balayer plusieurs temps de passage et plusieurs ratios).

4.2.3 Réacteur semi-batch

L'estimation préliminaire pour le réacteur semi-batch a été réalisée en prélevant 8 échantillons de 5 à 40 minutes avec un ratio initial de tyrosine de 50%. Au bout de 10 minutes, un équivalent molaire d'iode a été coulé sur une durée de 10 minutes comme le montre la Figure 4.15.

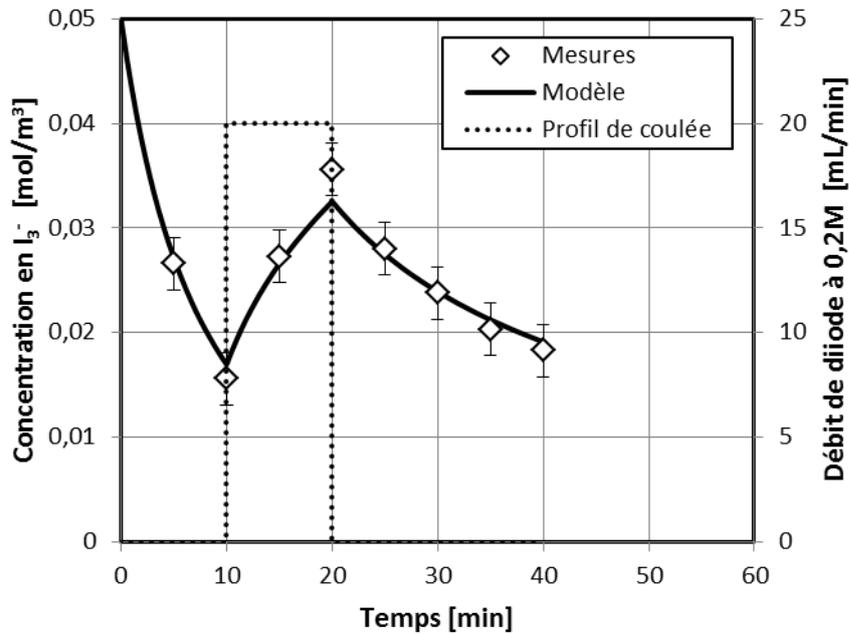


Figure 4.15: Régression paramétrique pour l'expérience préliminaire en réacteur semi-batch.

La Figure 4.16 montre comment sont répartis les points de prélèvements au cours des itérations. Là encore, leur répartition n'est pas intuitive. De plus, les profils de coulée à chaque itération ne sont pas non plus intuitifs (voir Figure 4.17, les profils ont été décalé pour plus de lisibilité).

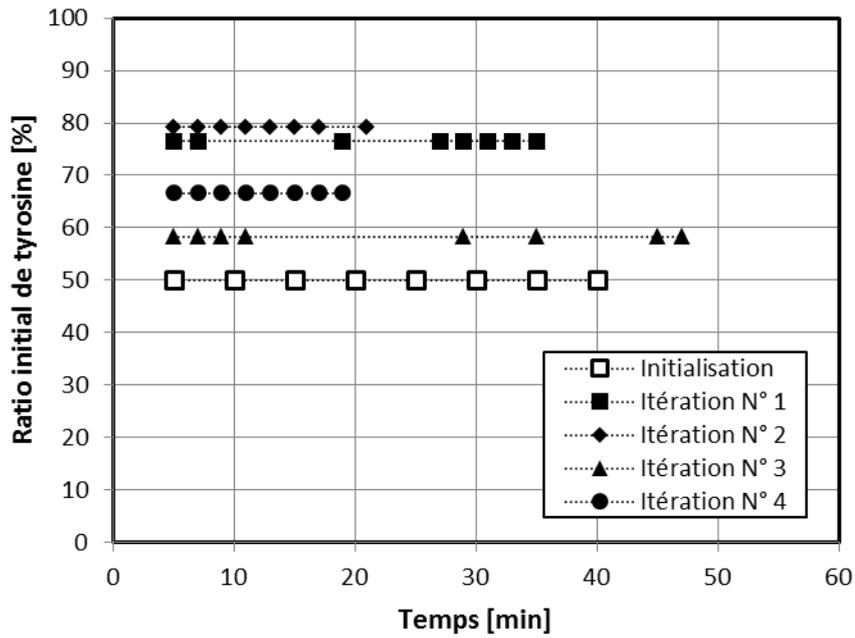


Figure 4.16: Répartition des échantillons après quatre expériences D-optimales en réacteur semi-batch (la concentration totale et les profils de coulée ne sont pas représentés).

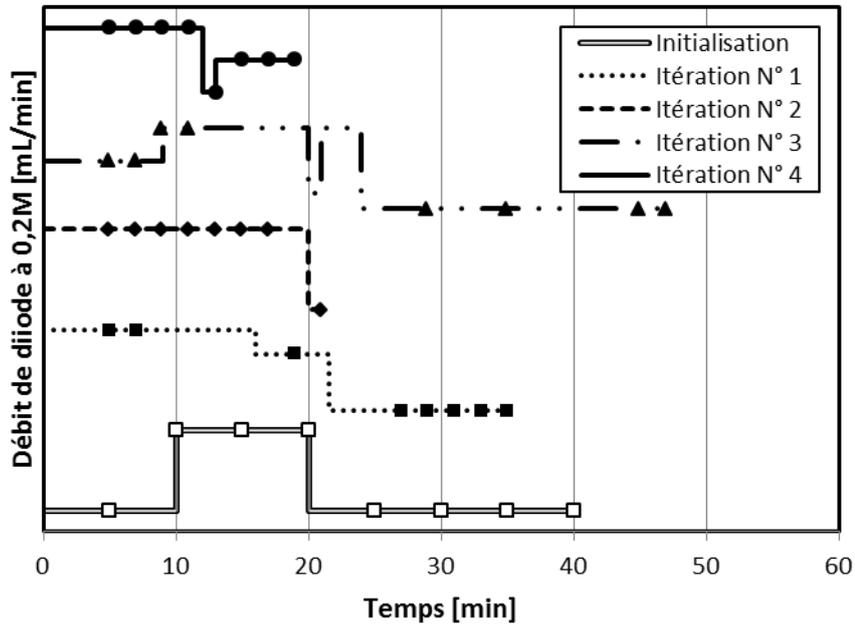


Figure 4.17 : Profil de coulée des quatre expériences D-optimales en réacteur semi-batch.

La Figure 4.18 montre un exemple de régression paramétrique avec un profil de coulée (itération N°1), le modèle semble bien représenter les mesures.

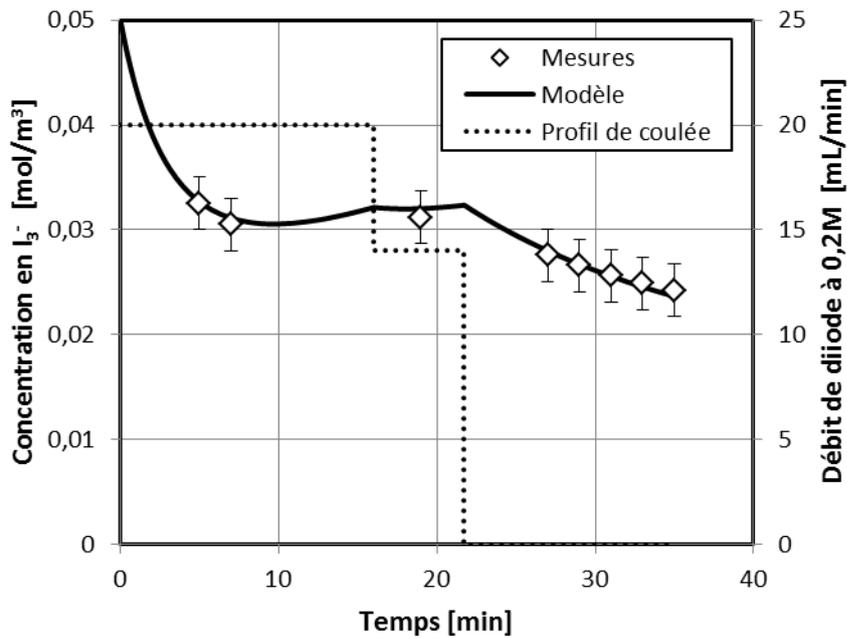


Figure 4.18: Régression paramétrique pour l'itération n°1 en réacteur semi-batch.

La Figure 4.19 montre l'évolution de l'estimation des paramètres. L'échelle est la même que pour le réacteur piston. La précision de l'estimation des paramètres obtenue en réacteur semi-batch est similaire à celle obtenue en réacteur piston. On peut remarquer que la taille de l'ellipsoïde de confiance ne diminue pas significativement au cours des itérations.

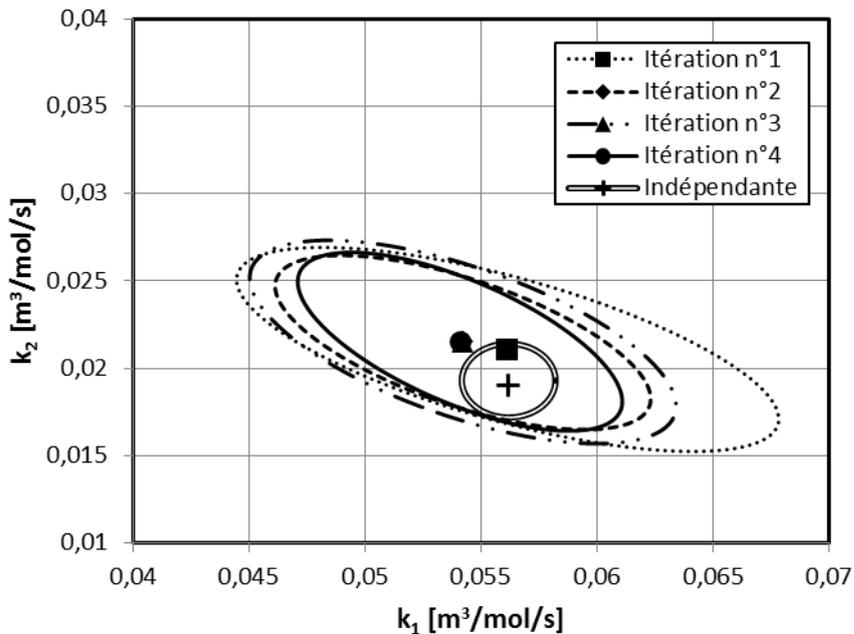


Figure 4.19: Évolution de l'ellipsoïde de confiance en réacteur semi-batch au cours des quatre itérations.

Les meilleurs résultats obtenus avec le semi-batch en comparaison avec le batch peuvent être expliqués par une analyse de sensibilité. La sensibilité du modèle par rapport au paramètre i dans les conditions expérimentales du point de mesure ξ est définie par :

$$s_i = \left. \frac{\partial f}{\partial \theta_i} \right|_{(\xi, \theta)} \quad (4.9)$$

La sensibilité est directement reliée à l'acquisition de données par la matrice d'information de Fisher (équations 1.54 et 1.55). Ainsi, à un point de mesure donné, plus le modèle est sensible aux paramètres, plus l'acquisition de données est efficace. La Figure 4.20 présente la sensibilité du modèle dans les conditions expérimentales de l'expérience préliminaire en réacteur batch. Comme la sensibilité dépend de la valeur des paramètres utilisés, on utilise l'estimation indépendante des paramètres pour tracer cette figure.

Il apparaît clairement que k_1 est plus sensible que k_2 . Le paramètre k_1 atteint son maximum de sensibilité dans les 5 premières minutes alors que k_2 atteint son maximum après 15 minutes. Au moment où k_2 est le plus sensible, la sensibilité de k_1 reste encore importante. Ceci a pour conséquence de corrélérer l'estimation des deux paramètres. Ainsi, on ne peut estimer chaque paramètre de manière indépendante, les deux paramètres sont fortement couplés. C'est pourquoi la région de confiance des paramètres estimés après l'expérience préliminaire en réacteur batch est fortement inclinée (voir Figure 4.11).

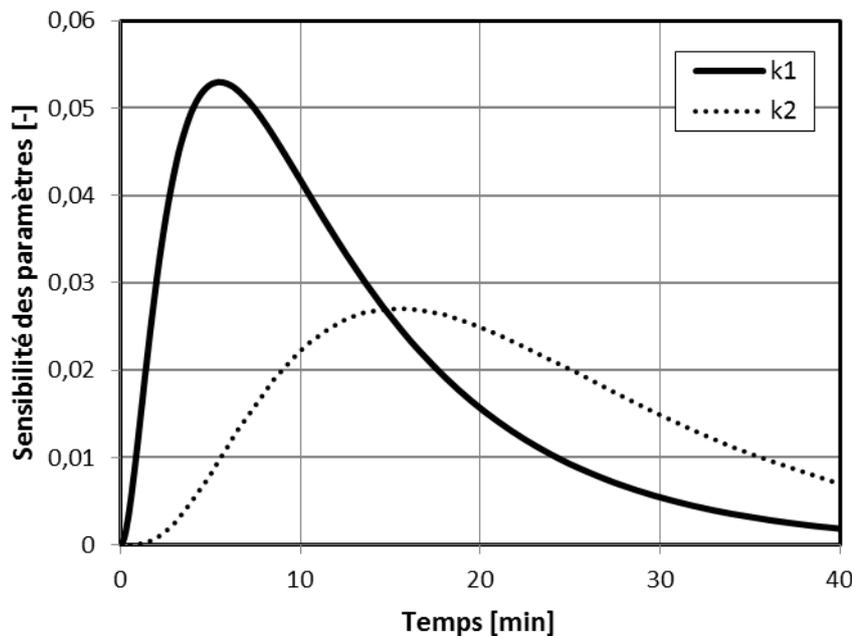


Figure 4.20: Sensibilité du modèle par rapport aux paramètres dans les conditions de l'expérience préliminaire en réacteur batch.

La Figure 4.21 présente les sensibilités du modèle dans les conditions de l'expérience préliminaire en réacteur semi-batch (on utilise aussi l'estimation indépendante des paramètres pour le calcul des sensibilités). Cette figure montre clairement que le fait d'introduire du diiode au cours de la réaction (grâce au profil de coulée) permet d'obtenir une sensibilité sur k_2 nettement supérieure que la même expérience sans profil de coulée. De plus, au maximum de sensibilité de chaque paramètre, la sensibilité du second paramètre est très faible, ce qui a pour conséquence de décorréler les paramètres. On pourra remarquer que lors des itérations n°1 et 3, les temps de prélèvements préconisés par le critère D-optimal sont situés au tout et tout à la fin de la réaction. Ainsi, le réacteur semi-batch permet dans ce cas d'avoir une estimation de k_1 et de k_2 plus fiable que le réacteur batch.

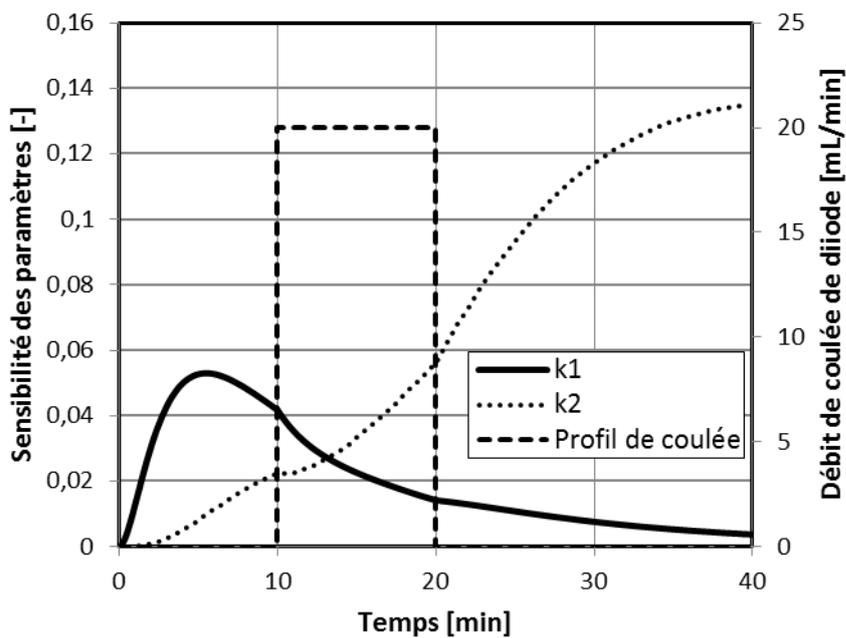


Figure 4.21: Sensibilité du modèle par rapport aux paramètres dans les conditions de l'expérience préliminaire en réacteur semi-batch.

4.2.4 Libre choix des réacteurs

L'estimation préliminaire des paramètres est la même dans ce cas que dans le cas du réacteur batch. Le choix du réacteur étant libre, le réacteur utilisé à chaque itération est choisi comme étant celui qui est le plus adapté pour l'acquisition de données. C'est-à-dire celui qui permet la meilleure expérience du point de vue du critère D-optimal. De cette manière, les deux premières expériences ont été faites en piston, et les deux dernières en semi-batch, comme le montre la Figure 4.22.

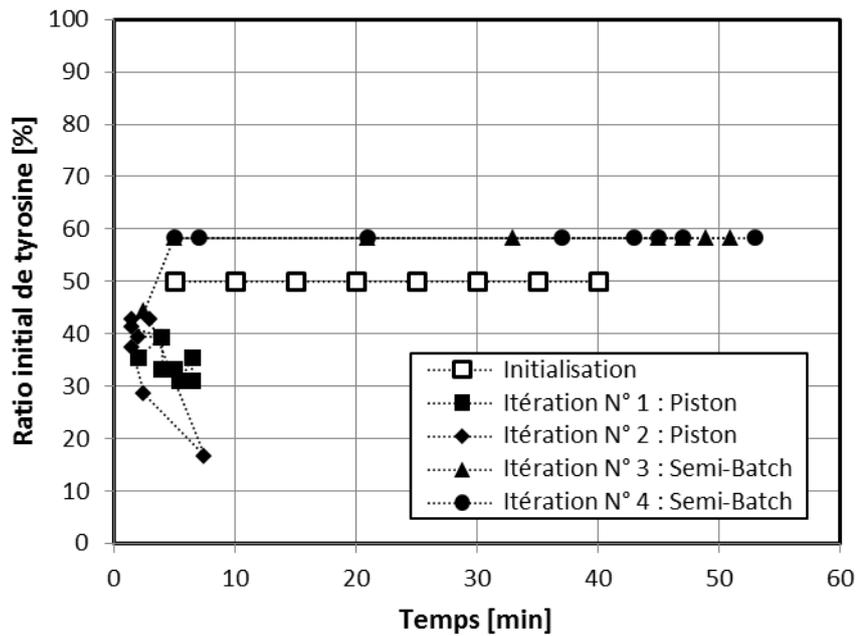


Figure 4.22: Répartition des échantillons après quatre expériences *D*-optimales avec libre choix du réacteur (la concentration totale et les profils de coulée ne sont pas représentés).

La Figure 4.23 représente l'évolution de l'estimation des paramètres en utilisant plusieurs réacteurs. L'échelle est la même que précédemment. Au bout de 4 itérations, l'ellipsoïde de confiance est du même ordre de grandeur que l'ellipsoïde de confiance obtenue lors de l'estimation indépendante.

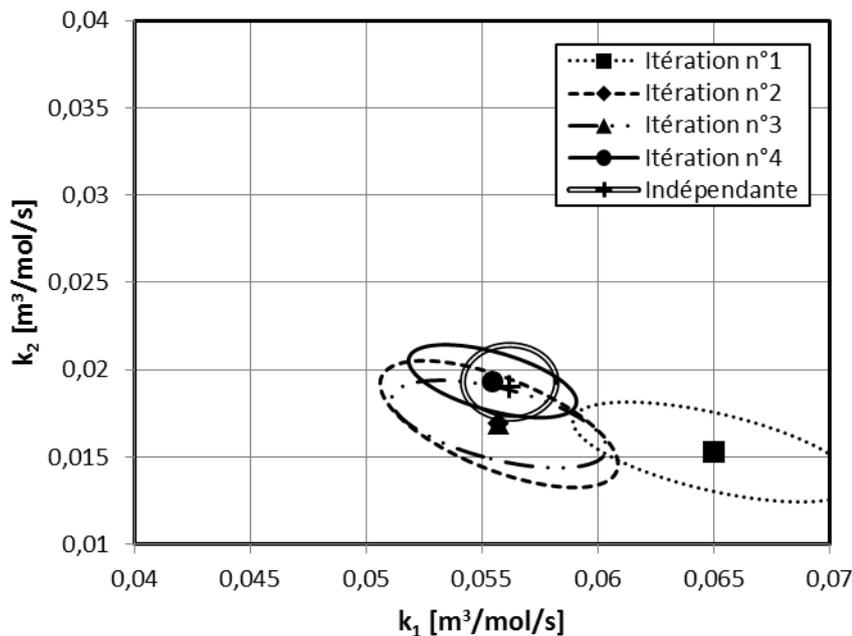


Figure 4.23: Évolution de l'ellipsoïde de confiance avec le libre choix du réacteur au cours des quatre itérations.

Le fait d'utiliser plusieurs réacteurs pour l'acquisition de données semble permettre une meilleure estimation des paramètres que le fait d'utiliser une technologie exclusivement. En effet, lorsque l'on compare les ellipsoïdes de confiance après 4 itérations (voir Figure 4.24 et Figure 4.25), la stratégie qui combine les réacteurs apparaît comme étant nettement meilleure.

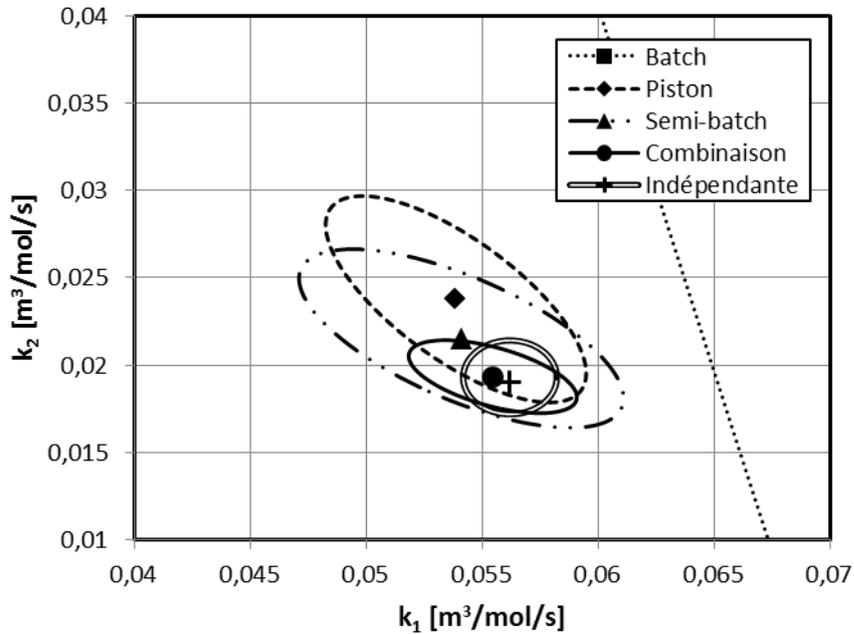


Figure 4.24: Comparaison des ellipsoïdes de confiance obtenus pour chaque réacteur après quatre itérations.

