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

Ecole Doctorale SIMPPÉ (SCIENCES ET INGÉNIERIE DES MOLÉCULES, DES PRODUITS, DES PROCÉDÉS, ET DE L'ÉNERGIE)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE l'UNIVERSITE DE LORRAINE

Mention: «Génie des Procédés, des Produits et des Molécules»

par **Zhengkun JIANG**

Développement et implémentation d'une stratégie systématique de développement de modèle se basant sur la planification optimale d'expériences

Le 22/11/2019

Membres du jury:

Rapporteurs:	Mme. Isabelle PITAULT	Chargée de recherche, LAGEP, Lyon
	M. Michel CABASSUD	Professeur, Université de Toulouse, Toulouse
Examinateurs:	Mme. Caroline GENTRIC	Professeur, Université de Nantes, Saint-Nazaire
	M. Laurent FALK	Directeur de recherche, LRGP, Nancy
	M. Jean-François PORTHA	Maître de Conférences, Université de Lorraine, Nancy
	M. Jean-Marc COMMENGE	Professeur, Université de Lorraine, Nancy, Directeur de thèse
Membres invités:	M. Régis PHILIPPE	Chargé de recherche, LGPC, Lyon
	M. Sébastien LEVENEUR	Maître de Conférences, INSA Rouen, Rouen

Abstract

Adequate and accurate models describing quantitatively the syntheses of fine and pharmaceutical chemicals are essential to optimize the performances of chemical processes. However, it is difficult, time consuming and experimentally expensive to develop such models. Appropriate, efficient and systematic strategies for model development are therefore required. In this context, the aims of this work consist in methodological development, numerical implementation and experimental validation of a systematic model-development strategy.

In the first stage of this work, a methodologically systematic model-development strategy, consisting of initial data acquisition, model development, model identification, model validation and model refining modules, is developed. In the initial data acquisition module, preliminary experiments are designed and performed to provide the basic information for the initial model development. Module development module is composed of three steps: model structure development, model structure analysis and model parameter development. Model structure development is based on the reaction network proposed within the reaction supernetwork containing all feasible chemical reactions and mass transfers. For model identification, validation and refining, the model-based experimental design is performed by taking into consideration several reactors, which enlarges the explored experimental windows.

In the second stage of this work, in order to facilitate the application of the strategy, a software, integrating model parameter estimation, model evaluation, model-based experimental design for model refining and performance optimization, is developed using MATLAB R2014a. The initial version of the software is suitable for the liquid-phase reaction systems, 4 ideal reactors are taken into consideration: batch stirred-tank reactor, semi-batch stirred-tank reactor, continuous stirred-tank reactor and continuous tubular reactor.

In the third stage of this work, the feasibilities and generalities of the developed strategy and strategy-based software are demonstrated with two experimental case studies, relating to the valorization of sunflower oil, namely, NaOH-catalyzed ethanolysis of sunflower oil and epoxidation of sunflower oil by performic acid generated in situ.

Keywords: model development, reaction supernetwork, experimental design, multi-equipment strategy, ethanolysis, epoxidation

Résumé

Des modèles adéquats et précis décrivant quantitativement les synthèses de produits chimiques fins et pharmaceutiques sont essentiels pour optimiser les performances de procédés chimiques. Cependant, il est difficile, expérimentalement long et coûteux de développer de tels modèles. Des stratégies appropriées, efficaces et systématiques servant au développement du modèle sont donc nécessaires. Dans ce contexte, les objectifs de ce travail consistent au développement méthodologique, à l'implémentation numérique et à la validation expérimentale d'une stratégie systématique de développement du modèle.

Lors de la première phase de ce travail, une stratégie méthodologiquement systématique servant au développement du modèle comprenant les modules d'acquisition initiale de données, de développement, d'identification, de validation et d'affinage du modèle, est développée. Dans le module d'acquisition initial de données, des expériences préliminaires sont planifiées et effectuées pour fournir les informations de base pour le développement initial du modèle. Le module de développement du modèle est composé de trois étapes: le développement de la structure du modèle, l'analyse de la structure du modèle et le développement des paramètres du modèle. La structure du modèle est développée en se basant sur le réseau réactionnel extrait du super-réseau réactionnel contenant toutes les réactions chimiques et les transferts de matière faisables. Pour l'identification, la validation et le raffinage du modèle, la planification expérimentale basée sur le modèle est effectuée en prenant en compte plusieurs réacteurs, ce qui élargit les fenêtres expérimentales.

Lors de la deuxième phase de ce travail, afin de faciliter l'application de la stratégie, un logiciel, intégrant l'estimation des paramètres, l'évaluation du modèle et la planification expérimentale pour l'affinage du modèle et l'optimisation des performances, est développée en utilisant Matlab R2014a. La version initiale du logiciel convient aux systèmes réactionnels liquides, quatre réacteurs idéaux sont pris en compte: réacteur fermé en cuve agitée, réacteur semi-fermé en cuve agitée, réacteur continu en cuve agitée et réacteur tubulaire continu.

Lors de la troisième phase de ce travail, les faisabilités et généralités de la stratégie et du logiciel basé sur la stratégie sont validées par deux cas d'étude expérimentaux, relatifs à la valorisation de l'huile de tournesol. Il s'agit de l'éthanolyse de l'huile de tournesol catalysée par NaOH et de l'époxydation de l'huile de tournesol par acide performique généré in situ.

Mots clés: développement de modèle, super-réseau réactionnel, planification expérimentale, stratégie multi-équipements, éthanolyse, époxydation.