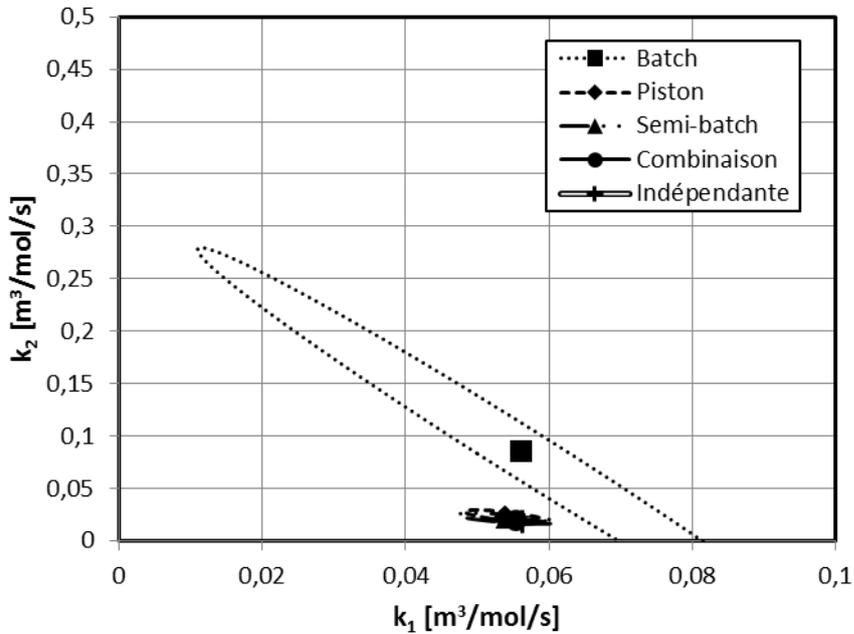


Figure 4.25 : Même figure que Figure 4.24 mais l'échelle est dézoomée.

Ce travail permet de mettre en lumière l'efficacité de l'utilisation de différentes technologies pour acquérir des données. L'importance de la fenêtre expérimentale et la flexibilité de l'échantillonnage est démontrée par la supériorité du réacteur piston sur le réacteur batch. La supériorité du réacteur semi-batch sur le réacteur batch et le réacteur piston montre qu'être dans un état transitoire en utilisant un profil d'alimentation complexe est important pour l'acquisition de données.

4.3 Validation expérimentale de la stratégie pour l'optimisation de performances

Dans ce paragraphe, nous appliquons la stratégie dans son ensemble avec pour but de choisir la technologie optimale afin d'améliorer les performances de la synthèse. Ces performances n'ont aucune réalité industrielle, puisque le système réactionnel est un cas d'étude développé pour la démonstration de l'efficacité de la stratégie.

4.3.1 Initialisation

Le domaine expérimental est le même que celui défini au paragraphe 4.1.3 à ceci près que la concentration totale est fixée à $C_{TOT}^0 = 0,1 \text{ mol/m}^3$. L'objectif est de synthétiser la 3-iodo-tyrosine en un minimum de temps. Ainsi, la performance du système est évaluée par l'agrégation (présentée au paragraphe 1.1.5) de deux critères : le temps d'opération, et le nombre de moles de 3-iodo-tyrosine ayant été créées sur le nombre total de moles de réactifs (tyrosine et diiode) introduites dans le réacteur. Le Tableau 4.4 donne les valeurs « objectifs » et « véto » de chaque critère.

Tableau 4.4 : Critères de performance de la synthèse de la 3-iodotyrosine.

		Critère 1 [s]	Critère 2 [%]
Performance	γ_i	t_f	$Y_{TI} = -100 \times \frac{1}{2} \cdot \frac{n_{TI}}{n_{I_3^0} + n_{I_3^0} + n_{I_3^E}}$
Objectif	o_i	0	- 95
Veto	v_i	1200	- 5

Pour pouvoir mesurer le critère 2, il est nécessaire de pouvoir mesurer la concentration en produit intermédiaire (la 3-iodo-tyrosine), or la méthode d'analyse utilisée pour cela est destructive (voir §4.1.2). Ainsi, pour les expériences d'acquisition de données, seules l'ion triiodure est mesuré. Tandis que pour les expériences de validation, la

réaction est stoppée par ajout de thiosulfate de sodium dans le milieu réactionnel et seule la 3-iodo-tyrosine est mesurée.

4.3.2 Évolution de la stratégie

Pour réaliser l'expérience préliminaire nous avons utilisé un réacteur batch. Les 8 échantillons ont été prélevés à cinq minutes d'intervalle entre $t = 5$ min et $t = 40$ min. Le ratio molaire de tyrosine a été fixé à 50%. La régression paramétrique aboutit à la Figure 4.26.

La stratégie nécessite deux itérations avant que le modèle ne soit suffisamment précis pour que le test de convergence de l'étape 3 soit satisfait (voir le schéma complet de la stratégie Stanislas Figure 2.2 page 70). Le Tableau 4.5 présente les conditions des expériences proposées par la stratégie pour améliorer l'estimation paramétrique. On peut noter que la première expérience d'acquisition de données est réalisée en réacteur piston, alors que la seconde est réalisée en réacteur semi-batch. Pour plus de lisibilité, nous avons tracé les régressions paramétriques de l'expérience préliminaire (Figure 4.26) et de l'expérience D-optimale de l'itération 2 (Figure 4.28). Pour l'itération 1, l'expérience étant faite en réacteur piston, nous avons choisi de ne représenter que la répartition des échantillons dans l'espace des conditions expérimentales (Figure 4.27).

Tableau 4.5 : Conditions expérimentales préliminaires et des itérations jusqu'à convergence.

	Technologie [–]	Temps [min]	Ratio en T [%]	Coulée [mL/min]
<i>Préliminaire</i>	Batch	5 – 10 – 15 – 20 25 – 30 – 35 – 40	50	aucun
<i>Itération 1</i>	Piston	3,75 – 8,5 – 4,5 – 4,75 8,5 – 4,25 – 4,5 – 4,5	45 – 30 – 60 – 55 30 – 55 – 60 – 55	aucun
<i>Itération 2</i>	Semi-batch	5 – 25 – 30 – 35 40 – 45 – 50 – 55	40	4 paliers (cf. Figure 4.28)
<i>Validation</i>	Piston	6,5	55	aucun

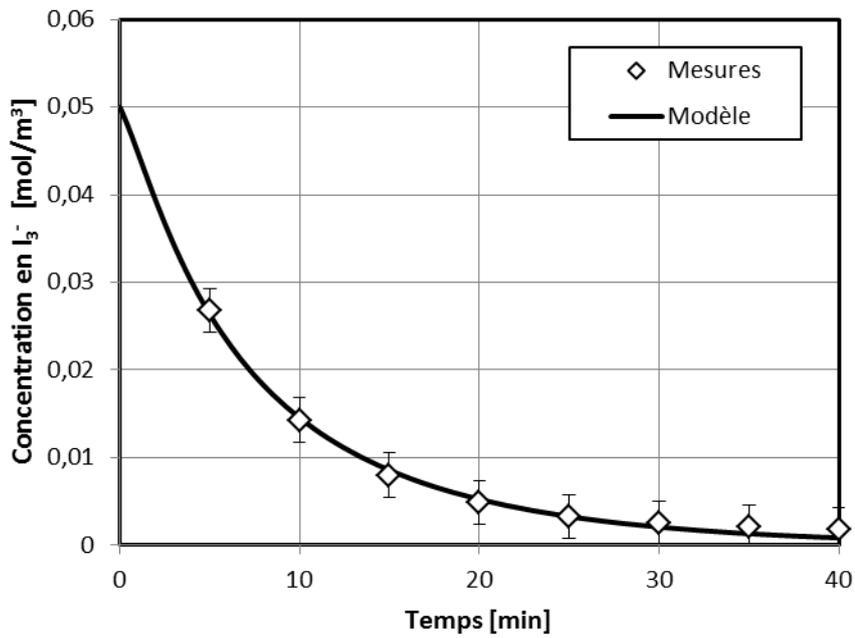


Figure 4.26 : Régression paramétrique de l'expérience préliminaire.

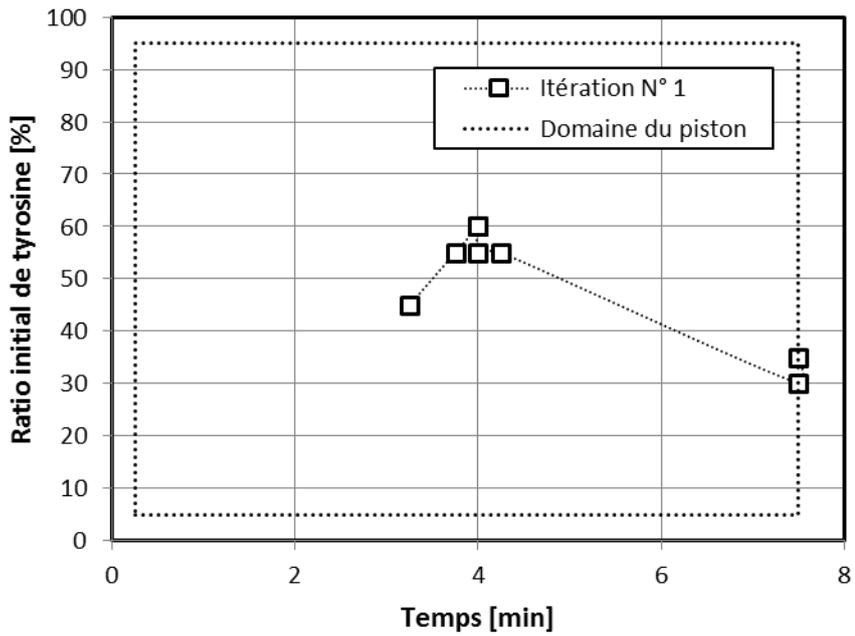


Figure 4.27 : Répartition des échantillons dans l'espace des conditions expérimentales du réacteur piston (itération 1).

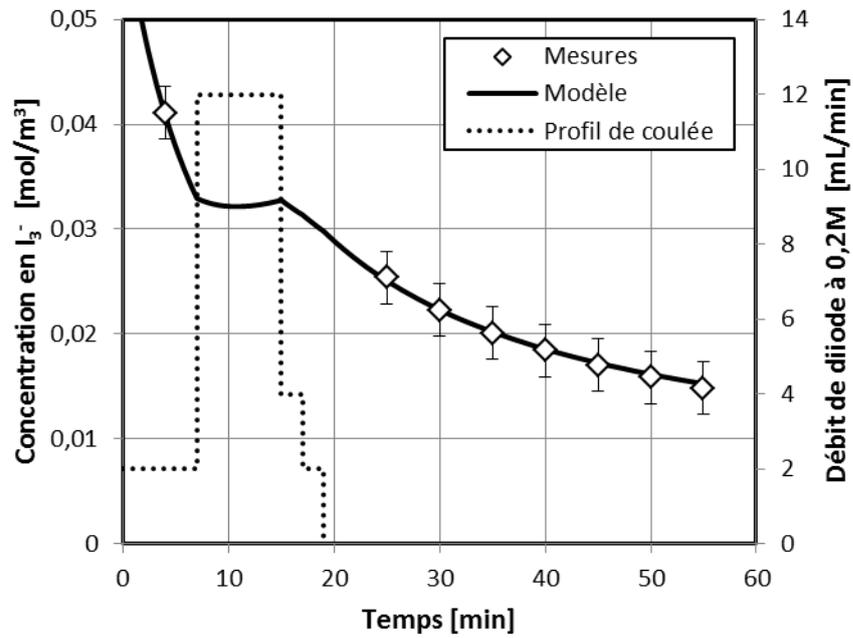


Figure 4.28 : Régression paramétrique de l'expérience D-optimale de l'itération 2.

La Figure 4.29 présente l'évolution de la région de confiance des paramètres au cours des deux itérations. On peut noter que la valeur estimée à l'itération 2 est proche de l'estimation indépendante des paramètres détaillée dans le Tableau 4.3.

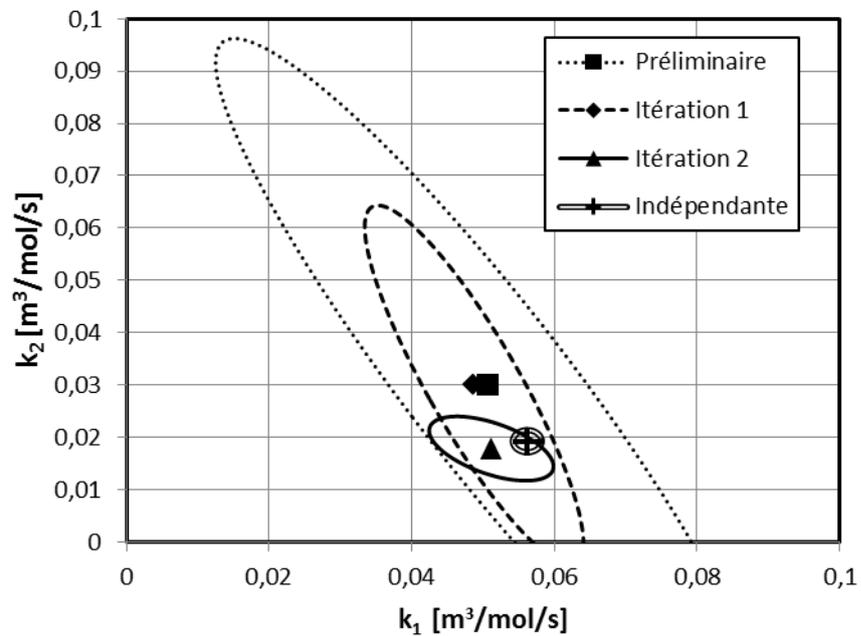


Figure 4.29 : Évolution de la région de confiance des paramètres au cours des itérations.

La Figure 4.30 montre la valeur du rendement optimal de 3-iodotyrosine prédit par le modèle à chaque itération. La précision de la prédiction décroît significativement jusqu'à être inférieure à la précision de la mesure du rendement à l'itération 2.

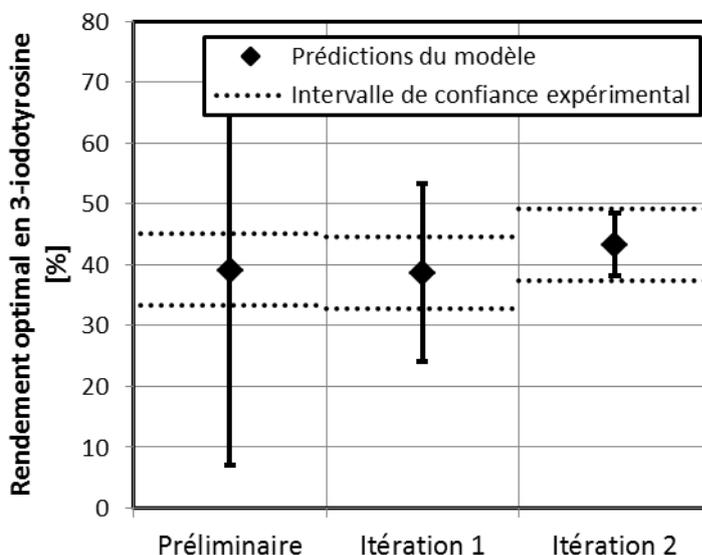


Figure 4.30 : Évolution de la prédiction du rendement optimal en 3-iodotyrosine au cours des itérations.

Au bout de la seconde itération, le modèle est jugé suffisamment précis pour passer à l'étape de validation. Nous avons donc fait une expérience dans les conditions optimales prédites par le modèle (le piston est trouvé comme étant la meilleure technologie avec $t = 6,5$ min et $R_T = 55\%$, voir plus haut le Tableau 4.5). Pour mesurer le rendement, nous devons mesurer la concentration en 3-iodotyrosine. Pour cela, nous utilisons la méthode décrite au paragraphe 4.1.2. Le mélange réactionnel a été « trempé » avec une solution de thiosulfate de sodium. On mesure alors le spectre UV de la solution décolorée puis on le décompose comme étant la superposition des trois molécules (tyrosine, 3-iodo-tyrosine et 3,5-diiodo-tyrosine) sous leur forme acide et basique (voir Figure 4.31).

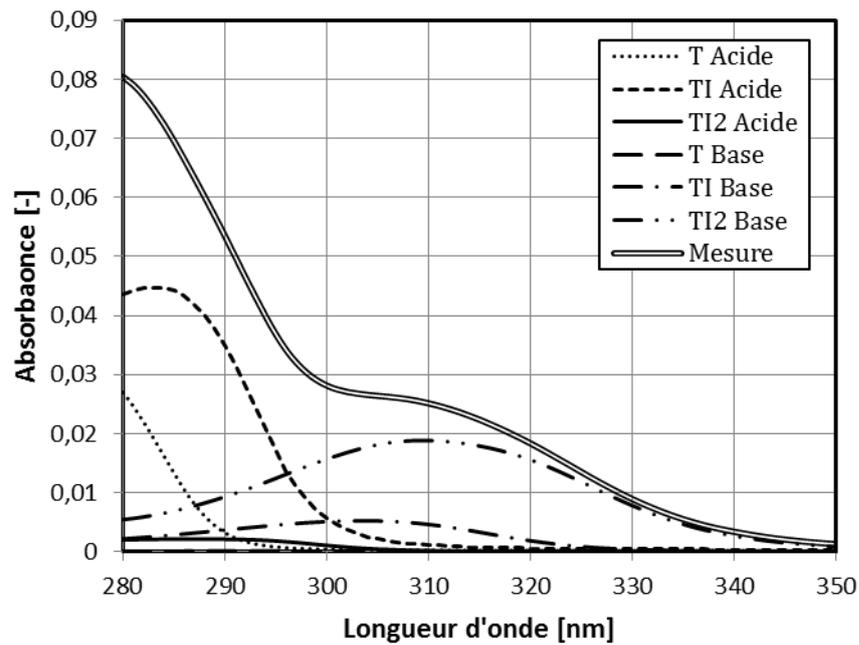


Figure 4.31 : Spectre du mélange réactionnel trempé obtenu lors de l'expérience de validation.

Les valeurs des concentrations ainsi mesurées sont données dans le Tableau 4.6 et comparées avec les valeurs prédites par le modèle. La mesure du rendement est en accord avec la prédiction du modèle. Ainsi, la technologie optimale d'un point de vue performance ainsi que les conditions expérimentales associées sont validées.

Par ailleurs, nous avons lancé l'optimisation de la performance en utilisant les valeurs des constantes obtenues de manière indépendantes (voir Tableau 4.6). Les résultats obtenus sont concordant : le rendement mesuré est dans l'intervalle de confiance du rendement prédit et le rendement estimé de manière indépendante est dans l'intervalle de confiance de la mesure.

Tableau 4.6 : Comparaison entre valeurs mesurées, valeurs prédites par la stratégie et valeurs obtenues par l'estimation indépendante des paramètres.

	$[T]$ [mol/m ³]	$[TI]$ [mol/m ³]	$[TI_2]$ [mol/m ³]	Rendement [%]
Mesurées	0,0292	0,0219	0,0038	42,8 ± 5,9
Prédites	0,0308	0,0216	0,0025	43,3 ± 5,1
Indépendantes	0,0298	0,0224	0,0026	44,9 ± 1,0

4.4 Conclusion

Le système expérimental de l'iodation de la tyrosine nous a permis de confirmer de manière expérimentale les résultats obtenus lors de l'application virtuelle de la stratégie (Chapitre 3). Cette synthèse possède l'avantage d'être simple à mettre en œuvre tant sur la manipulation des produits que sur les mesures analytiques. De plus, les molécules mises en jeu ne possèdent pas de toxicité particulière. Le système réactionnel est relativement complexe si l'on prend en compte les différents équilibres ioniques, mais en gardant des conditions expérimentales constantes dans les différentes expériences, nous avons pu simplifier le problème de manière à avoir un système réactionnel de type consécutif-parallèle.

L'utilisation du schéma réactionnel consécutif-parallèle a permis de montrer dans une première partie l'intérêt d'utiliser plusieurs technologies de réacteur pour rendre efficace l'acquisition de données cinétiques. Dans un second temps, nous avons appliqué la stratégie dans son ensemble ayant pour but de trouver la technologie et les conditions expérimentales associées pour maximiser le rendement en produit intermédiaire tout en minimisant le temps de synthèse. En seulement quatre expériences, (une préliminaire en batch, deux d'acquisition de donnée en piston puis en semi-batch et une dernière pour l'étape de validation en piston), la stratégie a convergé. Le résultat obtenu semble en accord avec l'estimation indépendante des conditions optimales.

Chapitre 5

Application à un cas industriel

Nous avons appliqué la stratégie Stanislas à un cas industriel, la synthèse de la cocopropylamine-bétaïne, plus couramment appelée cocobétaïne. Cette molécule entre dans la composition de nombreux cosmétiques, en particulier, la plupart des gels douches et shampooings. Les nombreuses applications cosmétiques de ce produit impliquent des spécifications très rigoureuses sur la composition du produit final, surtout sur les taux de réactifs résiduels. Industriellement, cette synthèse est réalisée dans un atelier fonctionnant en batch en un peu plus de seize heures. L'objectif est donc d'étudier différentes technologies pour réduire le temps de synthèse tout en respectant les spécificités sur la qualité du produit. Nous avons donc appliqué la stratégie Stanislas pour ce faire.

Avant d'appliquer la stratégie, nous détaillerons, dans une première section, la chimie du système réactionnel, les méthodes analytiques et les objectifs/contraintes de l'étude. Dans une seconde section, nous développerons l'application de la stratégie. Puis, dans une dernière section nous discuterons des résultats obtenus.

5.1 Description du cas industriel

5.1.1 La chimie

La cocobétaïne est un tensio-actif zwitterionique (qui possède des charges opposées au sein de la même molécule). Elle est utilisée dans la composition des gels douches et shampooings pour son puissant effet moussant. La partie lipophile de la cocobétaïne est composée d'acides gras issus de l'huile de noix de coco. Sa partie aquaphile est composée d'un ammonium quaternaire et d'un acide organique.

La synthèse de la cocobétaïne fait intervenir, en réalité, plusieurs étapes :

- L'extraction des acides gras de l'huile de noix de coco. De ce fait, la longueur des chaînes grasses varie d'une molécule à l'autre selon les proportions données dans le Tableau 5.1. Le glycérol obtenu comme coproduit de cette étape n'est pas séparé et restera présent jusque dans le produit final à hauteur d'une dizaine de pourcent massique.
- L'addition de la diméthylaminopropylamine. L'amine primaire de cette molécule réagit avec l'acide gras pour former une liai-

son amide. Cette synthèse constitue la première étape de la formation de la tête aquaphile du surfactant.

- Enfin, la formation du zwitterion. Elle est obtenue par l'alkylation d'acide monochloroacétique sur l'amine tertiaire située en bout de chaîne.

C'est sur cette dernière étape que se focalise notre étude. Le mélange obtenu à la fin de la synthèse est directement le produit commercial, il n'y a pas d'étape de séparation.