Nomenclature

Abbreviations

A	Alcohol
AMA	Alkali-Metal Alkoxide
AMH	Alkali-Metal Hydroxide
AMSFA	Alkali-Metal Salt of Fatty Acid
BSTR	Batch Stirred-Tank Reactor
CA	Carboxylic Acid
CO_2	Carbon dioxide
CPOC	Constant Practical Operating Condition
CSTR	Continuous Stirred-Tank Reactor
CTR	Continuous Tubular Reactor
DAE	Data Acquisition Experiment
DAEMD	Data Acquisition Experiments for Model Discrimination
DAEPPI	Data Acquisition Experiments for Parameter Precision Improvement
DB	Double-Bond-containing compound (Chapter 3) / pseudo species (Chapter 5)
DG	Di-Glyceride
DiC38	Dinonadecanoin
Dim	Dimer (Chapter 3) / Pseudo Dimer (Chapter 5)
DL	Di-Linolein
DO	Di-Olein
DP	Di-Palmitin
E	Ethanol
EL	Ethyl Linoleate

EO	Ethyl Oleate
Ep	Epoxide-containing compound (Chapter 3) / pseudo species (Chapter 5)
EP	Ethyl Palmitate
EVO	Epoxidized Vegetable Oil
FA	Formic Acid
FAAE	Fatty Acid Alkyl Ester
FAEE	Fatty Acid Ethyl Ester
FFA	Free Fatty Acid
G	Glycerol
GC	Glycerol Carbonate
HP	Hydrogen Peroxide
IS	Internal Standard
Ket	Ketone-containing compound (Chapter 3) / pseudo species (Chapter 5)
LLLPTC	Liquid-Liquid-Liquid Phase Transfer Catalysis
LLPTC	Liquid-Liquid Phase Transfer Catalysis
MBED	Model-Based Experimental Design
MDS	Model-Development Strategy
M_F	Fisher Information Matrix
MG	Mono-Glyceride
ML	Mono-Linolein
MonoC19	Mononononadecanoin
MO	Mono-Olein
MP	1, Mono-Palmitin; 2, Mass Percentage
MUT	Model Under Test
NaOH	Sodium hydroxide
O_2	Oxygen

OED	Optimal Experimental Design
ODE	Ordinary Differential Equation
OH	Hydroxyl-containing compound (Chapter 3) / pseudo species (Chapter 5)
OH_{FA}	Formiate-containing OH
OH_{HP}	Perhydroxyl-containing OH
OH_{PFA}	Performiate-containing OH
OH_{W}	Diol-containing OH
PE	Preliminary Experiment
PAA	PerAcetic Acid
PAE	Parameter to be Actually Estimated
PCA	PerCarboxylic Acid
PFA	PerFormic Acid
POC	Practical Operating Condition
PTC	Phase Transfer Catalysis
RNUT	Reaction Network Under Test
SA	Sulfuric Acid
SBSTR	Semi-Batch Stirred-Tank Reactor
SMUT	Set of Models Under Test
SSR	Sum of Squares of Residuals
STR	Stirred-Tank Reactor
TFA	Target Factor Analysis
TG	Tri-Glyceride
TL	Tri-Linolein
TO	Tri-Olein
TP	Tri-Palmitin
TriC57	Trinonadecanoin

Trim	Trimer (Chapter 3) / Pseudo Trimer (Chapter 5)
TR	Tubular Reactor
UMV	Unconventional Measurable Variable
VEPO	Validation Experiment for Performance Optimization
VPOC	Variable Practical Operating Condition
W	Water
WSSR	Weighted Sum of Squares of Residuals

Greek symbols

γ	kinetic order
Δh	molar reaction enthalpy
$\Delta\widehat{\boldsymbol{\theta}}$	vector of parameter confidence intervals
$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$	parameter variance-covariance matrix
$\boldsymbol{\Sigma}_{\hat{\boldsymbol{y}}}$	model prediction variance-covariance matrix
$\boldsymbol{\Sigma}_{\mathbf{y}}$	variance-covariance matrix of the residuals
$\overline{\boldsymbol{\Sigma}_{\mathbf{y}}}$	average variance-covariance matrix of the residuals
θ	fraction of the sites occupied
$\boldsymbol{\theta}$	vector of model parameters
$\widehat{\boldsymbol{\theta}}$	vector of parameter estimates
$\boldsymbol{\Theta}$	space of model parameters
$\boldsymbol{\xi}$	matrix of operating conditions
$\boldsymbol{\Xi}$	space of operating conditions
ν	stoichiometric coefficient
τ	experiment duration
φ	criterion
χ	molar fraction

Latin symbols

A	frequency factor
c	concentration
Cp	heat capacity
Ea	activation energy
f	model function
F	molar flow rate
G	Jacobian matrix
h	function of parameter variance-covariance matrix
H	enthalpy
J	1, mass-transfer rate; 2, performance function
k	reaction rate constant
K	overall mass-transfer coefficient
L	adsorption rate
m	1, distribution coefficient ; 2, mass
M	molar mass
M_T	transformation matrix
n	mole number
n_{exp}	number of experimental runs
n_{oc}	number of operating conditions
n_p	number of phases
n_{perf}	number of performances
n_r	number of chemical reactions
n_{resp}	number of measured response variables
n_s	number of species

n_{sp}	number of sampling points
n_{sv}	number of time-dependent state variables
n_u	number of parameters
o	objective value
q	volumetric flow rate
Q	heat flow rate
Q	sensitivity matrix
r	reaction rate
S	heat transfer area
t	1, time; 2, t-value
T	temperature
U	global heat-transfer coefficient
v	1, molar volume; 2, veto value
V	volume
x	vector of time-dependent state variables
$\dot{\mathbf{x}}$	vector of first derivatives of time-dependent state variables
X	conversion
y	vector of measured response variables
$\hat{\mathbf{y}}$	vector of measurable variables predicted by the model
Y	selectivity

Subscripts

e	e^{th} experiment
eq	equivalent
f	f^{th} sample
i	i^{th} species

j	j^{th} reaction
k	k^{th} phase
l	l^{th} phase
max	maximal
min	minimal
p	p^{th} performance
q	q^{th} reactor
r	1, r^{th} measured response variable; 2, reactor
ref	reference
rm	reaction mixture
s	s^{th} measured response variable
sys	system
u	u^{th} parameter

Superscripts

a	adsorption
ac	accumulation
aq	aqueous phase
d	desorption
exch	exchange
in	inlet of the reactor
int	initial
n	normalized
opt	optimal
org	organic phase
out	outlet of the reactor

p	production

Table of Contents

Abstract i

Résumé ii

Nomenclature iii

Table of Contents xi

List of Tables xviii

List of Figures xxi

General introduction 1

Chapter 1. Literature survey on model-development strategies 3

1.1 Thesis object 3

1.2 Model development 5

1.2.1 Model structure development 5

1.2.1.1 Target factor analysis technique 6

1.2.1.2 Reaction supernetwork 6

1.2.1.3 Automated system 7

1.2.2 Model structure analysis 8

1.2.3 Model parameter estimation 9

1.2.3.1 Model parameter estimation criteria 9

1.2.3.2 Model parameter estimation methods 10

1.3 Model evaluation 12

1.3.1 Model adequacy evaluation 12

1.3.2 Model accuracy evaluation 12

1.4 Post-evaluation 13

1.4.1 Model identification 13

1.4.2 Model refining 16

1.4.3 Model validation 17

1.5 Thesis context and objectives 18
References 19
Chapter 2. Systematic strategy for model development 23
2.1 Strategy description 23
2.1.1 Model development module and initial data acquisition module 26
2.1.2 Model identification module 27
2.1.3 Model validation module 28
2.1.4 Model refining module 29
2.2 Model structure development 29
2.2.1 Reaction system description 30
2.2.1.1 Reaction supernetwork and reaction network 30
2.2.1.2 Rate law expression 31
2.2.2 Reactor type 33
2.3 Model parameter development 34
2.3.1 Model parameter reformulation 34
2.3.2 Model parameter pre-assignment 35
2.3.3 Model parameter estimation 35
2.4 Model evaluation 35
2.5 Model-based experimental design 36
2.6 Optimization algorithm 37
2.7 Conclusion 38
References 38
Chapter 3. Preparation works for strategy application 40
3.1 Software description 40
3.1.1 Module A: Project initialization 42
3.1.1.1 Submodule A1: Define project name 42
3.1.1.2 Submodule A2: Characterize reactors 42

3.1.1.3 Submodule A3: Define species 42
3.1.1.4 Submodule A4: Define performance criteria 43
3.1.2 Module B: Model structure development 43
3.1.3 Module C: Experimental pre-design 44
3.1.3.1 Submodule C1: Specify constraints on operating conditions 44
3.1.3.2 Submodule C2: Define measurable variables 44
3.1.4 Module D: Data acquisition 45
3.1.5 Module E: Model parameter development 45
3.1.6 Module F: Model evaluation 45
3.1.7 Module G: Results 46
3.2 Screening of validation reaction systems 46
3.2.1 Screening criteria 46
3.2.2 Candidate reaction systems 47
3.2.2.1 Reaction system 1: Acid perhydrolysis of carboxylic acid 47
3.2.2.2 Reaction system 2: First step of lorcaserin synthesis 47
3.2.2.3 Reaction system 3: Saponification of methyl benzoate by soda 48
3.2.2.4 Reaction system 4: Homogeneous alkali-catalyzed alcoholysis of vegetable oil 49
3.2.2.5 Reaction system 5: Epoxidation of vegetable oil by percarboxylic acid generated in situ 51
3.2.2.6 Reaction system 6: Liquid-liquid phase transfer catalysis 54
3.2.2.7 Reaction system 7: Liquid-liquid-liquid phase transfer catalysis 54
3.2.2.8 Reaction system 8: Synthesis of glycerol carbonate from glycerol and dimethyl carbonate 55
3.2.3 Validation reaction systems 55
3.2.3.1 Case study 1: NaOH-catalyzed ethanolysis of sunflower oil 56
3.2.3.2 Case study 2: Epoxidation of sunflower oil by performic acid generated in situ 57
3.3 Conclusion 57
Reference 58

Chapter 4. Case study 1: NaOH-catalyzed ethanolysis of sunflower oil 62
4.1 Materials and methods 62
4.1.1 Materials 62
4.1.2 Set-up 63
4.1.3 Experimental protocols 64
4.1.4 Analytical methods 64
4.1.4.1 Mass percentages of esters in the treated light phase 64
4.1.4.2 Mass percentages of glycerides in the treated light phase 66
4.2 Strategy application 68
4.2.1 Stage I: Software application 68
4.2.1.1 Submodule A1: Name project 68
4.2.1.2 Submodule A2: Characterize reactors 69
4.2.1.3 Submodule A3: Define species 69
4.2.1.4 Submodule A4: Define performance criteria 70
4.2.1.5 Submodule B1: Develop reaction supernetwork 71
4.2.1.6 Submodule B2: Propose candidate reaction networks 72
4.2.1.7 Submodule B3: Select model parameters to be estimated 72
4.2.1.8 Submodule C1: Specify constraints on operating conditions 72
4.2.1.9 Submodule C2: Define measurable variables 76
4.2.1.10 Submodule D1: Design data acquisition experiments 77
4.2.2 Stage II: PEs 78
4.2.3 Stage III: Software application 79
4.2.3.1 Submodule D2: Input experimental data 79
4.2.3.2 Submodule E1: Specify constraints on model parameters 80
4.2.3.3 Submodule E2: Estimate model parameters 80
4.2.3.4 Submodule F1: Model adequacy pre-test 82
4.2.3.5 Submodule F2: Model accuracy pre-test 83

4.2.3.6 Submodule D1: Design data acquisition experiment....84
4.2.4 Stage IV: First DAEPPI....85
4.2.5 Stage V: Software application....85
4.2.5.1 Submodule D2: Input experimental data....85
4.2.5.2 Submodule E2: Estimate model parameters....85
4.2.5.3 Submodule F1: Model adequacy pre-test....86
4.2.5.4 Submodule F2: Model accuracy pre-test....86
4.2.5.5 Submodule D1: Design data acquisition experiment....86
4.2.6 Stage VI: Second DAEPPI....87
4.2.7 Stage VII: Software application....87
4.2.7.1 Submodule D2: Input experimental data....87
4.2.7.2 Submodule E2: Estimate model parameters....87
4.2.7.3 Submodule F1: Model adequacy pre-test....87
4.2.7.4 Submodule F2: Model accuracy pre-test....87
4.2.8 Stage VIII: VEPO....88
4.2.9 Stage IX: Software application....88
4.2.9.1 Submodule F3: Model adequacy and accuracy post-tests....88
4.2.9.2 Module G: Results....88
4.3 Strategy feasibility analysis....89
4.4 Software practicability analysis....89
4.5 Conclusion....90
References....91
Chapter 5. Case study 2: Epoxidation of sunflower oil by performic acid generated in situ....92
5.1 Materials and methods....92
5.1.1 Materials....92
5.1.2 Experimental protocols....93
5.1.3 Analytical methods....94