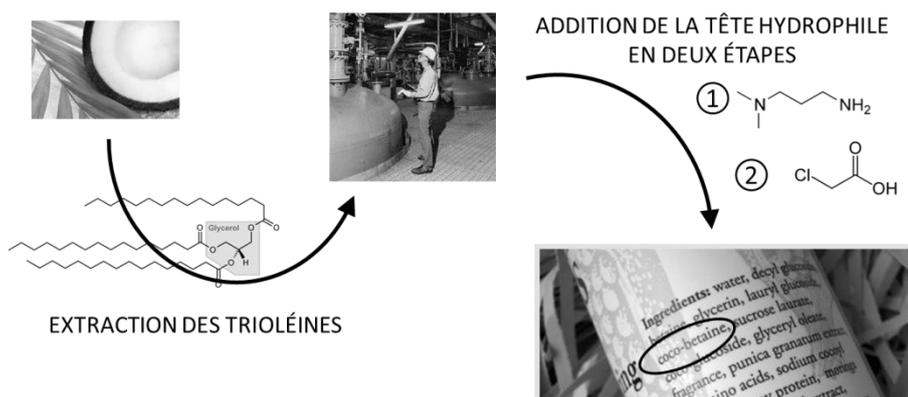


Figure 5.1 : Synthèse de la cocobétaine à partir de l'huile de noix de coco.

Le schéma réactionnel n'est pas connu, nous avons donc utilisé nos connaissances sur la réactivité des espèces chimiques pour élaborer un premier schéma réactionnel (voir section 5.2.2.5). Dans les paragraphes 5.1.1.1 et 5.1.1.2 nous détaillons successivement les molécules et le protocole de synthèse pour argumenter notre choix.

5.1.1.1 Les molécules

L'amido-amine (AA)

Cette molécule constitue la partie apolaire du produit final, elle est par conséquent insoluble dans l'eau. La chaîne grasse provient de l'huile de noix de coco, ainsi, cette molécule est en réalité un mélange de molécules dont la base majoritaire est l'acide laurique (voir Tableau 5.1). La diméthylaminopropylamine est greffée sur l'acide gras de manière à avoir une fonction amine tertiaire en bout de chaîne. Comme mentionné plus haut, le glycérol issu de la rupture de la trioléine en acide gras est présent dans le mélange à hauteur de 10,90% en masse. Le pKa de l'amine est d'environ 9,8. Ce produit est irritant pour la peau et les yeux. Jusqu'à aujourd'hui, il n'a jamais été mis en évidence que ce produit présentait des risques cancérigène, reprotoxique et mutagène pour l'homme. En revanche, plusieurs études lui ont montré un fort potentiel allergisant (Foti et al. 2003; Brey et Fowler Jr 2004).

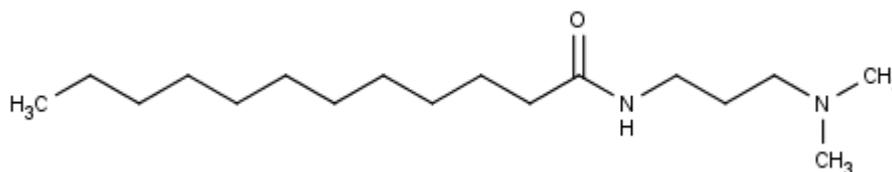


Figure 5.2 : Structure de la N-[3-(Diméthylamino)propyl]dodecanamide.

Tableau 5.1 : Composition en acide gras de l'huile de noix de coco.

Acide gras	Fraction massique [%]	Fraction molaire laire [%]	Masse molaire [g/mol]	Formule	(C: D n - x)
Laurique	47,6	48,9	183,3	$C_{12}H_{24}O_2$	(12: 0)
Myristique	18,2	16,4	211,4	$C_{14}H_{28}O_2$	(14: 0)
Palmitique	8,8	7,1	239,4	$C_{16}H_{32}O_2$	(16: 0)
Caprylique	7,8	11,1	127,2	$C_8H_{16}O_2$	(8: 0)
Caprique	6,7	8,0	155,3	$C_{10}H_{20}O_2$	(10: 0)
Oléique	6,2	4,5	265,5	$C_{18}H_{34}O_2$	(18: 1 n - 9)
Stéarique	2,6	1,9	267,5	$C_{18}H_{36}O_2$	(18: 0)
Linoléique	1,6	1,2	263,4	$C_{18}H_{32}O_2$	(18: 2 n - 6)
Caproïque	0,5	0,9	99,2	$C_6H_{12}O_2$	(6: 0)

Le MCA

L'acide monochloroacétique est le second réactif utilisé pour cette synthèse. Cet acide possède un pKa de 2,86 ce qui fait de lui un acide faible. Pour cette synthèse, il est utilisé en solution aqueuse à 80% massique. Le MCA à 80% n'est pas miscible avec la LAPA. C'est un produit très toxique, particulièrement sous cette forme. En effet, en solution concentrée, le contact avec la peau provoque des brûlures dues à son acidité, mais aussi et surtout il diffuse facilement dans les tissus. Étant un agent alkylant, il perturbe les mécanismes de production d'énergie au niveau des cellules et modifie les équilibres ioniques du corps. Il a été reporté plusieurs cas de décès liés à l'exposition de ce produit sur une surface de peau supérieur à 6% de la surface du corps (International Programme on Chemical Safety (IPCS) 2012). Jusqu'à aujourd'hui, il n'a jamais été prouvé un risque de cancérogénicité lié à cette molécule.

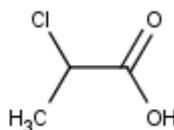


Figure 5.3 : Structure de l'acide monochloroacétique.

La soude

La soude joue le rôle de co-réactif dans cette synthèse, elle permet de maintenir un pH élevé tout au long de la synthèse. Elle est utilisée en solution à 50% massique. La soude ne possède pas de toxicité chronique, mais présente des risques de graves brûlures.



Figure 5.4 : Structure de la soude.

La cocobetaine

La molécule synthétisée est un mélange de surfactants zwitterioniques, de type bétaine, dont la longueur des chaînes grasses varie (même proportion molaire que dans l'amidoamine). Ces molécules sont considérées comme étant l'espèce active et doivent être présentes à environ 30% en masse dans le produit final. Cette molécule est faiblement toxique ($Oral LD_{50}^{Rat} > 2000 \text{ mg/kg}$) et n'est pas considérée comme étant cancérogène par les principales agences de sécurité et de santé.

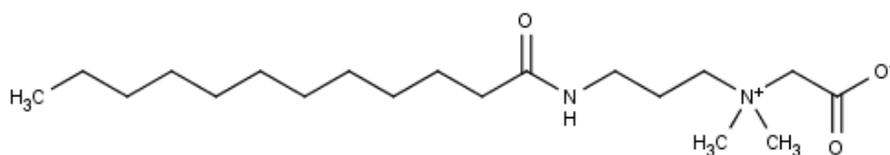


Figure 5.5 : Structure de l'acétate de 2-[(3-dodecanamidopropyl) diméthylazaniumyl].

Les sous-produits

Deux sous-produits identifiés sont l'acide glycolique et l'acide diglycolique. L'acide glycolique se forme suite la substitution du chlore du MCA par l'ion hydroxyde. L'acide diglycolique se forme dans un second temps, c'est le groupement alcool de l'acide glycolique qui se substitue au chlore du MCA. Leur production est favorisée par un pH fortement basique. C'est deux produits ne présentent pas de toxicité particulière si ce n'est leur acidité.

5.1.1.2 La synthèse actuelle

Industriellement, la synthèse étudiée se fait en trois étapes (voir Figure 5.6), le chargement, la réaction et le conditionnement. L'étape de chargement dure 3 heures, on commence par charger l'eau, puis l'aminoamine, le MCA et la soude (remplacée par de la potasse dans la synthèse similaire présentée sur la Figure 5.6). Les proportions molaires et massiques des différents réactifs sont données dans le Tableau 5.2. On remarquera qu'il y a un très léger excès de MCA par rapport à l'AA ($R_{MCA/AA} = 1,006$), de même qu'il y a un peu plus de NaOH que de MCA ($R_{NaOH/MCA} = 1,003$).

Tableau 5.2 : Composition d'entrée du procédé industriel.

Produits	Fraction massique [%]	Fraction molaire [%]
<i>EAU</i>	62,85	92,701
<i>AA</i>	23,66	2,200
<i>MCA</i>	7,87	2,214
<i>NAOH</i>	3,33	2,221
<i>GLYC</i>	2,33	0,674

Le temps de coulée est principalement limité par le débit des pompes. L'exothermie dégagée par la réaction acide-base entre la soude et le MCA est utilisée pour atteindre le premier palier de température de l'étape de réaction (180°F soit 85°C). À la fin de l'ajout de soude le milieu réactionnel atteint 11 unités de pH.

La réaction dure 5 heures, on fait un premier palier de température de 2 heures à 85°C, puis un second à 95°C. Entre les deux paliers, on ajoute à nouveau de la soude pour ramener le pH au-dessus de 10. Enfin un dernier palier d'une heure. À la fin de l'étape de réaction, un échantillon est prélevé pour valider sa conformité (QC pour « quality control » sur la figure). Si un des réactifs est en excès par rapport aux spécifications du produit, l'autre réactif est alors ajouté dans des quantités stœchiométriques, puis on attend 2 heures avant le prochain point de contrôle.

Lorsque les spécifications du produit sont jugées valides, la température est ramenée à 50°C et à $pH = 6,5 \pm 0,3$ (par des ajouts successifs d'acide chlorhydrique) avant d'être conditionné. Cette dernière étape dure 6 heures. Il est à noter que : les deux réactifs principaux ne sont pas miscible et le produit final est monophasique, nous avons donc une synthèse qui passe d'un système biphasique à un système monophasique.

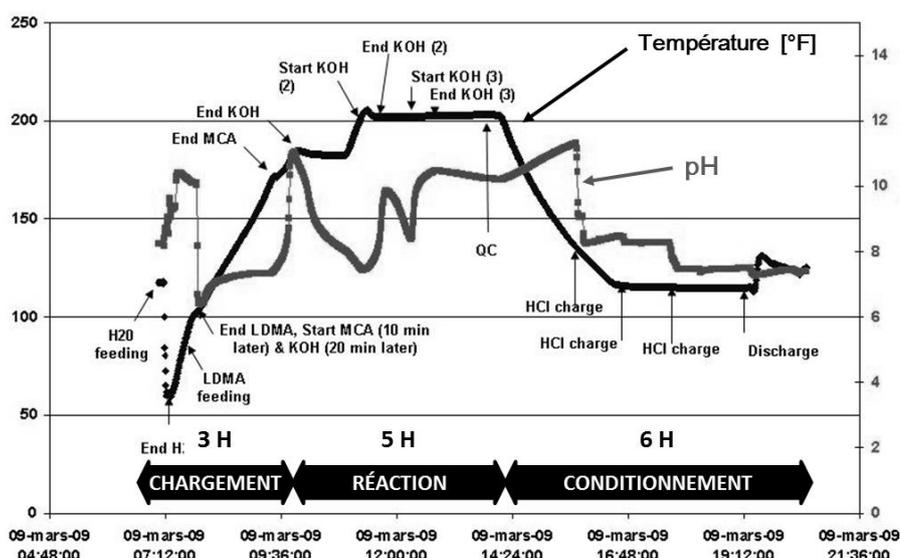


Figure 5.6 : Exemple de suivi de température et de pH lors d'une synthèse industrielle de laurylbétaine (similaire à la cocobétaine) à l'usine Winder – Novicare production – Georgia – USA (2009).

5.1.2 Objectifs et contraintes

L'objectif industriel est d'accroître la productivité volumique de la chaîne de production. L'idée étant de produire plus avec la même emprise au sol. Les solutions proposées doivent donc être limitées à des modifications mineures de l'unité de production actuelle.

La contrainte est de respecter la qualité du produit. C'est-à-dire qu'il doit impérativement répondre aux spécifications présentées dans le Tableau 5.3 :

Tableau 5.3 : Spécification requise pour la cocobétaine.

N°	Spécifications		Méthode d'analyse	Résultats attendus
1	MCA résiduel	[ppm]	HPLC	200 max
2	Amidoamine	[%]	HPLC	0,70 max
3	Apparence	[-]	Visuelle	Propre et limpide
4	Couleur	[ALPHA]	Colorimétrie	125 max
5	Chlorure de sodium	[%]	Potentiométrie	5,25 max
6	Taux de solide	[%]	Pesée	34,5 – 36,0
7	Actifs	[%]	Déduit de 5 et 6	29,0 – 31,0
8	pH (à 25°C)	[-]	pH-métrie	6,2 – 6,8

Les deux premières spécifications (MCA et amidoamine résiduelles) sont les plus importantes et les plus difficiles à respecter, puisqu'elles dépendent fortement de la manière dont est conduite la réaction. Le produit ne doit pas être chauffé trop long-

temps car une légère coloration intervient. Cette coloration est mesurée et ne doit pas dépasser 125 ALPHA. La concentration en chlorure de sodium et le taux de solide (c'est-à-dire la masse sèche) permettent de calculer la quantité d'actifs dans le produit final. Ces paramètres ne dépendent que des conditions opératoires (en particulier la quantité totale de soude et d'amidoamine introduite dans le milieu réactionnel). Enfin, le pH est ajusté par ajout d'acide chlorhydrique en fin de réaction.

Ainsi, le Tableau 5.4 donne les critères que nous avons choisis, leurs seuils vetos et objectifs. Le calcul de la fraction massique en chlorure de sodium est réalisé à partir de la quantité de sodium totale injectée dans le mélange réactionnel, elle ne dépend donc pas des mesures. La fraction massique en bétaine est estimée à partir de la quantité totale de LAPA injectée et de la quantité résiduelle.

Tableau 5.4 : Définition des critères pour la synthèse de la cocamido-propyl bétaine.

	Critère 1	Critère 2	Critère 3	Critère 4	Critère 5	
<i>Performance</i>	γ_i	t_f	ω_{MCA}	ω_{LAPA}	ω_{NaCl}	$ \omega_{BET} - 30 $
<i>Objectif</i>	o_i	0 s	0 ppm	0 %	0 %	0 %
<i>Veto</i>	v_i	1200 s	200 ppm	0,7 %	5,25 %	1 %

Pour augmenter la productivité, nous nous concentrons sur l'étape de réaction. Les modes d'intensification envisagés sont : le passage à un mode de production continue en cuve agitée ou en réacteur piston, l'augmentation de la température de réaction et enfin sur le mode d'introduction des réactifs (profil de coulée en semi-batch). Nous ne considérons pas, dans cette étude, l'usage d'un catalyseur pour accélérer la réaction, bien que cette solution ait déjà été mise en œuvre par Evonik (Herrwerth et al. 2013).

5.1.3 Méthodes analytiques

Comme nous l'avons vu au paragraphe 5.1.2, seul le dosage des réactifs résiduels est nécessaire pour mesurer les spécifications du produit commercial, les autres spécifications dépendant uniquement des quantités de réactif introduit dans le milieu. Ainsi, Solvay nous a proposé deux analyses HPLC distinctes pour doser les réactifs résiduels. Il est important de noter que le produit principal n'est pas dosé. Les chaînes HPLC utilisées sont de la marque Shimadzu.

5.1.3.1 Dosage de l'amidoamine résiduelle

Les amidoamines résiduelles sont dosées par HPLC sur une colonne à phase inverse, la Supelcosil LC-ABZ (250 x 4,6 mm – 5 μ m). Étant donné que les amidoamines sont un mélange de plusieurs molécules, nous ne dosons, en réalité, que la longueur de chaîne majoritaire, la chaîne en C12 appelée LAPA (pour LauryAmidoPropylAmine). Connaissant ses proportions dans le mélange de chaînes (48,9% en nombre), on peut re-

monter à la quantité totale d'amidoamines résiduelles. Mais cela suppose que la réactivité des différentes chaînes soit la même et donc que le ratio molaire en chaîne C12 est conservé tout au long de la réaction.

$$C_{AA} = \frac{1}{0,489} C_{LAPA} \quad (5.1)$$

L'échantillon est dilué 20 fois dans du méthanol avant d'être analysé. Cela permet de limiter la quantité de produit introduit dans la colonne (pour ne pas détériorer la durée de vie de la colonne). Le détecteur utilisé pour cette analyse est un détecteur évaporatif à diffusion de lumière, DEDL (ou en anglais ELSD pour Evaporative Light Scattering Detector). Ce détecteur à l'avantage d'être très sensible, universel et très peu influencé par les gradients de solvant. Or, la méthode choisie utilise justement un gradient de solvant. Le solvant principal utilisé est de l'eau à 0,1% d'acide trifluoroacétique (THF), le modificateur organique est du méthanol à 0,1% de THF. Initialement, nous avons implémenté le gradient proposé par la méthode Solvay, mais la résolution des pics n'était pas suffisante comme le montre les figures 5.7 et 5.8.

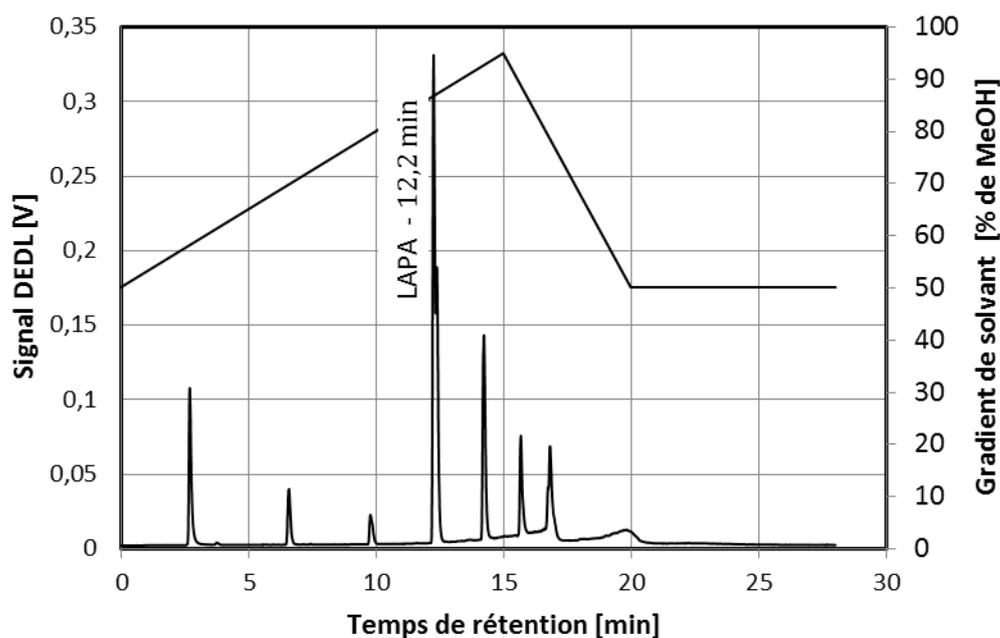


Figure 5.7 : Chromatogramme avec le gradient proposé par Solvay.

Sur la Figure 5.8, nous avons tracé le chromatogramme de trois échantillons pris aux temps de réactions 5, 10 et 15 minutes. Dans l'échantillon à 5 minutes, il reste d'avantage de LAPA que dans celui à 10 minutes et que dans l'échantillon à 15 minutes. Dans l'échantillon à 15 minutes, la quantité de produit est plus grande que dans celui à 10 et à 5 minutes. Or, on observe bien la disparition d'un pic et l'apparition d'un autre, mais les deux pics se chevauchent. Cela a pour effet de réduire la résolution de l'analyse.

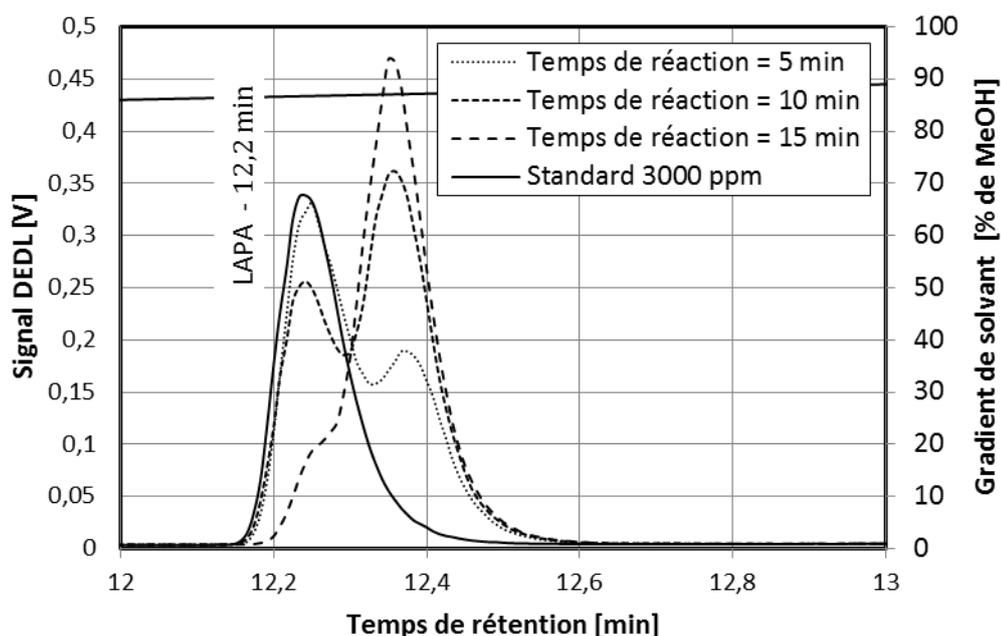


Figure 5.8 : Chromatogrammes d'échantillons pris à 5, 10 et 15 de temps réaction, réalisé avec le gradient proposé par Solvay, zoom sur l'épaulement du pic de la LAPA.

Nous avons donc modifié le gradient pour améliorer la résolution. L'idée est de faire un gradient plus doux pour mieux séparer les molécules. On part donc avec une concentration en modificateur organique très faible (5% de MeOH), puis on atteint 95% de méthanol au bout de 20 minutes (au lieu de 15 précédemment). On effectue un palier de 2 minutes pour rincer la colonne. Puis, on revient à 5% de méthanol pour stabiliser la colonne avant de passer un nouvel échantillon. Comme le montre les Figure 5.9 et Figure 5.10. En procédant ainsi, nous avons pu séparer les pics.

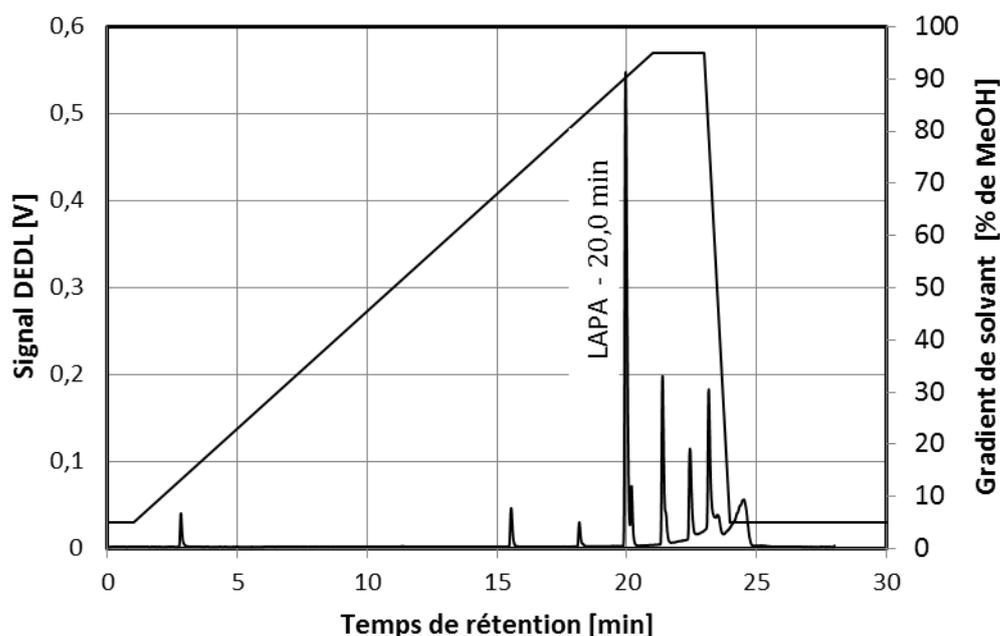


Figure 5.9 : Chromatogramme avec le gradient modifié.