5.1.3.1 Concentration of hydrogen ion in the aqueous phase 95
5.1.3.2 Concentration of hydrogen peroxide in the aqueous phase 95
5.1.3.3 Concentration of double bond in the organic phase 96
5.1.3.4 Concentration of epoxide in the organic phase 97
5.2 Strategy application 97
5.2.1 Stage I: Software application 98
5.2.1.1 Submodule A1: Name project 98
5.2.1.2 Submodule A2: Characterize reactors 98
5.2.1.3 Submodule A3: Define species 98
5.2.1.4 Submodule A4: Define performance criteria 101
5.2.1.5 Submodule B1: Develop reaction supernetwork 102
5.2.1.6 Submodule B2: Propose candidate reaction networks 104
5.2.1.7 Submodule B3: Select model parameters to be estimated 106
5.2.1.8 Submodule C1: Specify constraints on operating conditions 106
5.2.1.9 Submodule C2: Design measurable variables 109
5.2.1.10 Submodule D1: Design data acquisition experiments 109
5.2.2 Stage II: Preliminary experiments 110
5.2.3 Stage III: Software application 112
5.2.3.1 Submodule D2: Input experimental data 112
5.2.3.2 Submodule E1: Specify constraints on model parameters 112
5.2.3.3 Module E2: Estimate model parameters 114
5.2.3.4 Submodule F1: Model adequacy pre-test 115
5.2.3.5 Submodule F2: Model accuracy pre-test 118
5.2.3.6 Submodule D1: Design data acquisition experiment 120
5.2.4 Stage IV: First DAEPPI 121
5.2.5 Stage V: Software application 122
5.2.5.1 Submodule D2: Input experimental data 122

5.2.5.2 Submodule E2: Estimate model parameters 122
5.2.5.3 Submodule F1: Model adequacy pre-test 122
5.2.5.4 Submodule F2: Model accuracy pre-test 123
5.2.5.5 Submodule D1: Design data acquisition experiment 123
5.2.6 Stage VI: Second DAEPPI 124
5.2.7 Stage VII: Software application 124
5.2.7.1 Submodule D2: Input experimental data 124
5.2.7.2 Submodule E2: Estimate model parameters 124
5.2.7.3 Submodule F1: Model adequacy pre-test 125
5.2.7.4 Submodule F2: Model accuracy pre-test 125
5.2.8 Stage VIII: VEPO 125
5.2.9 Stage IX: Software application 125
5.2.9.1 Submodule F3: Model adequacy and accuracy post-tests 125
5.3 Strategy feasibility analysis 127
5.4 Software practicability analysis 128
5.5 Conclusion 129
References 129
Conclusion and Perspectives 131
Conclusion 131
Perspectives 132
Appendix 1: General mass, volume and heat balances 134
General mass balance 134
General volume balance 135
General heat balance 135
Appendix 2: Experimental responses for the case study 1 137
Appendix 3: Supplementary information for the case study 2 138

List of Tables

Table 1.1: Comparison of different models (↑ = increase; ↓ = decrease). 5
Table 1.2: Overview of the improved criteria for discrimination of two models in literature. 15
Table 1.3: Systematicness assessment for various model-development strategies..... 19
Table 2.1: Concrete optimization algorithms and the corresponding Matlab functions. 38
Table 3.1: Operating conditions in general form (× = needs to be defined). 43
Table 3.2: Submodules following and followed by various model-evaluation submodules. 45
Table 3.3: Candidate reaction systems..... 47
Table 3.4: Reaction supernetwork for the homogeneous alkali-catalyzed alcoholysis of vegetable oil. 50
Table 3.5: Survey on the reported reaction networks for the homogeneous alkali-catalyzed alcoholysis of vegetable oil. 51
Table 3.6: Reaction supernetwork for the epoxidation of vegetable oil by percarboxylic acid generated in situ. 53
Table 3.7: Survey on the reported reaction networks for the epoxidation of vegetable oil by percarboxylic acid generated in situ. 54
Table 3.8: Result of candidate reaction system assessment in terms of screening criteria (+ = expected; ± = neutral; − = unexpected). 56
Table 4.1: Information about the reagents used for the case study 1. 62
Table 4.2: Characteristics for each reactor. 64
Table 4.3: Conditions for gas chromatography analysis of esters in the treated light phase. 65
Table 4.4: Identified fatty acid ethyl esters in sample. 66
Table 4.5: Conditions for gas chromatography analysis of glycerides in the treated light phase. 67
Table 4.6: Information about all possible species in the reaction system for the case study 1. 70
Table 4.7: Performance criteria for the case study 1. 70
Table 4.8: Reaction supernetwork for the case study 1. 71
Table 4.9: Reaction networks under test for the case study 1. 72
Table 4.10: Bounds on practical operating conditions for each considered reactor. 74
Table 4.11: Bounds on practical operating variables for each considered reactor. 75
Table 4.12: Bounds on operating conditions in general form for each considered reactor. 75
Table 4.13: Definition of the low and high levels for each considered factor. 77
Table 4.14: Fractional factorial experimental design for case study 1. 78