Sur la Figure 5.10, où sont représentés les chromatogrammes des mêmes échantillons mais obtenu avec le gradient modifié, on observe que les deux pics (LAPA et produit) ne se chevauchent plus. Nous avons donc utilisé ce gradient pour toutes les mesures de la concentration en LAPA.

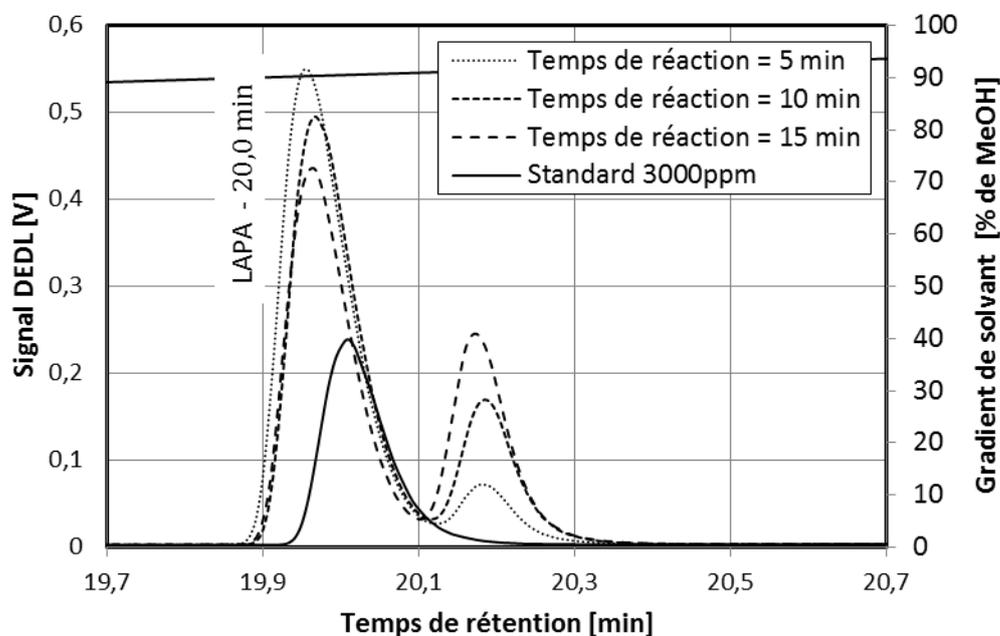


Figure 5.10 : Chromatogrammes d'échantillons pris à 5, 10 et 15 de temps réaction, réalisé avec le gradient modifié, zoom sur le pic de la LAPA.

La calibration se fait au moyen d'une loi de type puissance, $y = bx^m$ (ou $\ln y = \ln b + m x$), comme le montre la Figure 5.11. Pour faire la calibration, nous avons

utilisé de la LAPA à 97% de pureté achetée chez Sigma-Aldrich. À chaque série d'analyses, une nouvelle gamme d'échantillonnage est réalisée.

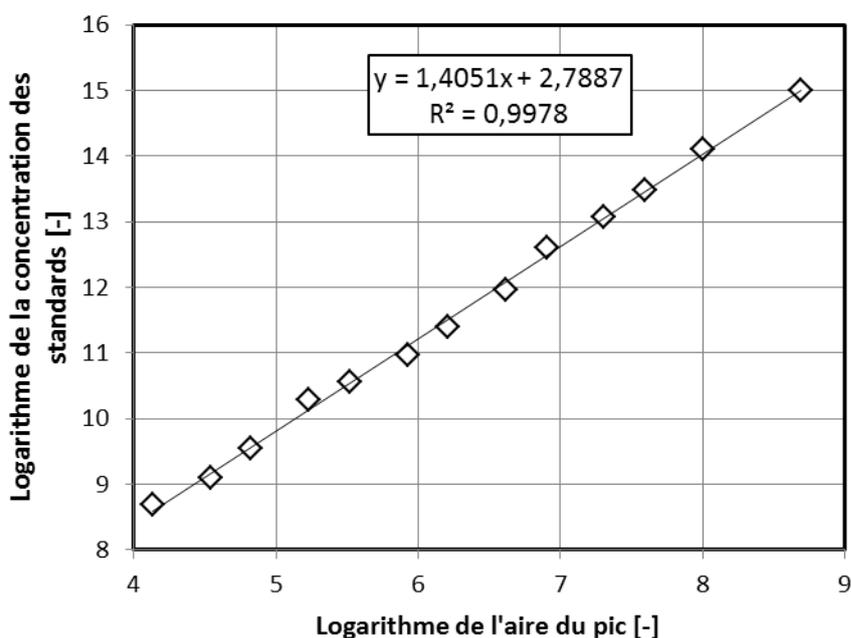


Figure 5.11 : Calibration du dosage des amines résiduelles.

Par ailleurs, nous avons observé une diminution de la résolution au fur et à mesure du nombre d'échantillons analysés. Pour remédier à cela, nous avons instauré après chaque série d'analyses un lavage de la colonne comme le montre le chromatogramme de la Figure 5.12, où l'on observe un important relargage de la colonne qui diminue très fortement après trois cycles. Nous n'avons pas de spectroscopie de masse sur la chaîne pour nous donner des informations sur la nature des composés accumulés dans la colonne.

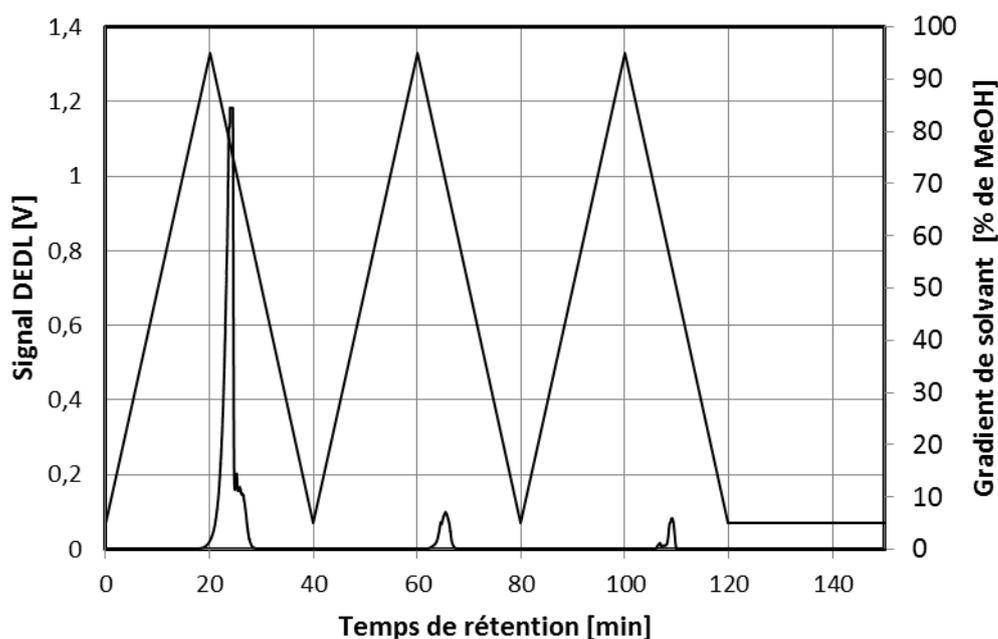


Figure 5.12 : Chromatogramme obtenu lors d'un cycle de lavage de la colonne.

5.1.3.2 Dosage des acides organiques

Le MCA résiduel est dosé en même temps que l'acide glycolique et que l'acide diglycolique par HPLC sur une colonne à échange d'ions (BioRad Aminex HPX-87). Pour cette analyse, nous utilisons de l'eau acidifiée à 0,007 N de H_2SO_4 en isocratique (i.e. : pas de gradient). Le détecteur que nous utilisons est un détecteur à barrette de diode nous permettant d'acquérir le spectre complet à chaque instant. Cela est très utile pour vérifier la pureté des pics et s'assurer qu'il n'y a pas d'épaulement. Cependant, pour la quantification, nous n'utilisons que la longueur d'onde à 210 nm. Les échantillons de calibration contiennent les trois acides (monochloroacétique, glycolique et diglycolique). La Figure 5.13 montre le chromatogramme obtenu pour le standard à 2000 ppm et la Figure 5.14 les droites de calibrations de ces trois acides.

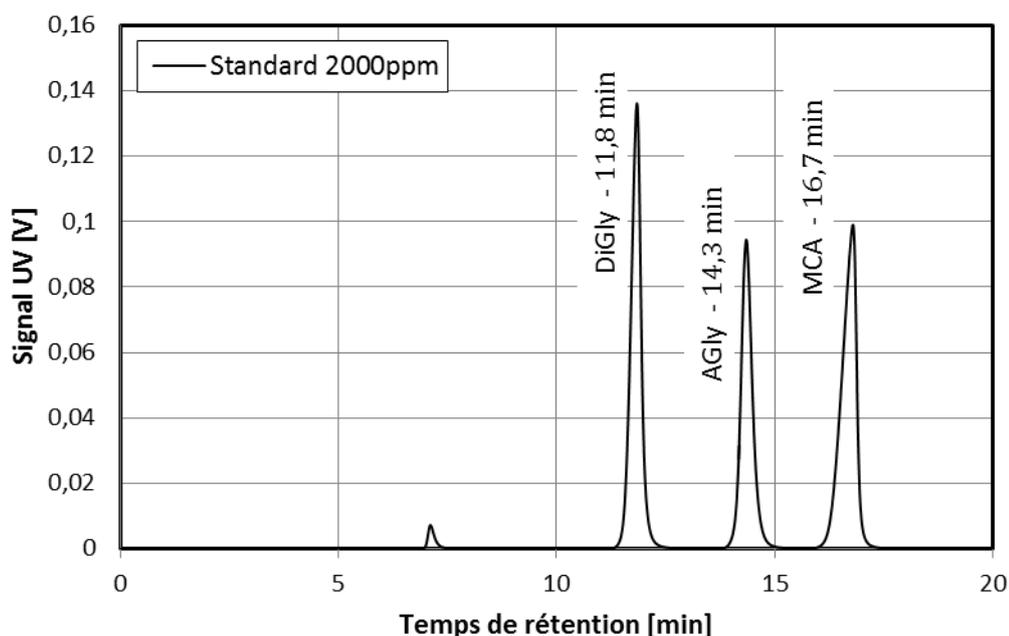


Figure 5.13 : Chromatogramme d'un standard d'acide monochloroacétique, glycolique et diglycolique à 2000ppm.

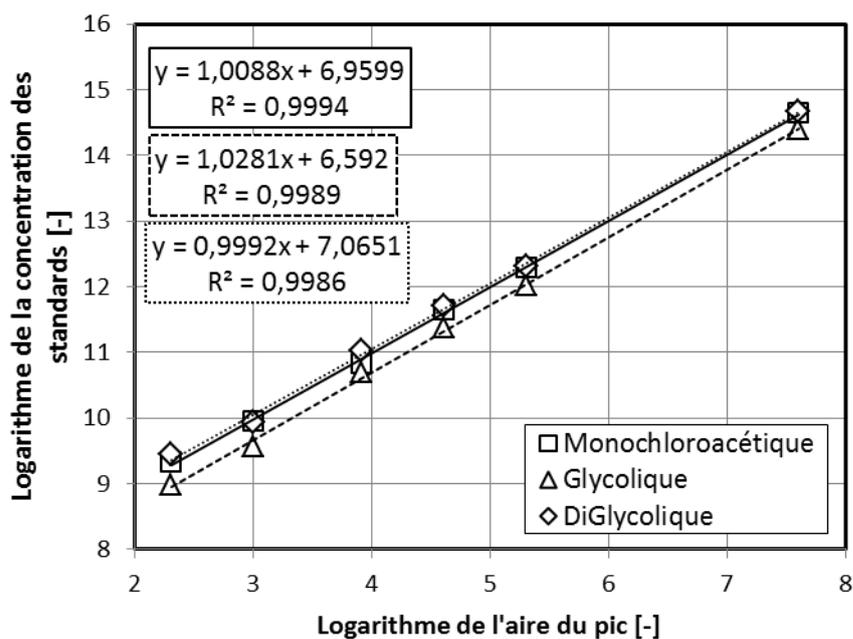


Figure 5.14 : Droites de calibration des acides monochloroacétique, glycolique et diglycolique.

Dans la Figure 5.15, nous avons tracé les chromatogrammes des mêmes échantillons que ceux utilisés dans le paragraphe 5.1.3.1 pour le dosage de l'amidoamine. On observe bien une décroissance du pic du MCA correspondant à sa consommation par la réaction principale. On note une très légère bosse à 14,3 minutes correspondant à des traces d'acide glycolique. Le pic à 7,4 minutes n'a pas été identifié, mais il doit très probablement correspondre à la bétaine.

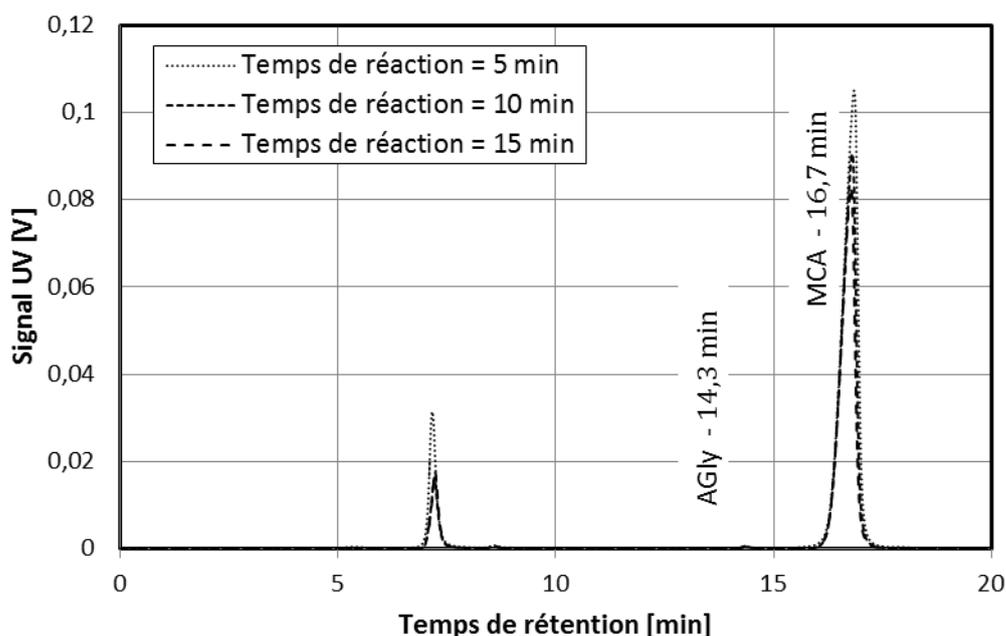


Figure 5.15 : Chromatogrammes d'échantillons pris à 5, 10 et 15 de temps réaction (les même que dans la Figure 5.10).

5.2 Application de la stratégie

À ce point, nous avons décrit les molécules qui jouent un rôle dans la synthèse qui nous intéresse, les méthodes analytiques utilisées, le protocole industriel de synthèse et les objectifs de Solvay. Nous pouvons donc appliquer la stratégie Stanislas. Nous commencerons par l'étape d'initialisation qui consiste en particulier à définir les technologies prises en compte et le modèle phénoménologique du système. Ensuite, nous détaillerons les itérations successives de la stratégie, puis nous conclurons.

5.2.1 Domaines expérimentaux

Pour cette étude nous avons à disposition les deux technologies du banc d'essai présenté au paragraphe 2.1 (un réacteur agité et un réacteur piston). Nous n'avons pas considéré le réacteur agité dans la configuration batch. Ainsi, trois réacteurs sont disponibles : un réacteur semi-batch, un réacteur agité continu (noté RPAC) et un réacteur piston. Leur domaine des conditions expérimentales sont définies dans le Tableau 5.5.

Le nombre de flux d'entrées est limité par le nombre de pompes, c'est-à-dire 2. Or, dans le procédé industriel quatre flux sont utilisés : l'eau, l'amidoamine, le MCA à 80% et la soude 50%. Il a donc été décidé de pré-mélanger, pour les deux technologies continues (RPAC et piston), l'amidoamine et la soude d'un côté et le MCA avec l'eau de l'autre. Ainsi, au cours d'une même expérience, on ne pourra pas faire varier le débit de soude indépendamment du débit d'amidoamine. Pour le semi-batch, l'eau et l'amidoamine sont chargés avant le début de l'expérience tandis que le MCA et la soude

sont coulés pendant. Par ailleurs, la stabilité (3 fois le temps de séjour) du réacteur piston étant d'au moins une demi-heure nous ne pouvons générer que quatre échantillons par expérience contre 7 en semi-batch. En cuve agitée continue, le temps de stabilisation est supérieur à 6 heures, ce qui ne laisse la possibilité de générer qu'un seul échantillon par expérience.

Enfin, la soude à 50% en masse est proche de la saturation ce qui présente l'inconvénient majeur d'avoir des cristaux en suspension. En effet, la présence de ces particules peut endommager les pistons des pompes utilisés. Pour éviter ce risque, nous avons travaillé avec une solution de soude à 20%.

Tableau 5.5 : Domaine expérimental disponible.

Semi-batch	t	T^0	M_{EAU}	M_{AA}	F_{NaOH}	F_{MCA}
Domaine	3 - 120	70 - 130	50 - 300	50 - 300	2 paliers	2 paliers
RPAC	t	T^0	w_{EAU}	w_{AA}	w_{NaOH}	w_{MCA}
Domaine	30 - 120	70 - 130	0 - 1	0 - 1	0 - 1	0 - 1
Piston	t	T^0	w_{EAU}	w_{AA}	w_{NaOH}	w_{MCA}
Domaine	0,5 - 7,5	70 - 130	0 - 1	0 - 1	0 - 1	0 - 1

5.2.2 Détermination du modèle

Pour construire le modèle nous avons utilisé la méthode basée sur l'analyse des temps caractéristiques détaillée au Chapitre 2 paragraphe 2.1.2.1. D'après toutes nos connaissances a priori, nous pouvons supposer qu'il y a potentiellement une influence du transfert à l'interface liquide-liquide au moins au début de la réaction, une influence du pH à la vue des espèces mises en jeu et une influence de la température sur la cinétique de réaction. Pour pouvoir appliquer le protocole d'initialisation, nous avons besoin d'utiliser une expérience de référence. Cette expérience de référence est l'adaptation du protocole de synthèse industriel à l'échelle du réacteur de laboratoire (en quelque sorte un « *scale-down* »). Nous avons donc conservé les proportions de chaque réactif. L'introduction de l'eau et de la LAPA se font sans coulée, tandis que le MCA et NaOH sont coulés simultanément dans le réacteur en 4 minutes. Nous ne faisons pas de palier de température, la double enveloppe est maintenue tout au long de la synthèse à 95°C. La vitesse d'agitation est calculée de manière à obtenir une énergie dissipée similaire à celle du procédé industriel.

Tableau 5.6 : Scale-in du mélange de la cuve agitée.

Échelle		Industriel	Laboratoire
Énergie dissipée	[W/kg]	0,19	0,21
Nombre de puissance	[-]	4	4
Vitesse d'agitation	[1/s]	0,83	10
Diamètre agitateur	[m]	1,33	34×10^{-3}
Volume réacteur	[m ³]	50	$0,8 \times 10^{-3}$

La Figure 5.16 montre l'évolution des concentrations mesurées (AA, MCA et Agly) pour cette expérience de référence.

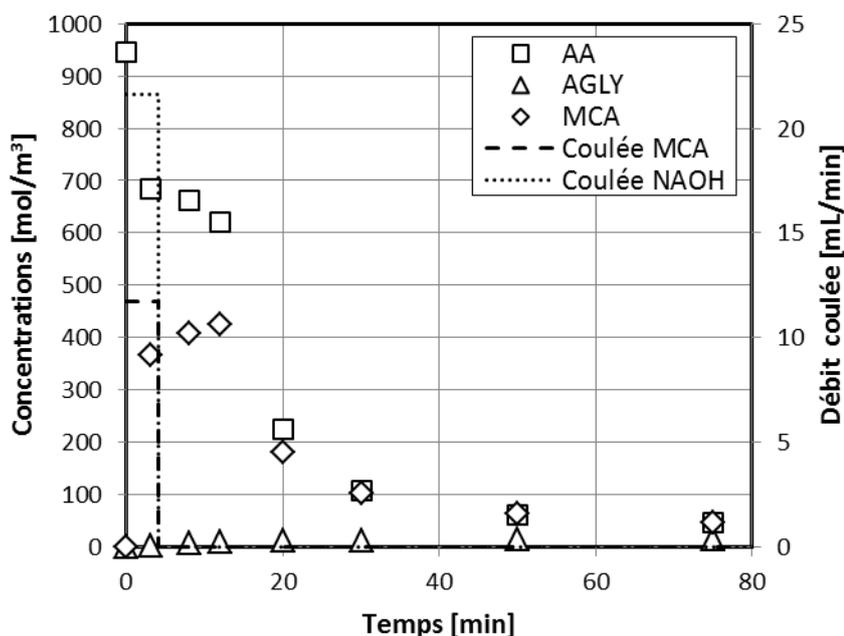


Figure 5.16 : Synthèse de la cocamidopropylbétaine dans les conditions de références.

5.2.2.1 Influence du transfert à l'interface

Le système réactionnel est biphasique, tout du moins au début de la synthèse. Il peut donc potentiellement avoir une limitation de transfert de matière à l'interface. Or, d'après le Tableau 2.4 page 73, pour faire varier ce temps caractéristique, on peut jouer soit sur le diamètre de l'agitateur, soit sur sa vitesse de rotation. Nous avons choisi de faire varier la vitesse d'agitation. Ainsi, nous avons réalisé la même expérience que l'expérience de référence mais en divisant par 10 la vitesse.

La Figure 5.17 montre l'évolution de la concentration en AmidoAmine en ayant divisée par 10 la vitesse d'agitation (le MCA n'a pas pu être analysé pour cette expérience). On observe clairement que la conversion est fortement impactée par la faible

agitation. Par ailleurs, en observant de plus près la conversion de l'expérience de référence (voir Figure 5.16), on s'aperçoit que la courbe de conversion subit une sorte de cassure aux alentours de 15 minutes : la réaction semble s'accélérer. Nous pouvons interpréter ces deux points en supposant que le système devient monophasique à partir d'une certaine concentration en bétaine dans le milieu réactionnel. Ainsi, dans la première phase de réaction, le système est limité par un transfert à l'interface, puis lorsque la concentration en bétaine est suffisante, le système devient monophasique et seule la cinétique décrit l'évolution du système.

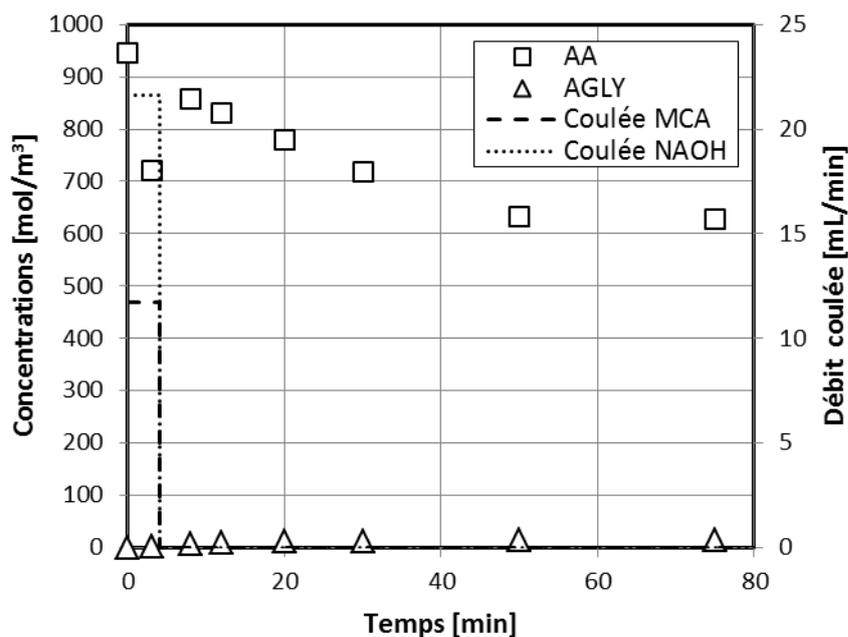


Figure 5.17 : Synthèse de la cocamidopropylbétaine à 90°C sous très faible agitation.

5.2.2.2 Influence du pH

Il semble que le pH joue un rôle prépondérant dans la cinétique de réaction. En effet, dans le protocole industriel, de la soude est ajoutée après 3 et 5 heures de réaction afin que la réaction puisse se faire préférentiellement basique. En regardant de plus près la nature chimique des deux réactifs principaux (le MCA et l'AA), on s'aperçoit que seule leur forme basique peut réagir pour former la bétaine (voir Figure 5.18).

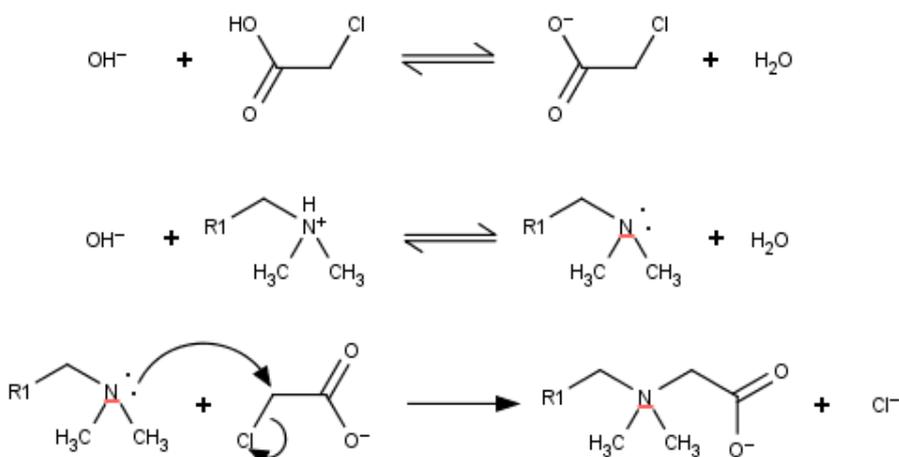


Figure 5.18 : équilibres ioniques mise en jeu dans la synthèse.

La pKa du MCA est de 2,86 donc la disponibilité de sa forme basique est largement majoritaire pour un pH supérieur à 5, or le pH au cours de la réaction est compris entre 6 et 12 unités (voir le suivi de température et de pH d'une synthèse industrielle Figure 5.6 page 170). En revanche, la forme basique de l'AA n'est pas forcément majoritaire tout au long de la réaction car son pKa est de 9,8. Ainsi, lorsque le pH devient trop bas, la réaction est fortement ralentie. D'ailleurs, on utilise ce principe pour tremper les échantillons avant analyse, ainsi, à chaque prélèvement, on ajoute de l'acide chlorhydrique concentré pour que le pH de l'échantillon soit inférieur à 3.

Le pH intervient aussi sur la réaction secondaire. En effet, la dégradation du MCA en acide glycolique se fait en milieu basique, car il s'agit de la substitution nucléophile du chlore par un ion hydroxyde (voir Figure 5.19).



Figure 5.19 : Formation du sous-produit.

Enfin, le pH joue un rôle sur la solubilité de l'AA, puisque sa forme basique est insoluble dans l'eau alors que sa forme acide l'est. Sur la Figure 5.20, nous avons tracé le coefficient de partage eau-octanol en fonction du pH de l'amidoamine, grâce au logiciel MarvinSketch 6.0.0 qui utilise une méthode de contribution de groupe décrite par (Csizmadia et al. 1997). On observe bien le fait qu'à pH supérieur à 10 l'AA est insoluble dans l'eau (logD = 3,97) et qu'à environ pH = 5, l'AA devient soluble, ce qui se traduit par le passage d'un système biphasique à un système monophasique.

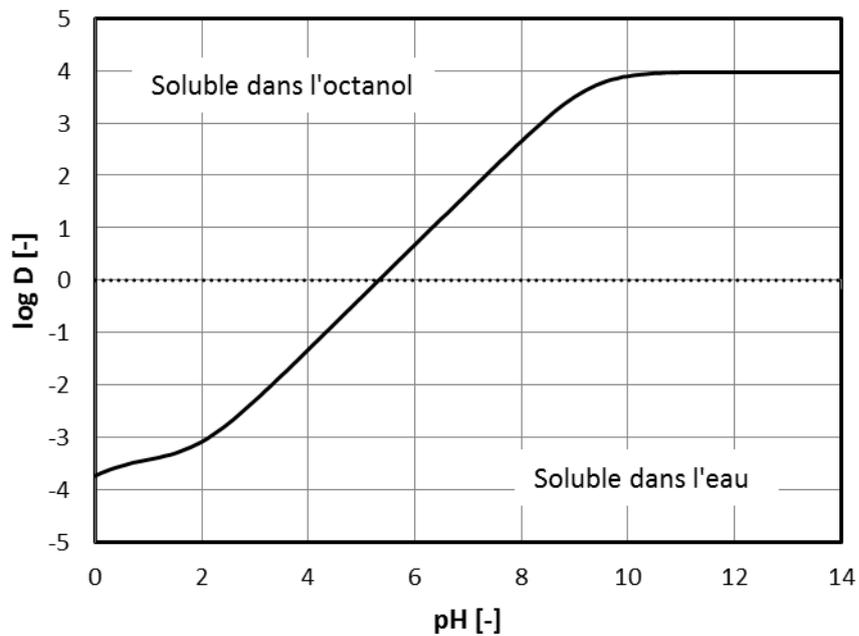


Figure 5.20 : Coefficient de partage de l'amidoamine entre l'eau et l'octanol.

5.2.2.3 Influence de la température

Dans le protocole industriel, la réaction se fait en deux paliers. La température joue donc vraisemblablement un rôle déterminant dans la conduite de la réaction. Afin de nous en assurer, nous avons réalisé deux essais complémentaires dans les mêmes conditions que l'essai de référence mais à 70°C (Figure 5.21) et à 120°C (Figure 5.22) au lieu de 95°C.

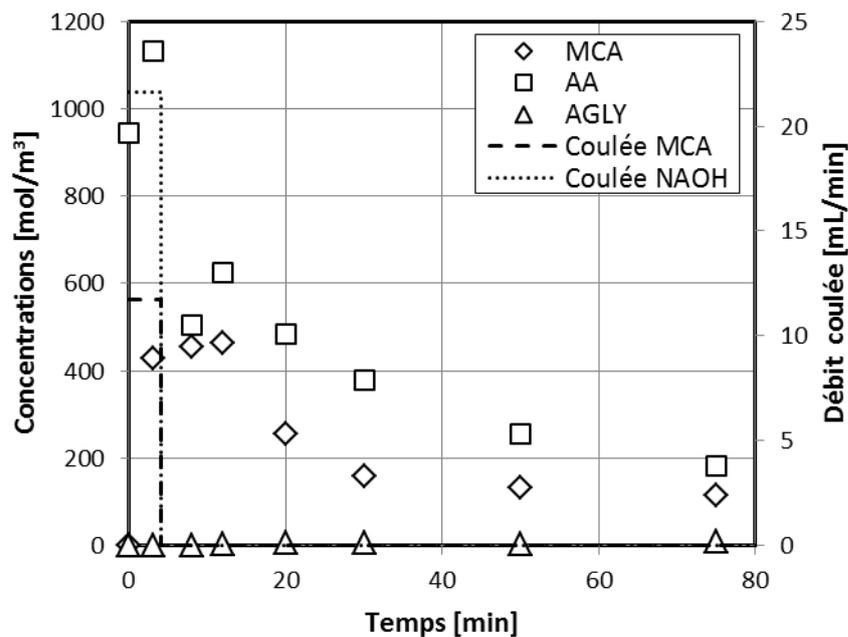


Figure 5.21 : Synthèse de la cocamidopropylbétaine à 70°C.

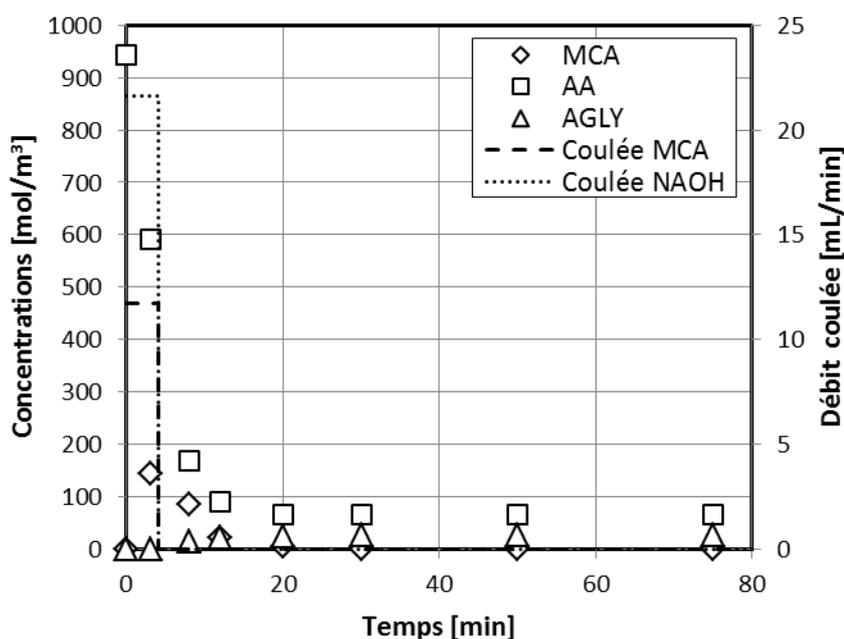


Figure 5.22 : Synthèse de la cocamidopropylbétaine à 120°C.

Clairement, la température accélère la réaction. Le temps caractéristique de synthèse (temps nécessaire pour atteindre 63% de conversion), passe d'une trentaine de minutes pour 70°C à moins de 5 pour 120°C. Il sera donc important de prendre en compte l'effet de température dans le modèle.

5.2.2.4 Influence de l'exothermicité

Solvay nous a fourni des résultats obtenus en DSC (Differential Screening Calorimetry). Ces essais nous fournissent des informations importantes sur les aspects thermiques. Cela nous donne accès aux chaleurs de réactions ($-42,2$ kJ/mol pour la réaction entre la soude et le MCA et $-77,1$ kJ/mol pour la réaction principale) et à la capacité calorifique du mélange réactionnel ($C_p = 500$ J/kg/K). D'autre part, nous observons que le produit final est stable jusqu'à 175°C, au-delà un dégagement gazeux apparaît.

5.2.2.5 Écriture du modèle

L'écriture du modèle doit être un compromis entre son niveau de détail et sa simplicité. Plus un modèle est complet, plus il est prédictif, mais plus le coût lié à l'acquisition de données est important. L'écriture d'un modèle capable de rendre compte des conditions de solubilité de l'amidoamine dans l'eau (pH, température, concentrations des différentes espèces), du phénomène de transfert de l'amidoamine dans l'eau, des équilibres ioniques et de la thermique du système est un travail ambitieux qui nécessite une excellente connaissance des phénomènes. Par ailleurs, un tel modèle comporterait un nombre très important de paramètres nécessitant un important effort d'identification, c'est pourquoi nous avons opté pour un modèle de compromis basé sur la simplification présentée ci-dessous.

Le premier constat que l'on fait porte sur l'influence de l'interface liquide-liquide. En comparant la Figure 5.16 (réaction à 90°C) et la Figure 5.22 (réaction à 120°C) on s'aperçoit que la cassure de pente liée au passage d'un système biphasique à un système monophasique n'est plus perceptible à 120°C. L'idée est donc de représenter le ralentissement de la réaction due aux transferts de matière par l'énergie d'activation de la réaction principale. Tout se passe comme si on avait un système homogène dès le départ et que l'on observait une réaction apparente.

Le deuxième point est la simplification des équilibres ioniques. En effet, la résolution dynamique d'un système comportant des équilibres ioniques couplés à des réactions irréversibles est assez lourd en soit. De plus, un modèle où la description des équilibres ioniques est détaillée sera probablement peu pertinent vu les niveaux de concentrations des espèces en solution et des températures atteintes. Dans la réaction principale seule la forme basique de l'amidoamine peut réagir. Or, le pH étant toujours aux alentours du pKa de l'amidoamine ($pK_{a_{AA}} = 9,8$) on approxime la concentration d'aminoamine sous forme basique par :

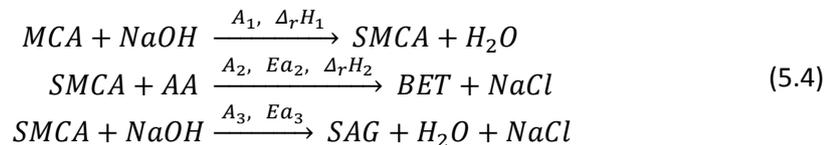
$$[AA^-] = [AA] \frac{[OH^-]}{10^{-pK_{b_{AA}}} + [OH^-]} \quad (5.2)$$

$$\text{Avec } pK_b = pK_e - pK_a$$

La réaction entre le MCA et la soude est totale, très rapide et exothermique. Sachant qu'elle est de toute façon plus rapide que les autres réactions, on la rend indépendante de la température ($E_{a_1} = 0 \text{ J/mol}$). De plus, on fixe arbitrairement les paramètres de cette réaction de telle sorte que le temps caractéristique de réaction soit proche de la seconde dans les concentrations de l'expérience de référence :

$$A_1 = \frac{1}{t_{car} \times C_{TOT}^0} = \frac{1}{1 \times 1800} \approx 5 \times 10^{-4} \text{ m}^3/\text{mol/s} \quad (5.3)$$

La réaction qui amène au sous-produit n'intervient que très faiblement, nous négligeons donc son exothermie. Une des spécifications sur le produit porte sur la quantité de chlorure de sodium, que nous faisons apparaître dans nos bilans de matières. Ainsi, la réaction entre le MCA et la soude produit de l'eau et du monochloroacétate de sodium (noté SMCA). C'est donc le SMCA qui réagit avec l'amidoamine pour former la bétaine et du NaCl. Au final, nous avons le schéma réactionnel suivant :



et les vitesses de réaction associées :

$$\begin{aligned}
 r_1 &= A_1 [NaOH][MCA] \\
 r_2 &= A_2 e^{(-Ea_2/RT)} [SMCA][AA] \frac{[NaOH]}{10^{pKb_{AA}} + [NaOH]} \\
 r_3 &= A_3 e^{(-Ea_3/RT)} [SMCA][NaOH]
 \end{aligned}
 \tag{5.5}$$

Ce modèle comporte donc quatre paramètres inconnus (A_2 , Ea_2 , A_3 et Ea_3), les enthalpies et le pKb de l'amidoamine étant connus de manière indépendante (voir plus haut). Pour des questions numériques, nous avons reparamétré les facteurs préexponentiels et les énergies d'activation comme présenté au paragraphe 1.1.4.3 (pour rappel, on cherche à estimer le logarithme des constantes de vitesse aux bornes du domaine de températures de travail).

Un autre point important du modèle est l'aspect thermique. Le bilan thermique utilisé est celui décrit au paragraphe 1.3.3. Les coefficients d'échange à la paroi utilisés pour la cuve agitée et pour le piston sont ceux donnés au paragraphe 2.2.2 ($U_{Cuve} = 300 \text{ W/m}^2/\text{K}$ et $U_{Piston} = 1220 \text{ W/m}^2/\text{K}$). Nous avons pris le parti de ne pas calculer les performances optimales en utilisant le ratio surface sur volume du réacteur industriel car nous n'aurions pas été en mesure de valider les prédictions dans notre réacteur de laboratoire. Enfin, à chaque fois que la stratégie propose une nouvelle expérience, nous vérifions que la limite haute de l'intervalle de confiance de la température maximale n'excède pas 150°C (soit la température limite de l'équipement) au cours de la réaction.

5.2.3 Déroulements de la stratégie

Lors de la campagne d'essais, pour des questions logistiques, nous ne pouvons réaliser que deux expériences par semaine. Nous avons donc choisi de réaliser chaque semaine, une expérience d'acquisition de données et une expérience de validation de performance bien que celle-ci n'était pas nécessaire d'après la stratégie. Ainsi, pour la régression paramétrique, nous utilisons sept échantillons issus des expériences d'acquisition de données (temps de prélèvement optimisés) et sept échantillons issus de l'expérience de validation (un prélèvement toute les 5 minutes). C'est donc une structure un peu différente de la stratégie qui a été utilisée, puisque l'on fait systématiquement l'étape de validation (voir Figure 5.23).

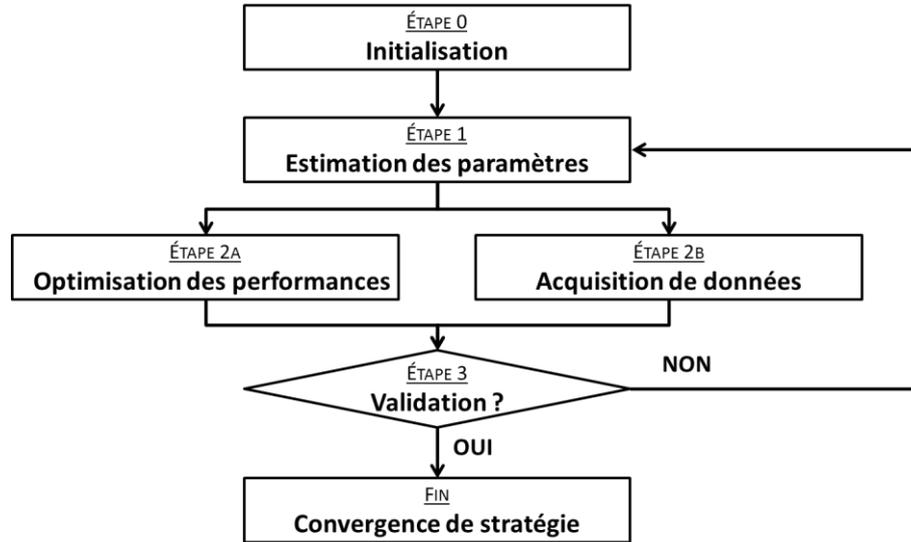


Figure 5.23 : Schéma de principe de la stratégie modifiée.

5.2.3.1 Expériences préliminaires

Nous n'avons pas besoin de refaire des expériences préliminaires, car nous avons déjà un jeu de trois expériences à trois températures différentes (70°C, 90°C et 120°C). La Figure 5.24 montre la régression paramétrique obtenue.

Mise à part pour l'expérience à 70°C, le modèle semble assez bien représenter les mesures expérimentales. Les valeurs estimées des paramètres pour chaque itération sont données dans le Tableau 5.7. Lorsque l'intervalle de confiance vaut une unité, cela veut dire que l'on a 95% de chance que la constante de vitesse (à T_{min} ou T_{max}) soit effectivement dix fois plus grande ou plus petite. En utilisant les paramètres estimés avec les expériences préliminaires nous avons déterminé des conditions optimales permettant de répondre aux critères définis au paragraphe 5.1.2 ainsi que l'expérience optimale d'un point de vue acquisition de données.

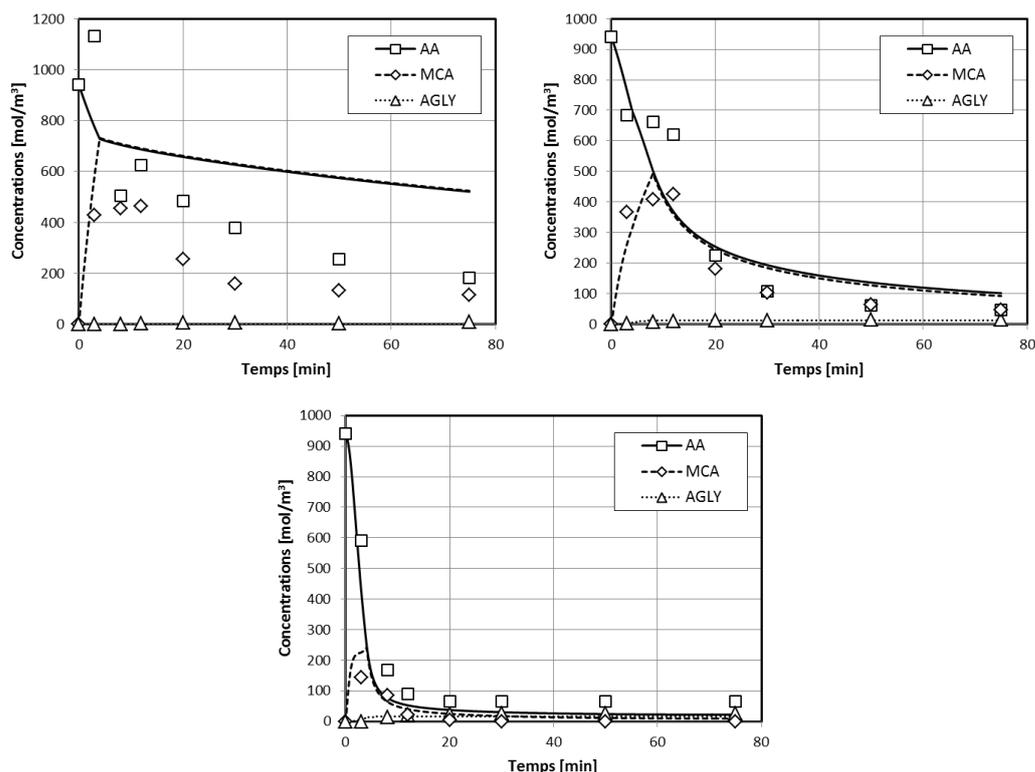


Figure 5.24 : Régression paramétrique de l'expérience à 70°C (en haut à gauche) à 90°C (en haut à droite) et à 120°C (en bas).

Tableau 5.7 : Estimation des paramètres et leur intervalles de confiance à chaque itération.