Table 4.15: Designed operating variables of the PEs for case study 1. 78
Table 4.16: Practical operating variables of the PEs for the case study 1. 79
Table 4.17: Evolution of parameter estimations for the case study 1. 81
Table 4.18: Results of model-based experimental design for the VEPO. 84
Table 4.19: Results of model-based experimental design for the 1st DAEPPI. 85
Table 4.20: Result of model-based experimental design for the 2nd DAEPPI. 87
Table 4.21: Comparison of pre-exponential factors and activation energies obtained from this work and from literature. 89
Table 4.22: Check of parameter estimates in this work and parameter confidence intervals derived from literature. 89
Table 5.1: Information about the reagents used for the case study 2. 93
Table 5.2: Survey on different combination modalities of solution constituents (SA = Sulfuric Acid). 93
Table 5.3: Information about various solutions used for analyses. 95
Table 5.4: Diversity of the products generated from tri-olein. 99
Table 5.5: Information about all possible species in the reaction system for the case study 2. 101
Table 5.6: Alternative performance criteria for the case study 2. 102
Table 5.7: Reaction supernetwork for the case study 2. 103
Table 5.8: Reaction networks under test for the case study 2. 104
Table 5.9: Model parameter reformulation for the case study 2. 106
Table 5.10: Bounds on practical operating conditions of each considered reactor for the case study 2. 106
Table 5.11: Bounds on operating conditions in general form of each considered reactor for the case study 2. 107
Table 5.12: Equality and inequality constraints among operating conditions in general form for the case study 2. 108
Table 5.13: Definition of the low and high levels for each considered experimental factor. 110
Table 5.14: Fractional factorial experimental design result for the case study 2 (+ = high-level; ± = average-level; - = low-level). 110
Table 5.15: Practical operating conditions of the PEs for the case study 2. 111
Table 5.16: Bounds on model parameters for the case study 2. 113
Table 5.17: Equality and inequality constraints among model parameters for the case study 2. 113
Table 5.18: Evolution of parameter estimations for the case study 2. 114
Table 5.19: Comparison of parameter estimation criteria. 115

Table 5.20: Results of model-based experimental design for the VEPO. 119
Table 5.21: Results of model-based experimental design for the DAEPPI. 120
Table 5.22: Various time periods of the experiment carried out in the SBSTR. 121
Table 5.23: Comparison of kinetic constants obtained from this work and from literatures.....127
Table A2.1: Experimental responses of the case study 1.....137
Table A3.1: Operating conditions in general form of the PEs for the case study 2.....138
Table A3.2: Experimental responses of the case study 2.....139
Table A3.3: Parameter estimation results using the experimental data of the PEs.....141

List of Figures

Figure 1.1: Workflow for the automated system for chemical reaction network elucidation. 8
Figure 1.2: Schematic of stepwise estimation (Brendel et al., 2006). 11
Figure 1.3: Geometrical interpretation of the standard criteria for the experimental design (Franceschini and Macchietto, 2008).....16
Figure 2.1: Systematic model-development strategy based on model-based experimental design. 25
Figure 2.2: General procedure for model development. 26
Figure 2.3: Model development module (in the left dotted box) and initial data acquisition module (in the right dotted box).....26
Figure 2.4: Model identification module (in dotted box). 27
Figure 2.5: Model validation module (in dotted box).....28
Figure 2.6: Model refining module (in dotted box). 29
Figure 3.1: Software workflow. 41
Figure 4.1: Experimental set-up.....63
Figure 4.2: Chromatographic analysis of fatty acid ethyl esters in the 3rd sample of the 4th PE. 65
Figure 4.3: Chromatographic analysis of glycerides in the 3rd sample of the 4th PE. 67
Figure 4.4: Operational sketch of strategy application for the ethanolysis case study. 68
Figure 4.5: Interface for the definition of the 4th performance criterion.....71
Figure 4.6: Illustration of the relationship between practical feeds and virtual feeds. 76
Figure 4.7: Interface for the selection of the measurable variables for the case study 1. 77
Figure 4.8: Evolution of the experimental responses for the 1st PE carried out at 60°C, NaOH concentration of 0.50 wt.% (per oil weight), ethanol-to-oil molar ratio of 6:1, using the BSTR.....79
Figure 4.9: Evolution of the various measurable variables for the 1st PE carried out at 60°C, NaOH concentration of 0.50 wt.% (per oil weight), ethanol-to-oil molar ratio of 6:1, using BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$). 82
Figure 4.10: Evolution of the various measurable variables for the 1st sample of the 5th PE carried out at 60°C, NaOH concentration of 2.34 wt.% (per oil weight), ethanol-to-oil molar ratio of 28:1, using CTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).....82
Figure 4.11: Concentration profiles for the VEPO using the BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$). 88

Figure 5.1: Molecule structure of the product N°8. 100
Figure 5.2: Molecule structure of the product N°19. 100
Figure 5.3: Interface for the customization of the chemical reaction orders for the RNUT6: (a) interface before the definition, (b) interface after the definition. 105
Figure 5.4: Illustration of the relationship between practical feeds and virtual feeds for the case study 2. 109
Figure 5.5: Interface for the selection of the measurable variables for the case study 2. 109
Figure 5.6: Evolution of the experimental responses for the 4th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 55°C, using the SBSTR. 111
Figure 5.7: Evolution of the various measurable variables for the 4th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 55°C, using the SBSTR (continuous lines: responses predicted by the SMUT[RNUT2,PEs]; singular points: experimental responses). 116
Figure 5.8: Evolution of the various measurable variables for the 3rd sample of the 7th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 75°, using the CTR (continuous lines: responses predicted by the SMUT[RNUT2,PEs]; singular points: experimental responses). 117
Figure 5.9: Evolution of the various measurable variables for the 1st DAEPPI (continuous lines: responses predicted by the SMUT[RNUT3,PEs+1st DAEPPI]; singular points: experimental responses). 123
Figure 5.10: Experimental responses for the VEPO using the SBSTR (continuous lines: responses predicted by the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI]; singular points: experimental responses; vertical bars around the singular points: experimental uncertainties). 126

General introduction

Models are important tools for reliable design, optimization and scale-up of chemical synthesis processes. Two essential elements of a model are the model structure and the model parameters. According to the amount of knowledge taken into consideration during the model structure development (i.e. the so-called model grey-level), models can be divided into three categories: white-box, grey-box and black-box models.

The white-box models (also called deterministic models and mechanistic models) developed on the basis of scientific laws (e.g. kinetics, mass balances, heat balances, thermodynamics) and reasonable assumptions (e.g. the species concentrations as well as the reaction temperature in a stirred-tank reactor are homogeneous along the reaction progress), are always preferred due to their reliable prediction capacities. However, it is difficult, time consuming and experimentally expensive to develop such models; a high level of user expertise is also required. Therefore, more and more attention has been paid to the Model-Development Strategies (MDSs), particularly the MDSs based on Model-Based Experimental Design (MBED), which guide the users to develop efficiently the adequate and accurate reactor model for a given chemical synthesis.

Ideal MDS should be both methodologically and practically systematic. A MDS can be considered methodologically systematic only if it provides the solutions to solve any problem met during model development, for instance, (i) which procedure has to be followed for the model structure development? (ii) if several rival models, based on different reaction networks, are available, how to identify efficiently the model with the most adequate structure? (iii) if model parameters are not accurate enough, how to improve efficiently model parameter precision?

Meanwhile, a practically systematic MDS should have the potential not only to be applied in any chemical reaction system (i.e. both homogeneous and heterogeneous reaction systems), but also to construct simultaneously the model structures for different reactors (e.g. stirred-tank reactor, tubular reactor, fixed-bed reactor). In addition, MDS is considered as only a guideline, however, in practice, the programming works aiming at solving the optimization problems, such as model parameter estimation and MBED, are required. Therefore, a practically systematic MDS should also be easily applied by the user: concretely, it is preferred to be numerically implemented as a software (or codes), integrating all (or part of) numerical calculations, which can reduce the user works.

Various MDSs have been developed in the past two decades, none of the reported MDSs is both methodologically and practically systematic, in other words, each of these MDSs has its own limits.

In this context, the objectives of this thesis consist of: (i) developing a methodologically systematic MDS based on MBED, guiding the user to develop the adequate and accurate models for all considered reactors for a given chemical synthesis, and further to identify the optimal reactor with its associated optimal operating conditions; (ii) developing the strategy-based software to facilitate the application of the strategy; (iii) checking experimentally the feasibilities and generalities of the developed strategy and strategy-based software by two case studies.

This thesis, presenting the works done to achieve the above objectives, is composed of five chapters. The outline of each chapter is as follow:

In chapter 1, firstly, the object of this thesis, i.e. the methodologically systematic MDS based on MBED, is presented. Then, a literature review on three essential modules (i.e. the model development module, the model evaluation module, and the so-called post-evaluation module) involved in such a MDS is given. Finally, the thesis context and objectives are presented.

In chapter 2, a methodologically systematic MDS based on MBED for simultaneous model development and reactor performance optimization is presented.

In chapter 3, the preparation works for strategy application, including: (i) the developed strategy-based software, facilitating significantly the user work; (ii) the screening procedure of the validation reaction systems, used to demonstrate the feasibilities and generalities of the developed strategy and strategy-based software, are presented.

In chapter 4, the feasibilities of the developed strategy and strategy-based software are firstly demonstrated by the case study 1, i.e. NaOH-catalyzed ethanolysis of sunflower oil.

In chapter 5, the feasibilities as well as the generalities of the developed strategy and strategy-based software are further demonstrated by the case study 2, i.e. epoxidation of sunflower oil by performic acid generated in situ.

Chapter 1. Literature survey on model-development strategies

In this chapter, firstly, system definition and classification as well as model definition and classification are presented, which leads to the object of this thesis, i.e. the systematic strategy used for developing the adequate and accurate mechanistic model of a chemical reaction system.

Subsequently, three essential modules involved in a systematic model-development strategy are described in detail: the model development module (see Section 1.2), the model evaluation module (see Section 1.3), and the so-called post-evaluation module (see Section 1.4), as their names suggest, cover respectively the activities on model development, model evaluation and the ones following model evaluation.

Finally, the reported model-development strategies are assessed in terms of systematicness. The assessment result shows that none of the reported strategies is both methodologically and practically systematic and leads to the necessity of the objectives of this thesis, which can be summarized in one phrase, i.e. methodological development, numerical implementation and experimental validation of a systematic model-development strategy (see Section 1.5).

1.1 Thesis object

In chemical engineering, the system is defined as the equipment together with its corresponding inside process, translating the system inputs into the system outputs. Depending on process functions, chemical engineering systems can be classified into three categories:

1. Chemical reaction systems, i.e. what are studied in this work, corresponding to a reactor together with the corresponding chemical reactions, transforming the reactants into the products;
2. Separation systems, such as distillation systems, corresponding to a distillation column together with the distillation process, separating for instance a binary solution (e.g. a mixture of product and solvent) into ideally the pure distillate (e.g. the solvent) and the pure residue (e.g. the product);
3. Other systems, such as heat-exchange systems, corresponding to a heat exchanger together with the heat-exchange process, transferring heat from the hot solution to the cold solution heated up to the desired temperature.

Models are an important tool for system understanding and further for reliable system design and optimization. The model of a system is defined as the mathematical function associating the inputs and the outputs of the system, and can be expressed as follows:

$$\mathbf{x} = \mathbf{f}_1(\boldsymbol{\xi}, \boldsymbol{\theta}) \quad (1.1)$$

$$\mathbf{y} = \mathbf{f}_2(\boldsymbol{\xi}, \mathbf{x}) \quad (1.2)$$

$$\mathbf{z} = \mathbf{f}_3(\boldsymbol{\xi}, \mathbf{x}, \mathbf{y}) \quad (1.3)$$

where:

$\boldsymbol{\xi}$ is the system input matrix, namely, the matrix of operating conditions, such as initial loading mass of each reactant, reaction temperature, reaction time, etc.,

x, **y**, **z** are the system output matrices, corresponding to respectively the matrix of time-dependent state variables (e.g. species concentration, system temperature, etc.), the matrix of variables to be actually measured (e.g. droplet size, refractive index, system temperature, etc.) and the vector of performances (e.g. residual concentration of one or several species, selectivity of the desired product, etc.),

$\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$ are the model structure, here, composed of three sets of mathematical equations,

$\boldsymbol{\theta}$ is the vector of model parameters, such as reaction rate constants, mass-transfer coefficients, etc.