Itération n°	$\log_{10} k_1^{min}$	$\log_{10} k_1^{max}$	$\log_{10} k_2^{min}$	$\log_{10} k_2^{max}$
Préliminaire	$-6,23 \pm 0,78$	$-4,56 \pm 0,90$	$-8,70 \pm 1,41$	$-4,45 \pm 0,60$
1	$-6,78 \pm 1,11$	$-4,44 \pm 1,15$	$-7,08 \pm 2,22$	$-4,89 \pm 2,59$
2	$-6,08 \pm 1,34$	$-4,67 \pm 1,00$	$-7,01 \pm 1,24$	$-4,84 \pm 1,13$
3	$-6,40 \pm 1,52$	$-3,56 \pm 1,34$	$-7,01 \pm 0,86$	$-4,29 \pm 1,25$
4	$-6,40 \pm 1,52$	$-3,56 \pm 1,34$	$-7,01 \pm 0,86$	$-4,29 \pm 1,25$

Il est important de noter que la recherche de conditions optimales ainsi que la planification d'expériences se fait à l'aide de l'estimation des paramètres de l'itération précédente. Par exemple, pour l'itération n°1, ce sont les paramètres issus de l'estimation préliminaire qui ont été utilisés pour rechercher les conditions optimales.

5.2.3.2 Itération n°1

Lors de cette première recherche de conditions optimales, le modèle n'a pas réussi à trouver des conditions permettant de répondre aux cinq critères simultanément : la

quantité de soude dépasse 5,25 % et la concentration en bétaine est légèrement en dehors de la gamme admissible. Le Tableau 5.8 compare les valeurs prédites par le modèle et les mesures expérimentales. Ainsi, le modèle prédit assez bien la concentration d'aminoamines résiduelles mais n'e prédit pas la concentration en MCA, puisqu'il reste environ sept fois trop de MCA.

Tableau 5.8 : Comparaison entre performances prédites et mesurées à l'itération n°1.

Performance	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
Objectif	10	0	0	0	0
Veto	120	1,00	5,25	0,70	200
Prédictions	35	3,64	15,47	$0,18 \pm 0,83$	203 ± 434
Mesures	35	3,64	15,47	$0,00 \pm 0,24$	4335 ± 126

La validation a échoué et nous avons ré-estimé les paramètres (voir Tableau 5.7, ligne « itération n°1 ») en incluant les nouveaux points expérimentaux de l'expérience de validation et de l'expérience d'acquisition de données (voir Figure 5.25). On peut remarquer que l'intervalle de confiance a augmenté significativement pour tous les paramètres. Cela est probablement lié au fait qu'avec les nouveaux paramètres, le modèle est moins sensible, donc la matrice d'information est plus petite (voir équation (1.53) page 48), et l'intervalle de confiance des paramètres est plus grand.

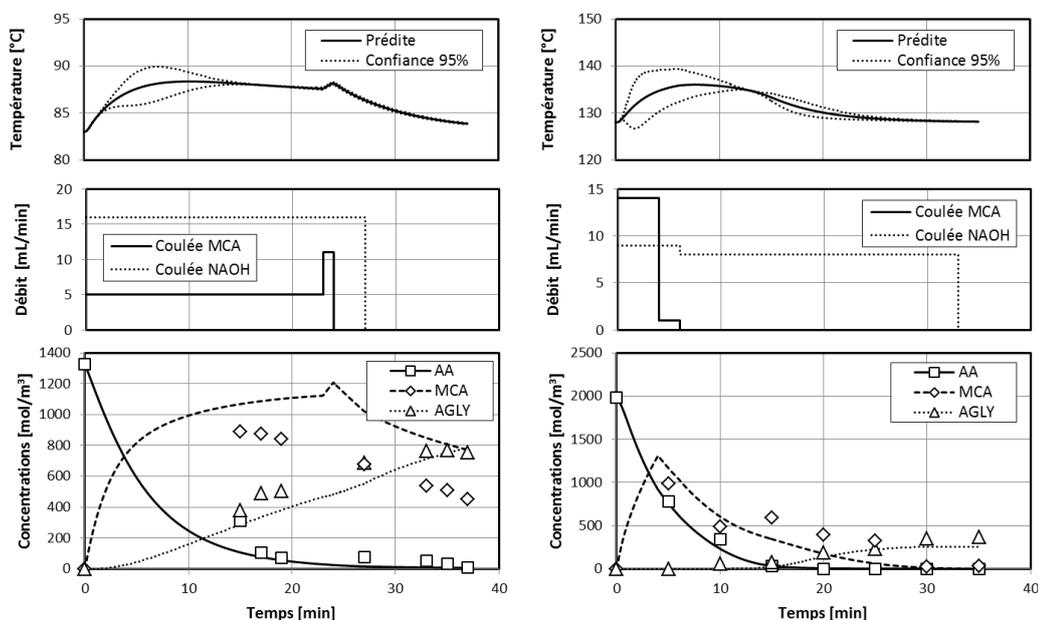


Figure 5.25 : Régression paramétrique de l'expérience d'acquisition de données (à gauche) et de validation (à droite) de l'itération n°1.

De l'expérience de validation (à droite), on peut remarquer que la stratégie propose de travailler à forte température, de maintenir un pH élevé tout au long de la réac-

tion en faisant une coulée de soude continue et de mettre un excès important en MCA ($R_{MCA/AA} = 1,336$). Mais cela a pour effet de favoriser la formation de sous-produits. De plus, on peut remarquer que la borne haute de l'intervalle de confiance de la température ne dépasse pas 150°C.

L'expérience d'acquisition de données propose une expérience plutôt à faible température, avec un très fort excès de MCA ($R_{MCA/AA} = 6,847$) et un fort excès de soude ($R_{NaOH/MCA} = 1,547$). Ces conditions favorisent la réaction secondaire et ainsi maximisent l'information cinétique de cette dernière. On peut remarquer un décalage significatif entre le modèle et les points expérimentaux pour le MCA. Cela est probablement lié au fait que le MCA se dégrade dans ces conditions fortement basiques en une espèce non-mesurée et non prise en compte dans le modèle.

5.2.3.3 Itération n°2

Avec les paramètres issus de l'itération n°1, la stratégie n'a pas réussi à trouver de conditions permettant de satisfaire l'ensemble des critères (les concentrations en bétaine dans le produit final et en MCA résiduelle sont trop élevées) comme le montre le Tableau 5.9. Les conditions mesurées sont un peu meilleures, mais pas suffisamment pour être validées.

Tableau 5.9 : Comparaison entre performances prédites et mesurées à l'itération n°2.

<i>Performance</i>	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
<i>Objectif</i>	10	0	0	0	0
<i>Veto</i>	120	1,00	5,25	0,70	200
<i>Prédictions</i>	30	7,30	12,66	$0,00 \pm 0,03$	53500 ± 59600
<i>Mesures</i>	30	7,30	12,66	$0,65 \pm 0,24$	15980 ± 126

La Figure 5.26 montre la régression paramétrique obtenue en complétant la base de données expérimentale par ces deux nouvelles expériences (voir le nouveau jeu de paramètres dans le Tableau 5.7, ligne « itération n°2 »). L'expérience optimale de performance proposée à cette itération est relativement semblable à celle de l'itération n°1 (forte température, excès de MCA, coulée de soude continue). En revanche, l'expérience d'acquisition de données est, contrairement à précédemment, à haute température (130°C contre 83°C). Les proportions de réactifs sont $R_{MCA/AA} = 1,880$ et $R_{NaOH/MCA} = 1,350$ et la soude est introduit avant le MCA. Il en résulte une forte production d'acide glycolique. On peut remarquer que jusqu'à présent, toutes les expériences proposées par la stratégie sont dans le réacteur semi-batch.

5.2. APPLICATION DE LA STRATÉGIE

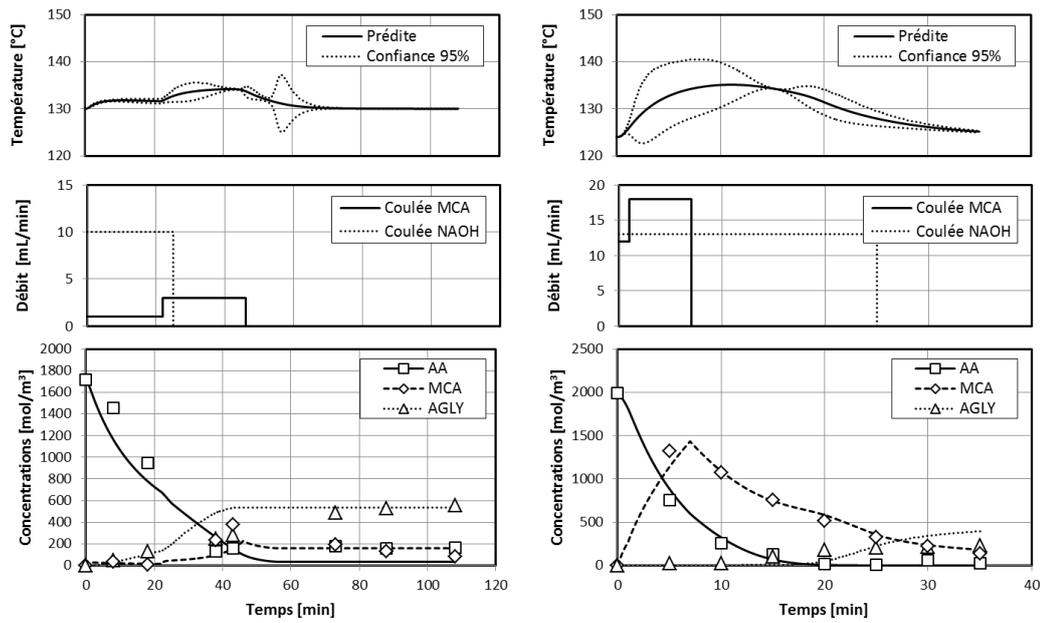


Figure 5.26 : Régression paramétrique de l'expérience d'acquisition de données (à gauche) et de validation (à droite) de l'itération n°2.

Comme précédemment, deux autres expériences sont alors proposées.

5.2.3.4 Itération n°3

Là encore, la stratégie ne propose pas d'expérience satisfaisant les critères de performances (voir Tableau 5.10) et une nouvelle régression paramétrique est réalisée (voir Figure 5.27).

Tableau 5.10 : Comparaison entre performances prédites et mesurées à l'itération n°3.

Performance	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
Objectif	10	0	0	0	0
Veto	120	1,00	5,25	0,70	200
Prédictions	35	12,08	11,18	$0,11 \pm 0,60$	27320 ± 18121
Mesures	35	12,08	11,18	$0,00 \pm 0,24$	15920 ± 126

Les conditions optimales trouvées par la stratégie sont toujours basées sur une température élevée, un excès de MCA ($R_{MCA/AA} = 1,336$) et une coulée de soude continue. L'interprétation que l'on peut donner est que l'excès de MCA accélère la réaction principale et que la soude permet d'éliminer le MCA résiduel en le transformant en acide glycolique sur lequel il n'y a pas de contrainte. L'expérience d'acquisition se déroule à basse température en ajoutant d'abord le MCA, puis dans un second temps un maximum de soude sur une courte durée (8 minutes). Soit une expérience à l'opposé de

l'expérience d'acquisition de données de l'itération n°2. Cela montre bien que la stratégie essaie de balayer le plus possible l'ensemble du domaine expérimental disponible.

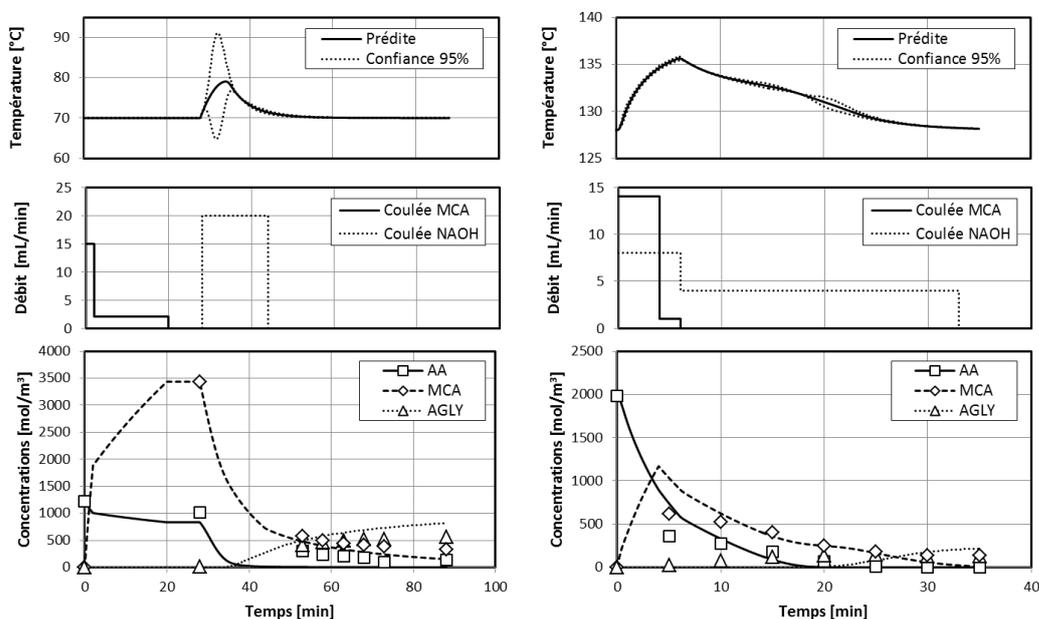


Figure 5.27 : Régression paramétrique de l'expérience d'acquisition de données (à gauche) et de validation (à droite) de l'itération n°3.

5.2.3.5 Itération n°4

Cette fois, les critères prédits, avec les paramètres issus de l'itération n°3 (voir Tableau 5.7, page 188), à l'optimum de performances respectent les contraintes, bien que leur intervalle de confiance dépasse les bornes véto (voir Tableau 5.11). Cependant, les mesures montrent que, expérimentalement, ces contraintes sont respectées. Nous avons donc trouvé des conditions qui permettent de passer d'un temps de synthèse industriel de 5 heures à 35 minutes, et ce tout en respectant les spécifications sur le produit final.

Tableau 5.11 : Comparaison entre performances prédites et mesurées à l'itération n°4.

Performance	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
Objectif	10	0	0	0	0
Veto	120	1,00	5,25	0,70	200
Prédictions	35	0,44	4,40	$0,15 \pm 0,81$	119 ± 576
Mesures	35	0,44	4,40	$0,54 \pm 0,24$	0 ± 126

La Figure 5.28 montre les prédictions du modèle par rapport aux points expérimentaux. Le profil de coulée des réactifs est différent de l'expérience précédente, le

ratio molaire est plus proche de la stœchiométrie ($R_{MCA/AA} = 1,076$ et $R_{NaOH/MCA} = 1,279$) et la température légèrement plus basse.

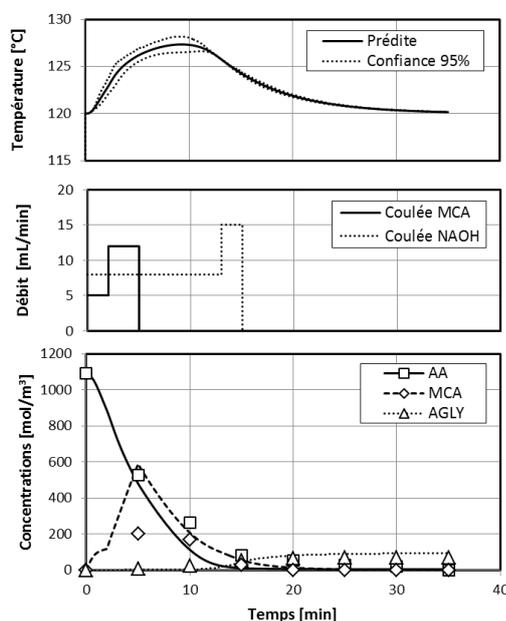


Figure 5.28 : Régression paramétrique de l'expérience de validation de l'itération n°4.

5.3 Discussion

En premier lieu, il est important de remarquer que le réacteur semi-batch est, dans cette étude, toujours le meilleur réacteur, que ce soit pour optimiser les performances ou pour améliorer l'acquisition de données. Par exemple, si l'on voulait trouver quelle est l'expérience optimale d'un point de vue acquisition de données, étant donné que l'on possède déjà la base de données expérimentale obtenue lors des 4 itérations (soit onze expériences en semi-batch), on obtiendrait les critères D-optimaux donnés dans le Tableau 5.12.

Tableau 5.12 : Comparaison des performances de chaque réacteur d'un point de vue acquisition de données.

Performances	Semi-batch	Piston	RPAC
$\ln j_{Dopt}$	-20,68	-18,19	-11,33

Pour rappel, le critère D-optimal est lié au volume de l'ellipsoïde de confiance des paramètres. Ainsi, plus le critère est petit, plus l'on aura confiance aux paramètres estimés. Dans ce cas, même si l'on a déjà réalisé onze expériences en semi-batch, le semi-batch reste encore le meilleur réacteur. Cela peut s'expliquer par le fait que les réactions sont globalement assez lentes, et donc que l'information cinétique se trouve après le temps maximal atteignable en réacteur piston. De plus, le nombre d'échantillons que

l'on peut générer avec les différents réacteurs entre aussi en compte dans cette explication (7 pour le semi-batch, 4 pour le piston et 1 pour la cuve agitée continue).

Le fait que les réactions soient lentes rend le réacteur semi-batch le meilleur réacteur aussi pour la performance. Mais ce n'est pas son seul atout : la possibilité de faire des profils de coulée complexes permet au réacteur semi-batch de favoriser la réaction principale par rapport à la réaction parasite en limitant la concentration du MCA dans le mélange réactionnel. Le Tableau 5.14 donne les meilleurs critères de performances estimés obtenus par chacun des trois réacteurs (calculés à l'aide des paramètres issus de l'itération n°4).

Tableau 5.13 : Comparaison des performances prédite avec chacun des trois réacteurs.

Performances	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
Objectif	10	0,0	0,0	0,00	0
Veto	120	2,0	5,4	0,70	200
Semi-batch	35	0,2	4,4	$0,15 \pm 0,81$	119 ± 576
Piston	8	0,1	9,9	$0,19 \pm 2,71$	77 ± 838
RPAC	645	0,2	9,5	$4,23 \pm 5,33$	1736 ± 1554

Ainsi, seul le semi-batch permet d'atteindre toutes les spécifications sur le produit. Le réacteur piston permet d'atteindre les spécifications sur les réactifs résiduels au détriment de la quantité de *NaCl* produite. En effet, pour atteindre la conversion exigée en 8 minutes, en plus de jouer sur la température de réaction, il est nécessaire d'utiliser un large excès de soude. La cuve agitée continue, quant à elle, ne permet pas d'atteindre les spécifications requises avec un délai raisonnable. Ceci se comprend très bien car les concentrations exigées en sortie sont très faibles, ce qui implique que la vitesse de la réaction principale dans ces conditions est très faible (à titre de comparaison $r_2(t_0) = 8,33 \text{ mol/m}^3/\text{s}$ et $r_2(t_f) = 3,12 \times 10^{-4} \text{ mol/m}^3/\text{s}$). Outre ces considérations, il faut aussi prendre en compte des aspects difficilement chiffrables tels que la flexibilité du semi-batch (on peut changer de synthèse facilement avec une cuve agitée contrairement à un réacteur piston), ainsi que l'automatisation de la chaîne de production.

Le second point que nous souhaitons aborder est l'aspect thermique. Les conditions optimales trouvées considèrent la thermique du réacteur de laboratoire et non celle du réacteur industriel. Cependant, on peut tout de même prédire le comportement du système en réacteur de dimension industrielle (cuve de 50 m^3 , $U = 300 \text{ W/m}^2/\text{K}$ et $A = 50 \text{ m}^2$) dans les conditions optimales en modifiant, dans le modèle, les caractéristiques thermique du réacteur. On obtient les performances données dans le Tableau 5.14. Les performances prédites sont meilleures dans le réacteur industriel (moins de

réactifs résiduels), cela s'explique notamment par le moins bon échange thermique qui provoque une élévation de la température plus importante (+19°C contre +8°C à l'échelle laboratoire) et le maintien d'une haute température en fin de réaction (voir Figure 5.29). Cela a pour effet de ne pas trop favoriser la réaction parasite en début de réaction (donc la sélectivité est conservée), mais en fin de réaction une forte température permet d'accélérer les réactions sans que cela induise un problème de sélectivité. D'où une meilleure performance en réacteur industriel. De ce résultat, on remarque que le profil de température, qui n'a pas été étudié expérimentalement dans ce travail, se révèle être un levier efficace pour l'optimisation des performances.

Tableau 5.14 : Comparaison des performances prédites en réacteur de laboratoire et dans le réacteur industriel.

Performances	t_f [min]	$ \omega_{BET} - 30 $ [%]	ω_{NaCl} [%]	ω_{AA} [%]	ω_{MCA} [ppm]
Objectif	10	0	0	0	0
Veto	120	1,00	5,25	0,70	200
Labo	35	0,44	4,40	$0,15 \pm 0,81$	119 ± 576
Industriel	35	0,44	4,40	$0,06 \pm 0,50$	0 ± 3

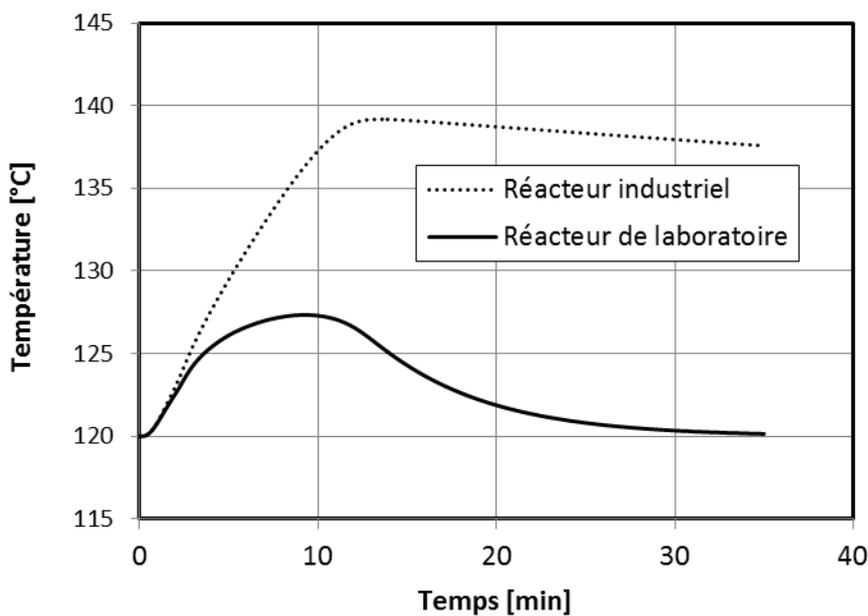


Figure 5.29 : Évolution de la température du milieu réactionnel en réacteur de laboratoire et en réacteur industriel.

Un autre point sur le réalisme des résultats porte sur le coût engendré par le surplus de matières premières utilisées. En effet, les conditions optimales que nous avons trouvées nécessitent un excès de MCA (environ 12%) qui peut être rédhibitoire en termes de coût. Seul une analyse économique incluant en particulier le coût de la matière première, de l'énergie et de la main d'œuvre permettrait de fixer des bornes rai-

sonnables sur l'excès de matière première à ne pas dépasser. Nous n'avons pas les moyens de cette étude, mais nous avons tout de même cherché des conditions optimales en imposant que les excès par rapport à la stœchiométrie soient inférieurs à 1% ($0,99 \leq R_{MCA/AA} \leq 1,01$ et $0,99 \leq R_{NaOH/MCA} \leq 1,01$). Avec ces nouvelles contraintes, le meilleur temps de synthèse trouvé qui respecte les spécifications sur le produit est de 1h55 avec une température de paroi de 119°C. Le profil de coulée utilisé est donné par la Figure 5.30 ainsi que l'évolution de la température (calculée en supposant un réacteur de taille industrielle) et des concentrations prédites.

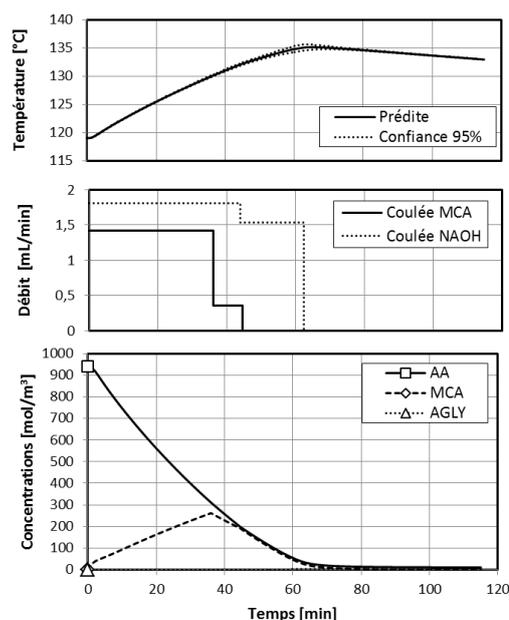


Figure 5.30 : Évolution de la température, du débit de coulée et des concentrations en fonction du temps pour les conditions optimales respectant les contraintes stœchiométriques.

Enfin, toujours du point de vue économique, le surcoût engendré par le passage à une cuve pressurisée permettant de travailler à une température au-dessus du point d'ébullition de l'eau pourrait être suffisant pour que le gain de productivité ne soit pas rentable. Nous avons donc envisagé un cas non pressurisé où la température maximale atteinte au cours de la réaction n'excède pas 100°C. Pour respecter les contraintes stœchiométriques ($\pm 1\%$ d'excès) et pour que la température maximale de réaction n'excède pas 100°C, il est nécessaire de faire la réaction pendant 5h00 avec comme température de paroi 82°C. L'évolution du profil de température, des profils de coulées et des concentrations sont donnés Figure 5.31.

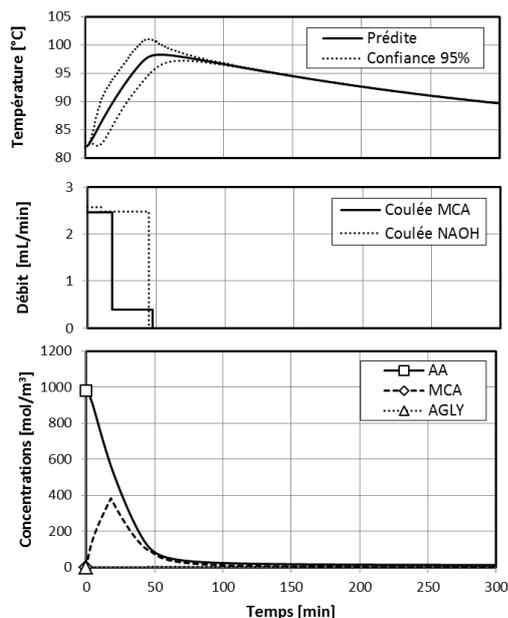


Figure 5.31 : Évolution de la température, du débit de coulée et des concentrations en fonction du temps pour les conditions optimales respectant les contraintes stœchiométriques et de température ($T_{max} \leq 100^{\circ}C$).

Les conditions optimales trouvées en évitant que la réaction ne provoque l'ébullition du mélange sont très proches du procédé actuel.

5.4 Conclusion

La stratégie que nous avons développée, qui consiste à alterner des phases d'acquisition de données pour améliorer la pertinence du modèle et des phases de validation pour vérifier les prédictions de performances optimales, a été appliquée à un cas réel. Il aura fallu trois expériences préliminaires pour élaborer un modèle et quatre itérations de deux expériences chacune pour trouver des conditions expérimentales permettant de réduire drastiquement le temps de synthèse (de 5 heures à 35 minutes) tout en maintenant la qualité du produit. Cette qualité est évaluée par la concentration en réactifs résiduels. Les réactifs étant relativement toxiques (acide monochloroacétique) ou allergènes (laurylamidopropyldiméthylamine) et le produit ayant des fins cosmétiques, leur taux résiduel acceptable est donc très faible. Toute la difficulté de cette synthèse réside ainsi dans le fait qu'il faille atteindre un taux de conversion très élevé ($> 0,995$). En outre, il existe une réaction parasite qui s'active à haute température.

Le modèle utilisé pour décrire le comportement du système a été développé par l'analyse de l'influence de la température, du pH, du transfert à l'interface (dont nous avons montré qu'il était fortement lié à la température). Ce modèle, basé sur des bilans de matière, d'énergie et d'une cinétique simplifiée permet une représentation assez fidèle du comportement réel. On observe toutefois que le modèle n'est pas capable de rendre compte précisément des points expérimentaux (voir les courbes des régressions).

Malgré tout, le modèle donne les bonnes tendances et permet d'améliorer le temps de synthèse. On aurait pu écrire un modèle plus complexe, mais le nombre de paramètres aurait été beaucoup plus grand que les quatre utilisés ici et nombre d'entre eux n'auraient pas été significatifs.

Ces réactions sont exothermiques, leur enthalpie nous avait été fournie de manière indépendante. Pour éviter tout risque d'emballement ou de dégradation du produit, toutes les expériences proposées par la stratégie ont respecté la règle de sécurité indiquant que la température maximale atteinte au cours de la réaction ne devait pas dépasser 150°C.

Cette stratégie a été développée pour montrer l'intérêt de changer de technologie, et il s'avère que dans ce cas précis, le réacteur semi-batch est le réacteur le plus efficace aussi bien pour l'acquisition de données que pour améliorer les performances. Les expériences proposées pour l'acquisition de données ont balayé une large gamme de températures (83, 130 et 70°C), de même pour les profils de coulées (rapide, lent, MCA avant la soude et inversement). Ceci montre toutefois l'intérêt de notre stratégie qui peut confirmer que des technologies conventionnelles peuvent être utilisées et adaptées. Gardons toutefois à l'esprit que le réacteur semi-batch est un réacteur dont on pilote les coulées, contrairement aux autres configurations étudiées, ce qui lui donne ici un avantage certain. La configuration d'un réacteur continu piston avec des alimentations multi-étagées pourrait également être étudié à partir de la configuration optimale du semi-batch identifié.

Deux leviers d'intensification ont été mis au jour : la température du milieu réactionnel et la concentration en soude. Nous avons arbitrairement limité la quantité de soude utilisée par la spécification du produit stipulant que la concentration en chlorure de sodium ne doit pas dépasser 5,4 % (soit un excès stœchiométrique d'environ 30%) et non par une valeur provenant d'une étude économique. Ainsi, les conditions optimales trouvées (et vérifiées expérimentalement) utilisent un excès de MCA et un excès de soude relativement important ($R_{MCA/AA} = 1,076$ et $R_{NaOH/MCA} = 1,279$). Pour se rapprocher d'une proposition plus réaliste, nous avons premièrement simulé le comportement du réacteur industriel, et non du réacteur de laboratoire, dans les conditions optimales. Le réacteur industriel évacuant moins bien la chaleur, la température du mélange en fin de réaction est plus élevée et la conversion encore meilleure que dans le réacteur de laboratoire. Puis, deuxièmement, nous avons proposé des conditions respectant des contraintes sur les excès stœchiométriques (excès $\leq \pm 1\%$) ce qui limite fortement la quantité de soude utilisée. Le temps de réaction passe ainsi de 35 minutes à 1h55. Enfin et dernièrement, nous avons envisagé un cas où la température ne devait pas dépasser 100°C. Les conditions ainsi obtenues sont très proches des conditions actuelles du procédé.

Pour aller plus loin, il faudrait catalyser la réaction comme l'a fait Evonik ou utiliser une combinaison de réacteurs. On pourrait imaginer utiliser une série de plusieurs réacteurs pistons ayant des températures de plus en plus élevées avec possiblement des injections de soude. Ainsi, cela compenserait la perte de vitesse liée aux faibles concen-

5.4. CONCLUSION

trations par l'activation par la température et par le pH. Ainsi, le modèle prévoit que la configuration proposée dans Figure 5.32 atteint les spécifications sur le produit en 21 minutes tout en respectant les contraintes stœchiométriques (pas d'excès supérieur $\pm 1\%$). Pour rappel, le modèle prévoyait 1h55 pour le réacteur semi-batch avec les mêmes contraintes. Cependant, ce ne sont que des prédictions d'un modèle développé en réacteur semi-batch, il faut donc être très prudent sur la validité de ces résultats, qu'il serait très intéressant de valider expérimentalement. Il est fort probable que la validation de ces prédictions montrerait que le modèle n'était pas quantitatif, pour autant, il permet de proposer qualitativement un concept de procédé.

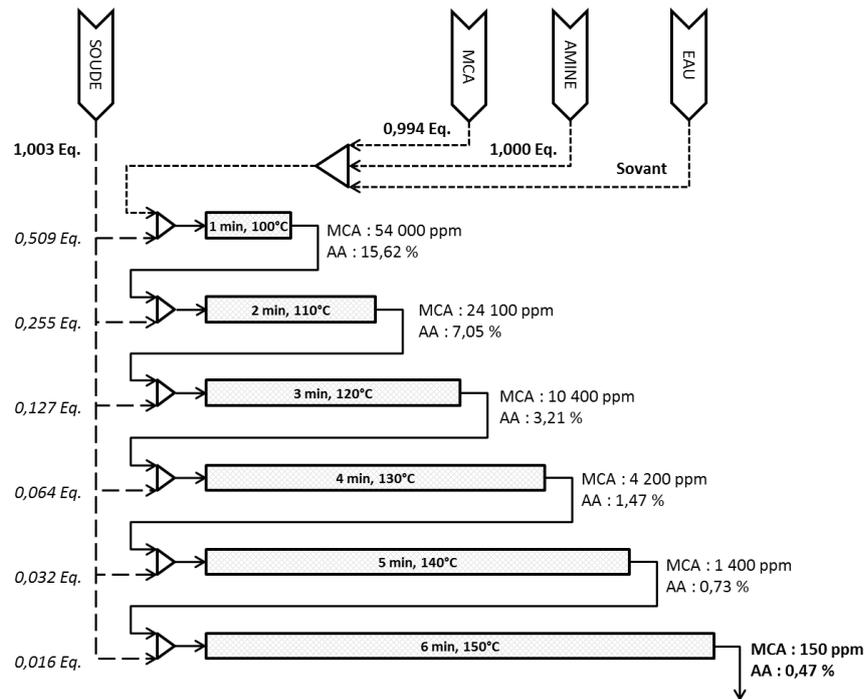


Figure 5.32 : Proposition de mise en série de réacteur piston de plus en plus chaud.

Conclusions et perspectives

La stratégie Stanislas a été développée dans le but de choisir la meilleure technologie disponible pour une synthèse donnée, l'objectif étant de minimiser l'effort expérimental nécessaire.

Le fait de prendre en compte plusieurs technologies implique de gérer, d'une part, un domaine de conditions opératoires complexe et d'autre part, la capacité, propre à chaque technologie, de générer plusieurs échantillons par expérience. Les stratégies expérimentales existantes, en particulier, celles reposant sur des modèles de tendance (type plan d'expériences), ne permettent pas de gérer convenablement ces deux aspects (domaine opératoire complexe et flexibilité d'échantillonnage).

Nous avons donc opté pour l'utilisation de modèles phénoménologiques. La principale difficulté liée à l'utilisation de ce type de modèle réside dans le choix de la structure du modèle, l'objectif étant d'avoir un modèle suffisamment détaillé pour rendre compte des principaux phénomènes sans pour autant avoir une trop grande complexité. Pour atteindre cet objectif, nous avons mis au point un protocole expérimental basé sur une analyse en temps caractéristiques. Ce protocole n'est pas automatique et exige de la part de l'utilisateur un certain savoir-faire. Une fois le modèle établi, la stratégie alterne des phases d'acquisition de données permettant d'améliorer les prédictions du modèle et des phases de validation des prédictions, jusqu'à ce que les critères de convergence soient atteints.

En parallèle de la stratégie, nous avons développé un banc d'essai expérimental permettant l'utilisation aisée de différentes technologies (cuve agitée fermée, semi-fermée, continue et réacteur piston). Ce banc d'essai, réalisé par l'atelier du laboratoire LRGP, a été caractérisé thermiquement. Les matériaux utilisés (acier inox 316L) ont été choisis pour leur qualité de résistance face à la corrosion, permettant l'utilisation de la plupart des chimies. Le système expérimental a été conçu pour atteindre des températures allant de -40°C à 165°C et des pressions allant de 1 à 10 bar.

La stratégie a été appliquée sur des systèmes virtuels, puis sur un cas d'étude simplifié mais réel et enfin sur un cas industriel avec de réelles problématiques industrielles. Ces différentes applications nous ont permis de mettre en évidence les points forts, les limites et le potentiel de la stratégie.

Les points forts

Son efficacité

Nous avons montré dans le chapitre 3 que la stratégie Stanislas permet de réduire significativement le nombre d'expériences et d'améliorer la précision des prédictions par rapport à l'Algorithme de Nancy et aux techniques de plan d'expériences conventionnelles. Le gain par rapport aux plans d'expériences conventionnels peut s'expliquer d'une part, grâce aux connaissances que l'utilisateur apporte à la stratégie et d'autre part, grâce au fait que la stratégie Stanislas a été élaborée en prenant en compte les domaines expérimentaux et la capacité d'échantillonnage propre à chaque technologie. Enfin, la stratégie Stanislas est itérative, elle s'arrête quand l'objectif est atteint, contrairement aux plans d'expériences qui proposent un nombre déterminé à l'avance d'expériences à faire.

La différence majeure avec l'Algorithme de Nancy, qui est aussi une stratégie itérative, est l'étape d'acquisition de données. En effet, la base de données expérimentale utilisée par l'Algorithme de Nancy pour identifier les paramètres est basée seulement sur les points à l'optimum proposés par la stratégie. Or, l'optimum de performance n'est pas forcément intéressant pour améliorer la connaissance des paramètres du modèle. C'est pourquoi la stratégie Stanislas propose de réaliser des expériences riches en informations jusqu'à ce que le modèle soit jugé suffisamment précis.

Un autre point important de la stratégie est sa robustesse. Comme nous l'avons montré dans le chapitre 3, la stratégie permis à trouver l'optimum de performance sur plusieurs systèmes réactionnels. Nous avons testé l'influence du nombre d'espèces mesurées, de la valeur des paramètres du modèle, ainsi que de l'erreur expérimentale. Dans la plupart des cas testés, la stratégie a convergé. Ce résultat ne donne pas les limites de la stratégie et il serait intéressant d'étendre le champ du test de robustesse jusqu'à ce que la stratégie ne converge plus.

Enfin, du point de vue de la sécurité des expériences proposées, la stratégie ne propose pas d'expérience susceptible d'induire un emballement thermique. Pour ce faire, la stratégie prédit la température maximale atteinte au cours de l'expérience proposée ainsi que son intervalle de confiance à 95%. Si la température maximale de sécurité fixée par l'utilisateur se trouve dans l'intervalle de confiance, alors il y a un risque de dépasser cette température et la stratégie ne proposera pas cette expérience. Ce point est très important, car lorsque les phénomènes dépendent de la température, il n'est pas rare que les expériences à hautes température soient intéressantes autant pour l'acquisition de données que pour optimiser les performances

Sa robustesse

Grâce à l'utilisation d'expériences virtuelles, qui permettent de générer un grand nombre de points expérimentaux en un temps réduit, nous avons pu tester la répétabilité de la stratégie. Ainsi, il a été mis en évidence que le niveau de précision avec lequel l'optimum de performance est trouvé dépend du système réactionnel étudié. Par ailleurs, nous avons testé la robustesse de la stratégie en faisant varier le nombre d'espèces mesurées, l'erreur expérimentale et les valeurs des différents paramètres de chaque système virtuel étudié. Cela nous a permis d'identifier les conditions dans lesquelles la stratégie ne réussit pas à trouver l'optimum de performance vrai. Les cas où la stratégie a échoué sont liés soit à une trop grande erreur expérimentale (supérieure à 20 ou 50%), soit au fait que certains phénomènes sont très indépendants des espèces mesurées.

Par ailleurs, la stratégie a réussi à converger sur le cas d'étude simplifié et le cas d'étude industriel, confortant les résultats obtenus avec les expériences virtuelles. La stratégie Stanislas semble donc pouvoir s'appliquer à une large gamme de réactions. Afin de définitivement valider sa robustesse, il est nécessaire d'étendre son application à d'autres synthèses homogènes ainsi qu'à d'autres aspects des procédés tels que la cristallisation, les synthèses hétérogènes, la séparation, ... C'est-à-dire, toutes les opérations unitaires, pourvu qu'elles puissent être modélisées par des équations différentielles ordinaires.

Les limites

Choix de l'utilisateur

La principale limite de cette stratégie provient de l'utilisateur lui-même. Ce sont les informations que l'on donne à la stratégie sous forme d'équations qui lui permettent de réduire l'effort expérimental par rapport aux stratégies utilisant des modèles de tendance. Or, l'utilisateur doit faire des choix subjectifs lors de la construction du modèle. En effet, même si nous avons établi un protocole expérimental permettant d'identifier les principaux phénomènes afin de les prendre en compte dans le modèle, le choix des phénomènes et leur mise en équation restent le fruit de l'utilisateur. Seul ce dernier peut décider ou non de modéliser tel ou tel phénomène. Si un phénomène majeur est omis dans le modèle, la stratégie risque de ne pas converger vers un optimum satisfaisant. De même, si trop de phénomènes sont pris en compte, le modèle sera trop complexe, il sera alors difficile d'identifier les paramètres et le modèle ne sera pas prédictif.

Dans le chapitre 3 où la stratégie a été testée sur des expériences virtuelles, le problème du choix du modèle ne se posait pas car le même modèle était utilisé pour générer les données expérimentales et pour le déroulement de la stratégie expérimentale. C'est pour cela que le chapitre utilisant des expériences virtuelles ont été complété

par les validations expérimentales des chapitres 4 et 5. Dans le chapitre 4, le modèle du système réactionnel a été établi sur des publications. Pour le chapitre 5, le modèle a été construit grâce au protocole d'initialisation que nous avons mis au point. Les choix de simplifier le modèle, en incluant par exemple les limitations de transfert dans une cinétique apparente ont tout de même permis de proposer des conditions expérimentales permettant de réduire significativement le temps de synthèse.

Pour pouvoir mesurer l'influence des choix de l'utilisateur, il serait intéressant d'appliquer la stratégie en utilisant un modèle plus complet. Cette étude pourrait également être menée de manière virtuelle. Par exemple, en générant des points expérimentaux par un code numérique inconnu par l'utilisateur.

Couplage de technologies

La stratégie a été développée en ne considérant les technologies qu'une à une, séparément. Or, une des conclusions importantes de l'application industrielle est de montrer que l'on peut considérablement améliorer le procédé en considérant un couplage d'équipements (série de réacteurs piston à différents niveaux de température). On pourrait considérer un ensemble d'équipements comme étant une seule et unique technologie. Cela implique trois complications majeures.

La première provient de la modélisation. En effet, il est assez facile de modéliser une cuve agitée ou un réacteur piston et de pouvoir calculer des sensibilités (car ces équipements sont décrits par des équations différentielles ordinaires), alors qu'une succession de différents réacteurs peut induire des singularités qui rendent beaucoup plus laborieux le calcul des sensibilités (sensibilités qui sont indispensables pour la planification d'expériences).

La deuxième complication provient de la démultiplication du nombre de technologies et de conditions expérimentales. Si l'on considère une configuration avec trois réacteurs qui peuvent être soit une cuve agitée continue, soit un réacteur tubulaire, on obtient 8 configurations possibles. Pour chacune de ces huit configurations, il faut définir de nombreuses conditions expérimentales (le temps de passage, la température, les débits d'injection de chacune des trois technologies). D'un point de vue numérique, cela devient extrêmement coûteux en temps de calcul et peut devenir très rapidement réducteur.

La troisième complication est d'ordre expérimental. Coupler plusieurs réacteurs nécessite une quantité d'appareillages non négligeable (pompes, bains thermostatés, réacteurs tubulaires de différentes longueurs, ...). Le coût, la réalisation et la mise en œuvre pratique d'un tel montage risqueraient d'annuler le bénéfice du gain de la stratégie en termes d'efforts expérimentaux.

Même en ne se focalisant que sur l'étape de réaction, considérer un couplage de plusieurs technologies est assez complexe. Il semble donc très difficile de pouvoir étendre cette stratégie à un procédé comprenant plusieurs opérations unitaires dans sa

totalité. Or, une des voies de l'intensification des procédés repose sur le réagencement d'opérations unitaires voire même la fusion de plusieurs opérations unitaires en une seule et même technologie (tel que la distillation réactive ou les réacteurs membranaires). Il y a donc un enjeu important sur ce point et développer une stratégie expérimentale pour répondre à cette problématique est très prometteur.

Le potentiel

Un outil d'aide à la conception

La stratégie Stanislas a été développée dans le but de choisir la meilleure technologie disponible. Mais, elle pourrait aussi servir à la conception de nouvelles technologies. En effet, en élargissant virtuellement les bornes des domaines de conditions opératoires, on pourrait savoir s'il est intéressant, aussi bien du point de vue de l'acquisition de données que de l'optimisation de performances, d'explorer expérimentalement des zones non atteignables par les technologies disponibles. Il faudrait alors imaginer une technologie qui permette d'atteindre ces zones potentiellement intéressantes. De la même manière que dans le chapitre sur la synthèse industrielle, une fois que le modèle est suffisamment prédictif, on peut alors le réinterpréter et proposer une solution technologique pas forcément prévue dans les technologies étudiées au départ.

Ainsi, dans le cas de la synthèse industrielle, l'analyse des résultats proposés par la stratégie nous a permis d'identifier une solution technologique que nous n'avions pas considérée dans un premier temps (la série de réacteurs piston). En allant plus loin, on pourrait imaginer que pour un système donné, la stratégie mette en évidence une limitation de transfert de matière liée à un système biphasique : l'utilisateur saura alors qu'il peut être intéressant pour lui de se lancer dans une étude pour tenter de passer d'un système biphasique à un système monophasique, en changeant de solvant par exemple.

L'automatisation complète et l'interfaçage graphique

À l'issue de ce travail, le concept de la stratégie a été bien développé et peut être appliqué à d'autres cas d'étude. Cependant, dans la pratique c'est l'utilisateur qui fait le lien entre le code Matlab, le banc expérimental et le système analytique. Un nouvel objectif pourrait consister à éviter à l'utilisateur ce travail pour qu'il puisse se concentrer sur la compréhension du système réactionnel. Pour cela, il serait fort intéressant de créer une interface entre le code de calcul, le banc expérimental et les moyens analytiques. On peut ainsi imaginer que le déroulement de la stratégie soit semi-automatique voire même complètement automatisé.