Two essential elements of a model are the model structure and the model parameters. According to the amount of knowledge taken into consideration during the model structure development, models can be divided into three types:

1. White-box models (also called deterministic models and mechanistic models) are developed on the basis of scientific laws (i.e. mass balances, heat balances, thermodynamics, etc.) and reasonable assumptions (e.g. the species concentrations as well as the reaction temperature in the stirred-tank reactor are homogeneous along the reaction progress), and further attempts to represent accurately the reaction system under investigation.
2. Black-box models forego any attempt to explain why the variables interact with each other, and simply attempt to describe the observed phenomena and quantify in a statistical sense the relationship between the operating conditions (inputs) and the experimental responses (outputs).
3. Grey-box models (also known as tendency models) converge towards a compromise between the two models mentioned above. The simplified representation allows to decrease the model-development difficulty, but reduces their explanatory power with respect to the white-box model.

Table 1.1: Comparison of different models (↑ = increase; ↓ = decrease).

Model type	Amount of considered knowledge	Difficulty of model development	Reliability of model predictions
White-box model Grey-box model Black-box model	↓	↑	↓

Table 1.1 shows the comparison of these different types of models. The logic fact, that the more knowledge taken into consideration while model structure development, the more reliable model predictions, is observed.

In order to describe a chemical reaction system, white-box models, which are considered in this work, are always preferred due to their reliable prediction and retro-prediction capacities. For instance, (i) given a set of operating conditions, these models can predict the behaviors taking place in the reaction system, which allows the user to evaluate whether the undesired phenomena, such as flammability limits, flash polymerization, are present or not; (ii) given a set of desired performances (i.e. the system outputs), these models can identify the corresponding operating conditions.

However, a white-box model (hereinafter referred to as model) is also the most difficult type to be developed. Therefore, more and more attention has been paid to methods enabling to develop efficiently an adequate and accurate model for a studied system. Model-development strategy is the result of these researches and also the object of this thesis.

1.2 Model development

The conventional model development is a sequential procedure consisting of three steps: (i) model structure development, (ii) model structure analysis, (iii) model parameter estimation.

1.2.1 Model structure development

Model structure depends on both (the type and the characteristic variables of) the reactor and the phenomena (e.g. chemical reactions, mass transfer, heat transfer, etc.) taken into consideration. For a given chemical reaction system, i.e. having specified the considered reactor as well as the reactants, the catalyst, the solvent and the desired product of the chemical synthesis, model structure development is approximatively equivalent to reaction network development regardless of rheological effects as well as those of pH, mixing, etc.

Reaction network (also called reaction pathway, reaction mechanism or reaction scheme) is defined as the set of reactions supposed to occur in the reaction system, and traditionally proposed by the

user based on his understanding of the studied reaction system. A high level of user expertise may be required to propose an adequate reaction network. In this context, some advanced techniques have been proposed to guide the user.

1.2.1.1 Target factor analysis technique

Target Factor Analysis (TFA) technique, initially proposed by Bonvin and Rippin (1990) and continuously improved by Fotopoulos et al. (1994), Amrhein et al. (1999), has been used successfully to elucidate the reaction network, without knowledge of reaction kinetics. TFA technique consists of two steps:

1. Factor analysis: determine the number of linearly independent reactions and derive an observed stoichiometric space from the experimental data;
2. Target testing: verify whether a target factor (i.e. a postulated chemical reaction network, which may stem from a priori knowledge or from deductive reasoning) is compatible with the stoichiometric space.

The major drawback of TFA technique is that target-testing work becomes very cumbersome when a large number of species are involved in the reaction system, namely, when lots of reaction networks are feasible.

1.2.1.2 Reaction supernetwork

In order to propose an adequate reaction network, one can refer to the concept of process superstructure used in process synthesis and design (Achenie and Biegler, 1990): this superstructure contains all possible alternatives of a potential process flowsheet to which the optimal solution belongs. By analogy, a reaction supernetwork including all feasible chemical reactions can be defined (Burnham et al., 2008; Zhang et al., 2015): any reaction network, including the appropriate reaction network to be finally identified and the "theoretically-right" reaction network, is a subset extracted from this reaction supernetwork. Elucidation of chemical reaction network within reaction supernetwork is a sequential procedure consisting usually of three steps:

1. Reaction supernetwork development: generate all possible reactions between all possible species (i.e. reactants, solvents, catalysts, and products) in the reaction system, according to the pre-defined criteria used for reaction generation, e.g. the reactions are at most the result of bimolecular collisions.

2. Reaction supernetwork reduction: remove from the reaction supernetwork the theoretically unfeasible reactions, e.g. those which do not satisfy the mass balance (Burnham et al., 2008) or the atomic balance (Zhang et al., 2015);
3. Extraction of reaction network: extract an adequate reaction network within the reduced reaction supernetwork through different manners, e.g. statistical inference (Burnham et al., 2008) and model discrimination (Zhang et al., 2015).

The major drawback of extracting the right reaction network within reaction supernetwork is the same as that of TFA technique, i.e. reaction network extraction work becomes increasingly cumbersome and time consuming when the number of species is large.

1.2.1.3 Automated system

Reaction network can also be elucidated through evolutionary algorithms, such as genetic programming (Koza et al., 2007), differential evolution (Searson et al., 2012) and genetic algorithm (Hii, 2017). The procedures are similar, and can be summarized in four iterative steps (illustrated in Figure 1.1):

1. Initialization: define the algorithm parameters (e.g. the number of individuals per generation, the maximum number of generations, etc.) and generate randomly the initial individuals (i.e. reaction networks) of the first generation;
2. Parameter estimation: estimate the parameters involved in the model structure constructed based on each individual and calculate the corresponding fitness function, which is used to evaluate how closely related are the predicted values with respect to the experimental data;
3. Evaluate whether the predefined termination criteria are reached. If reached, the fittest individual is identified, otherwise, new individuals have to be reproduced;
4. Reproduction: create the new individuals of the next generation, the precise individual-creation methods depending on the algorithms.

Elucidation of chemical reaction network through evolutionary algorithms can be computationally parallelized and automatized, which minimizes the human intervention; whereas the disadvantage of this method is also obvious, as the previous two methods, the computational cost becomes prohibitive, when the number of species in the reaction system increases.