Enfin, pour faciliter son utilisation à l'utilisateur, une interface graphique pourrait être développée. L'utilisateur devra alors renseigner les caractéristiques des technologies qu'il souhaite utiliser, les données physico-chimiques élémentaires des espèces mises en œuvre et ses critères de performances. En couplant l'interface graphique à l'automatisation complète (code – banc expérimental – moyens analytiques), on se rapprocherait alors d'un produit commercialisable.

Bibliographie

- Agarwal, A. K., et M. L. Brisk. 1985. « Sequential experimental design for precise parameter estimation. 1. Use of reparameterization ». *Industrial & Engineering Chemistry Process Design and Development* 24 (1): 203–207.
- Benyahia, B. 2009. « Modélisation, expérimentation et optimisation multicritère d'un procédé de copolymérisation en émulsion en présence d'un agent de transfert de chaîne ». Ph. D. thesis. France: National Polytechnic Institute of Lorraine, Nancy-University.
- Bonvin, D., et D. W. T. Rippin. 1990. « Target factor analysis for the identification of stoichiometric models ». *Chemical engineering science* 45 (12): 3417–3426.
- Box, G. E. P. 1957. « Evolutionary operation: A method for increasing industrial productivity ». *Applied Statistics*: 81–101.
- Box, G. E. P., et W. J. Hill. 1967. « Discrimination among mechanistic models ». *Technometrics*: 57–71.
- Box, G. E. P., J. S. Hunter, et W. G. Hunter. 2005. *Statistics for experimenters: design, innovation, and discovery*. Vol. 13. Wiley Online Library.
- Box, G. E. P., et H. L. Lucas. 1959. « Design of experiments in non-linear situations ». *Biometrika* 46 (1/2): 77–90.
- Brans, J. P., P. Vincke, et B. Mareschal. 1986. « How to select and how to rank projects: The PROMETHEE method ». *European journal of operational research* 24 (2): 228–238.
- Brendel, M., D. Bonvin, et W. Marquardt. 2006. « Incremental identification of kinetic models for homogeneous reaction systems ». *Chemical engineering science* 61 (16): 5404–5420.
- Brey, N. L., et J. F. Fowler Jr. 2004. « Relevance of positive patch-test reactions to cocamidopropyl betaine and amidoamine. » *Dermatitis: contact, atopic, occupational, drug: official journal of the American Contact Dermatitis Society, North American Contact Dermatitis Group* 15 (1): 7.
- Chapman, F. S., H. Dallenbach, et F. A. Holland. 1964. « Heat transfer in baffled, jacketed agitated vessels ». *Trans. Inst. Chem. Engng* 42: T398–T406.
- Chu, G. W., Y. H. Song, H. J. Yang, J. M. Chen, H. Chen, et J. F. Chen. 2007. « Micromixing efficiency of a novel rotor–stator reactor ». *Chemical Engineering Journal* 128 (2): 191–196.
- Churchill, Stuart W., et Humbert HS Chu. 1975. « Correlating equations for laminar and turbulent free convection from a horizontal cylinder ». *International Journal of Heat and Mass Transfer* 18 (9): 1049–1053.
- Coello, C. A. C., G. B. Lamont, et D. A. Van Veldhuizen. 2007. *Evolutionary algorithms for solving multi-objective problems*. Vol. 5. Springer.
- Commenge, J. M., L. Falk, J. P. Corriou, et M. Matlosz. 2002. « Optimal design for flow uniformity in microchannel reactors ». *AIChE Journal* 48 (2): 345–358.
- Conn, A. R., Gould, N. I., & Toint, P. (1991). "A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds". *SIAM Journal on Numerical Analysis*, 28(2), 545-572.

- Cook, R. D., & Nachtrheim, C. J. 1980. "A comparison of algorithms for constructing exact D-optimal designs". *Technometrics*, 22(3), 315-324.
- Craven, P., et G. Wahba. 1978. « Smoothing noisy data with spline functions ». *Numerische Mathematik* 31 (4): 377-403.
- Csizmadia, Ferenc, Anna Tsantili-Kakoulidou, Irene Panderi, et Ferenc Darvas. 1997. « Prediction of Distribution Coefficient from Structure. 1. Estimation Method ». *Journal of Pharmaceutical Sciences* 86 (7): 865-871. doi:10.1021/js960177k.
- De Panthou, F., et P. Giuliano. 2008. « Agitated continuous mini-reactors: an industrial option for Fine Chemicals ». *Chimica oggi* 26 (3): 6.
- Ehrfeld, W., V. Hessel, et V. Haverkamp. 2000. *Microreactors*. Wiley Online Library.
- Falk, L., & Commenge, J. M. 2010. "Performance comparison of micromixers". *Chemical Engineering Science*, 65(1), 405-411.
- Fedorov, V. V. 1972. « Theory of optimal experiments ».
- Foti, Caterina, Domenico Bonamonte, Giuseppe Mascolo, Angela Corcelli, Simona Lobasso, Luigi Rigano, et Gianni Angelini. 2003. « The role of 3-dimethylaminopropylamine and amidoamine in contact allergy to cocamidopropylbetaine ». *Contact Dermatitis* 48 (4): 194-198.
- Franceschini, G., et S. Macchietto. 2008. « Model-based design of experiments for parameter precision: State of the art ». *Chemical Engineering Science* 63 (19): 4846-4872.
- Garcia Molla, Victor. 2003. « Sensitivity Analysis for ODEs and DAEs - File Exchange - MATLAB Central ». <http://www.mathworks.com/matlabcentral/fileexchange/1480>.
- Gill, Philip E., Walter Murray, et Margaret H. Wright. 1981. « Practical optimization ». <http://www.citeulike.org/group/1938/article/1055426>.
- Goodwin, GC, et RL Payne. 1977. *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press.
- Greco, S. 2004. *Multiple criteria decision analysis: state of the art surveys*. Vol. 78. Springer.
- Herrwerth, S., H. Leidreiter, H. H. Wenk, M. Farwick, I. Ulrich-Brehm, et B. Grüning. 2013. « Highly Concentrated Cocamidopropyl Betaine—The Latest Developments for Improved Sustainability and Enhanced Skin Care ». Consulté le mai 31. <https://personal-care.evonik.com/product/personal-care/en/media-center/downloads/publications/Documents/tenside-tego-betain-c-60.pdf>.
- Hisamoto, H., T. Saito, M. Tokeshi, A. Hibara, et T. Kitamori. 2001. « Fast and high conversion phase-transfer synthesis exploiting the liquid-liquid interface formed in a microchannel chip ». *Chemical Communications* (24): 2662-2663.
- Hosten, L. H. 1974. « A sequential experimental design procedure for precise parameter estimation based upon the shape of the joint confidence region ». *Chemical Engineering Science* 29 (11): 2247-2252.
- Ince, N. H., G. Tezcanli, R. K. Belen, et İ G. Apikyan. 2001. « Ultrasound as a catalyzer of aqueous reaction systems: the state of the art and environmental applications ». *Applied Catalysis B: Environmental* 29 (3): 167-176.
- International Programme on Chemical Safety (IPCS). 2012. « Monochloroacetic acid (PIM 352) ». <http://www.inchem.org/documents/pims/chemical/pim352.htm>.

-
- Issanchou, S., P. Cognet, et M. Cabassud. 2003. « Precise parameter estimation for chemical batch reactions in heterogeneous medium ». *Chemical engineering science* 58 (9): 1805–1813.
- Jacquez, J. A., et P. Greif. 1985. « Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design ». *Mathematical Biosciences* 77 (1): 201–227.
- Lidström, P., J. Tierney, B. Wathey, et J. Westman. 2001. « Microwave assisted organic synthesis—a review ». *Tetrahedron* 57 (45): 9225–9283.
- MacKay, D. J. C. 1992. « Bayesian interpolation ». *Neural computation* 4 (3): 415–447.
- Maly, T., et L. R. Petzold. 1996. « Numerical methods and software for sensitivity analysis of differential-algebraic systems ». *Applied Numerical Mathematics* 20 (1): 57–79.
- Marchal-Brassely, S., J. Villermaux, J. L. Houzelot, et J. L. Barnay. 1992. « Optimal operation of a semi-batch reactor by self-adaptive models for temperature and feed-rate profiles ». *Chemical engineering science* 47 (9-11): 2445–2450.
- Martin, H., et K. Stephan. 1983. *Heat exchanger design handbook*. Hemisphere Pub. Corp.
- Mayberry, W. E., J. E. Rall, et D. Bertoli. 1964. « Kinetics of iodination. i. A comparison of the kinetics of iodination of N-acetyl-L-tyrosine and N-acetyl-3-iodo-L-tyrosine ». *Journal of the American Chemical Society* 86 (23): 5302–5307.
- McKay, M. D., R. J. Beckman, et W. J. Conover. 1979. « A comparison of three methods for selecting values of input variables in the analysis of output from a computer code ». *Technometrics*: 239–245.
- McMullen, J. P., et K. F. Jensen. 2010. « An automated microfluidic system for online optimization in chemical synthesis ». *Organic Process Research & Development* 14 (5): 1169–1176.
- Mhamdi, A, et W Marquardt. 1999. « An inversion approach to the estimation of reaction rates in chemical reactors ». In Karlsruhe, Germany.
- Mitchell, T. J. 1974. « An algorithm for the construction of “D-optimal” experimental designs ». *Technometrics* 16 (2): 203–210.
- Monnier, H., A. M. Wilhelm, et H. Delmas. 2000. « Effects of ultrasound on micromixing in flow cell ». *Chemical engineering science* 55 (19): 4009–4020.
- Oxley, P., C. Brechtelsbauer, F. Ricard, N. Lewis, et C. Ramshaw. 2000. « Evaluation of spinning disk reactor technology for the manufacture of pharmaceuticals ». *Industrial & engineering chemistry research* 39 (7): 2175–2182.
- Paul, Edward L., Victor Atiemo-Obeng, et Suzanne M. Kresta. 2004. *Handbook of industrial mixing: science and practice*. Wiley-Interscience.
<http://books.google.fr/books?hl=fr&lr=&id=s4SsrmvEc7QC&oi=fnd&pg=PR5&dq=handbook+of+mixing&ots=54Nueae8tU&sig=CMaUHbza5CFozypPpFJdQmoSzSc>.
- Pennemann, Helmut, Volker Hessel, et Holger Löwe. 2004. « Chemical microprocess technology—from laboratory-scale to production ». *Chemical Engineering Science* 59 (22–23) (novembre): 4789–4794. doi:10.1016/j.ces.2004.07.049.
- Perry, Robert H., et Don W. Green. 2008. *Perry's chemical engineers' handbook*. McGraw-Hill New York.
<http://www.lavoisier.fr/livre/notice.asp?id=OOLWRKA336KOWD>.
-

- Pohekar, S. D., et M. Ramachandran. 2004. « Application of multi-criteria decision making to sustainable energy planning—A review ». *Renewable and Sustainable Energy Reviews* 8 (4): 365–381.
- Pritchard, D. J., et D. W. Bacon. 1978. « Prospects for reducing correlations among parameter estimates in kinetic models ». *Chemical Engineering Science* 33 (11): 1539–1543.
- Pronzato, Luc, et Eric Walter. 1988. « Robust experiment design via maximin optimization ». *Mathematical Biosciences* 89 (2) (juin): 161–176. doi:10.1016/0025-5564(88)90097-1.
- Radhakrishna Rao, C. 1945. « Information and accuracy attainable in the estimation of statistical parameters ». *Bulletin of the Calcutta Mathematical Society* 37 (3): 81–91.
- Rao, D. P., A. Bhowal, et P. S. Goswami. 2004. « Process intensification in rotating packed beds (HIGEE): an appraisal ». *Industrial & engineering chemistry research* 43 (4): 1150–1162.
- Reyes-Sierra, M., et C. A. C. Coello. 2006. « Multi-objective particle swarm optimizers: A survey of the state-of-the-art ». *International Journal of Computational Intelligence Research* 2 (3): 287–308.
- Rimensberger, T., et D. W. T. Rippin. 1986. « “ Sequential experimental design for precise parameter estimation. 1. Use of reparameterization”. Comments ». *Industrial & Engineering Chemistry Process Design and Development* 25 (4): 1042–1044.
- Roberge, D. M., L. Ducry, N. Bieler, P. Cretton, et B. Zimmermann. 2005. « Microreactor technology: A revolution for the fine chemical and pharmaceutical industries? » *Chemical engineering & technology* 28 (3): 318–323.
- Roberge, D. M., B. Zimmermann, F. Rainone, M. Gottsponer, M. Eyholzer, et N. Kockmann. 2008. « Microreactor technology and continuous processes in the fine chemical and pharmaceutical industry: Is the revolution underway? » *Organic Process Research & Development* 12 (5): 905–910.
- Rosenbrock, H. H., et C. Storey. 1966. *Computational techniques for chemical engineers*. Pergamon Press New York.
- Rushton, J. H., E. W. Costich, et H. J. Everett. 1950. « Power characteristics of mixing impellers ». *Chem. Eng. Prog* 46 (8): 395.
- Saber, M., J. M. Commenge, et L. Falk. 2010. « Microreactor numbering-up in multi-scale networks for industrial-scale applications: Impact of flow maldistribution on the reactor performances ». *Chemical Engineering Science* 65 (1): 372–379.
- Schweich, Daniel. 2001. *Génie de la réaction chimique*. Tec & Doc.
- Shirure, V. S., S. Anand, et V. G. Pangarkar. 2005. « Intensification of precipitation using narrow channel reactors: magnesium hydroxide precipitation ». *Industrial & engineering chemistry research* 44 (15): 5500–5507.
- Srinivasan, B., C. J. Primus, D. Bonvin, et N. L. Ricker. 2001. « Run-to-run optimization via control of generalized constraints ». *Control Engineering Practice* 9 (8): 911–919.
- Walter, É, et L. Pronzato. 1994. Identification de modèles paramétriques à partir de données expérimentales. Masson.
- Walter, E., et L. Pronzato. 1990. « Qualitative and quantitative experiment design for phenomenological models—a survey ». *Automatica* 26 (2): 195–213.

-
- Yao, K. Z., B. M. Shaw, B. Kou, K. B. McAuley, et D. W. Bacon. 2003. « Modeling Ethylene/Butene Copolymerization with Multi-site Catalysts: Parameter Estimability and Experimental Design ». *Polymer Reaction Engineering* 11 (3): 563–588.
- Zhao, Hong, Lei Shao, et Jian-Feng Chen. 2010. « High-gravity process intensification technology and application ». *Chemical Engineering Journal* 156 (3) (février 1): 588-593. doi:10.1016/j.cej.2009.04.053.

RÉSUMÉ

DEPUIS QUELQUES ANNÉES, LES RÉACTEURS INTENSIFIÉS ET LES RÉACTEURS MICROSTRUCTURÉS SONT UNE ALTERNATIVE AUX RÉACTEURS EN CUVE AGITÉE. GRÂCE À LEURS PERFORMANCES ÉLEVÉES EN TERMES DE TRANSFERT DE MATIÈRE ET TRANSFERT DE CHALEUR, IL EST POSSIBLE DE MIEUX MAÎTRISER LES CONDITIONS OPÉRATOIRES ET DONC D'ATTEINDRE DES MEILLEURES QUALITÉS DE PRODUIT. CEPENDANT, IL EST DIFFICILE DE PRÉDIRE QUANTITATIVEMENT L'INTÉRÊT RÉEL DES RÉACTEURS INTENSIFIÉS OU DU PASSAGE BATCH/CONTINU, POUR UN SYSTÈME CHIMIQUE DONNÉ, SANS UN EFFORT EXPÉRIMENTAL IMPORTANT. AINSI, RÉDUIRE CETTE PHASE EXPÉRIMENTALE COÛTEUSE EN TEMPS ET EN MOYEN EST UNE PRIORITÉ.

POUR CELA, UNE STRATÉGIE ITÉRATIVE BASÉE SUR LA PLANIFICATION D'EXPÉRIENCE ET LA MODÉLISATION A ÉTÉ DÉVELOPÉE. ELLE PERMET DE PROFITER PLEINEMENT DE L'ÉLARGISSEMENT DES FENÊTRES OPÉRATOIRES LIÉES AUX NOUVELLES TECHNOLOGIES, TOUT EN MINIMISANT LE NOMBRE D'EXPÉRIENCES. CETTE STRATÉGIE UTILISE UN MODÈLE PHÉNOMÉNOLOGIQUE QUI DÉCRIT LES RÉACTIONS ET LES TRANSFERTS DE MASSE ET DE CHALEUR DANS CHACUN DES RÉACTEURS PROPOSÉS. À CHAQUE ITÉRATION, LA STRATÉGIE PLANIFIE L'EXPÉRIENCE QUI PERMET D'ACQUÉRIR LE PLUS D'INFORMATION EN S' Aidant DE CRITÈRES STATISTIQUE TEL QUE LE CRITÈRE D-OPTIMAL. PUIS, SI LA PRÉCISION DES PARAMÈTRES EST JUGÉE SUFFISANTE, ELLE PRÉDIT QUEL RÉACTEUR ATTEINT LA MEILLEUR PERFORMANCE ET PROPOSE DE VALIDER EXPÉRIMENTALEMENT LA PRÉDICTION. L'ESTIMATION DE LA PERFORMANCE PASSE PAR UNE FONCTION DE COÛT BASÉE SUR PLUSIEURS CRITÈRES QUI PEUVENT ÊTRE LE RENDEMENT, LE TEMPS D'OPÉRATION, LA CONSOMMATION ÉNERGÉTIQUE,... SI LA VALIDATION EXPÉRIMENTALE ÉCHOUE, LA STRATÉGIE PROPOSE DE CONTINUER LA PHASE D'ACQUISITION APRÈS QUE L'UTILISATEUR AIT APPORTÉ DES MODIFICATIONS AU MODÈLE.

LE BIENFONDÉ DE LA STRATÉGIE ET L'INTÉRÊT DE PRENDRE EN COMPTE PLUSIEURS TECHNOLOGIES POUR LES PERFORMANCES, MAIS AUSSI ET SURTOUT POUR L'ACQUISITION DE DONNÉES, A D'ABORD ÉTÉ DÉMONTRÉ EN UTILISANT DES SYSTÈMES CHIMIQUES VIRTUELS. PUIS DANS UN SECOND TEMPS, UNE RÉACTION CLASSIQUE DE TYPE PARALLÈLE/CONSÉCUTIVE A ÉTÉ UTILISÉE. POUR CELA, UN BANC EXPÉRIMENTAL SPÉCIFIQUE A ÉTÉ MISE EN ŒUVRE PERMETTANT L'UTILISATION DE TROIS TYPES DE RÉACTEURS : BATCH, SEMI-BATCH ET PISTON. ENFIN, SOLVAY, NOTRE PARTENAIRE INDUSTRIEL, NOUS A FOURNI UN CAS D'ÉTUDE RÉEL PORTANT SUR LA SYNTHÈSE DE SURFACTANT AYANT DES APPLICATIONS COSMÉTIQUES. CES TRAVAUX ONT PERMIS DE METTRE EN ÉVIDENCE L'INTÉRÊT D'UTILISER PLUSIEURS RÉACTEURS POUR L'ACQUISITION DE DONNÉES, DE DÉMONTRER L'IMPORTANCE DES PHASES D'ACQUISITION DE DONNÉES POUR L'OPTIMISATION DU PROCÉDÉ ET DE SOULIGNER LA SUPÉRIORITÉ DES MODÈLES PHÉNOMÉNOLOGIQUES PAR RAPPORT AUX MODÈLES DE TENDANCES, QUI SONT UTILISÉS DANS LES PLANS D'EXPÉRIENCES « CLASSIQUE ».

ABSTRACT

SINCE A FEW YEARS, INTENSIFIED AND MICROSTRUCTURED REACTORS ARE AN ALTERNATIVE TO STIRRED TANK REACTORS. THANKS TO THEIR HIGH HEAT AND MASS-TRANSFER RATES, A BETTER CONTROL OF OPERATING CONDITIONS IS POSSIBLE LEADING TO A BETTER PRODUCT QUALITY. HOWEVER, IT IS DIFFICULT TO PREDICT QUANTITATIVELY THE REAL INTEREST OF INTENSIFIED REACTORS AND ALSO TO TRANSPOSE A PRIORI A BATCH PROCESS TO A CONTINUOUS PROCESS, WITHOUT AN EXTENSIVE EXPERIMENTAL WORK AIMING AT SELECTING THE BEST REACTOR TECHNOLOGY AND OPTIMAL OPERATING CONDITIONS. THE ACCELERATION OF THIS EXPLORATORY STEP, COSTLY IN TIME AND MONEY, IS THEREFORE A CHALLENGING TASK.

HENCE, AN ITERATIVE STRATEGY BASED ON THE COMBINATION OF EXPERIMENTAL DESIGN AND MODELING HAS BEEN DEVELOPED TO TAKE ADVANTAGE OF NEW EXTENDED EXPERIMENTAL WINDOWS OFFERED BY INTENSIFIED REACTORS AND MINIMIZE THE NUMBER OF EXPERIMENTS. THIS METHOD CONSIDERS A PREDICTIVE PHENOMENOLOGICAL MODEL WHICH DESCRIBES REACTIONS, HEAT AND MASS-TRANSFER PROCESSES IN VARIOUS REACTORS. THE ITERATIVE METHOD ALTERNATES EXPERIMENTAL STEPS AND PREDICTION STEPS. DURING THE PREDICTION STEP, USED FOR DATA ACQUISITION, THE MODEL PROPOSES, USING THE D-OPTIMAL CRITERION, THE OPTIMAL EXPERIMENTS WHICH GIVE THE MAXIMAL INFORMATION ON THE MODEL PARAMETERS. AFTER EACH EXPERIMENT, THE STRATEGY DECIDES WHETHER A NEW ITERATION IS REQUIRED. IF THE PARAMETERS ACCURACY IS JUDGED SUFFICIENT, THEN THE STRATEGY LOOKS FOR OPTIMAL OPERATING CONDITIONS IN THE AVAILABLE REACTORS, WHICH MAXIMIZE SEVERAL PROCESS PERFORMANCE CRITERIA (E.G. YIELD, COST, OPERATION TIME, ETC.), AND PROPOSES OPTIMAL OPERATING CONDITIONS. IF THE NEW EXPERIMENTAL RESULTS ARE NOT VALIDATED, A NEW GLOBAL ITERATIVE STEP IS PERFORMED TO IMPROVE THE ACCURACY OF THE MODEL AND PROPOSE BETTER OPERATING CONDITIONS.

TO DEMONSTRATE ITS PERFORMANCE, THE STRATEGY HAS BEEN APPLIED TO A CLASSICAL PARALLEL-CONSECUTIVE REACTION SCHEME. A SPECIFIC EXPERIMENTAL SET-UP BASED ON THREE DIFFERENT REACTORS (BATCH, SEMI-BATCH AND PLUG-FLOW REACTORS) HAS BEEN DEVELOPED. THIS WORK HIGHLIGHTS THE INTEREST OF USING SEVERAL REACTORS FOR DATA ACQUISITION, DEMONSTRATES THE IMPORTANCE OF DATA ACQUISITION FOR PROCESS PERFORMANCE OPTIMIZATION AND EMPHASIZES THE SUPERIORITY OF MECHANISTIC MODELS ON BLACK-BOX MODELS FOR THE CHOICE OF REACTOR TECHNOLOGIES WHILE MINIMIZING EXPERIMENTATION STEP.