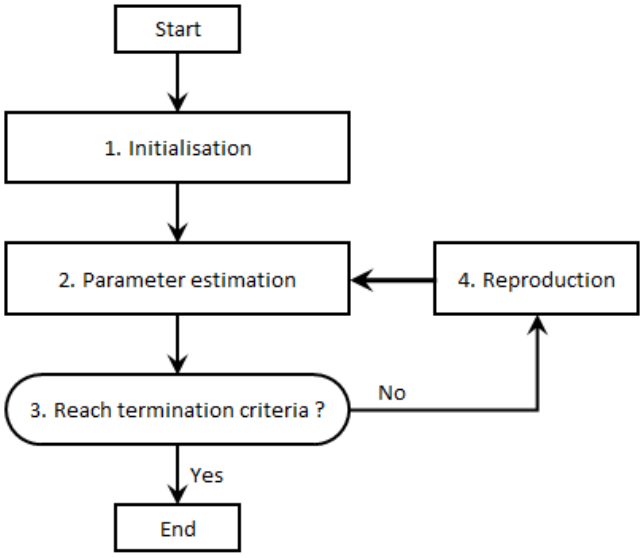


Figure 1.1: Workflow for the automated system for chemical reaction network elucidation.

1.2.2 Model structure analysis

The number of variables to be actually measured is commonly limited compared to the number of variables needed to be measured (i.e. all time-dependent state variables). Therefore, model structure analysis is required to evaluate model structure(s) in terms of identifiability (and distinguishability, if more than one candidate model structure), before any experimental activities followed by model parameter estimation. The definitions of identifiability and distinguishability are given below:

- Identifiability definition: a model is considered identifiable if, for any admissible system input matrix $\boldsymbol{\xi}$ and any two model parameter vectors $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ within the model parameter space $\boldsymbol{\Theta}$, $\mathbf{y}(\boldsymbol{\xi},\boldsymbol{\theta}_1) = \mathbf{y}(\boldsymbol{\xi},\boldsymbol{\theta}_2)$ holds only if $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$. In other words, the parameter vector $\boldsymbol{\theta}$ of the identifiable model can be uniquely identified from the given $\boldsymbol{\xi}$ and $\mathbf{y}$.
- Distinguishability definition: two models $\mathbf{y}_1$ and $\mathbf{y}_2$ are discernable if, for any admissible system input matrix $\boldsymbol{\xi}$ and any two model parameter vectors $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ within their associated parameter spaces $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ respectively, $\mathbf{y}_1(\boldsymbol{\xi},\boldsymbol{\theta}_1) \neq \mathbf{y}_2(\boldsymbol{\xi},\boldsymbol{\theta}_2)$ holds (Asprey and Macchietto, 2000; Violet, 2016).

According to whether algebraic manipulation is performed or not, the methods used for model structure analysis can be classified into two categories:

- Theoretical methods, such as the linearization method (Grewal and Glover, 1976), the power series expansion method (Pohjanpalo, 1978), the similarity transformation method (Vajda and Rabitz, 1989). For complex model structures, the tedious algebraic symbolic calculations

involved in such methods can be done with the help of computer algebra tools; the associated computational cost is high.

- Practical method, such as the method of Walter and Pronzato (1994), consists of:
  1. generate randomly a model parameter vector $\boldsymbol{\theta}$ within the model parameter space $\boldsymbol{\Theta}$;
  2. given a random admissible system input matrix, predict the corresponding fictive experimental response matrix;
  3. estimate the model parameters, verify whether the estimated model parameter vector is the same as the initially generated one. If yes, model is considered locally identifiable; otherwise, model is non-identifiable;
  4. repeat steps 1-3 in order to evaluate summarily and approximatively whether model is considered globally identifiable.

For more details about the theoretical and practical methods used for evaluating model identifiability and distinguishability, refer to the reviews given by Ollivier (1990) and by Miao et al. (2011).

1.2.3 Model parameter estimation

Once model structure has been verified structurally identifiable, the model parameters have to be estimated by following a two-stage procedure:

1. Define the model parameter estimation criterion, a function of model parameters, denoted as $\varphi(\boldsymbol{\theta})$, characterizing model adequacy, i.e. the deviation between model predictions and experimental data;
2. Seek a set of parameter estimates making the criterion defined above minimum, which can be mathematically expressed as:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \varphi(\boldsymbol{\theta}) \quad (1.4)$$

1.2.3.1 Model parameter estimation criteria

The least square criterion, i.e. the oldest criterion, is defined as the Sum of Squares of Residuals (SSR) and can be expressed as:

$$\varphi(\boldsymbol{\theta}) = \sum_{e=1}^{n_{exp}} \sum_{f=1}^{n_{sp_e}} \left[\mathbf{y}(\boldsymbol{\xi}_{e,f}) - \hat{\mathbf{y}}(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})\right]^T \left[\mathbf{y}(\boldsymbol{\xi}_{e,f}) - \hat{\mathbf{y}}(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})\right] \quad (1.5)$$

where:

n_{exp} represents the number of experimental runs,

n_{sp_e} represents the number of samples in the eth experiment,

$\mathbf{y}(\boldsymbol{\xi}_{e,f})$ is the system output vector measured under the system input vector $\boldsymbol{\xi}_{e,f}$, which is the fth system input vector of the eth experiment,

$\hat{\mathbf{y}}(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})$ is the corresponding predicted system output vector.

The parameter estimates obtained by minimizing the least square criterion may be often unsatisfactory for the following reasons (Bard, 1974):

- The various measurable variables may exhibit very different orders of magnitude. For instance, a concentration-type measurable variable, expressed in mole fraction, falls in the range 0–1; at the same time, a temperature-type measurable variable, measured in degrees centigrade, falls in the range 50–500. It clearly makes no sense to sum together squares of numbers of such disparate orders of magnitude; the residuals of the temperatures are likely to dominate those of the mole fractions and any information contained in the latter will be lost.
- Some observations may be known to be less reliable than others, it is desired that the parameter estimates are less influenced by those than by the more accurate ones.

Thus, integration of a non-negative weight factor (or matrix) into the least square criterion is necessary. The weighted least square criteria, the most widely used criteria, are defined on the basis of the least square criterion by integrating different weight factors (or matrices), and can be expressed by Eq. (1.6), if the most common weight matrix, i.e. the measurement variance-covariance matrix, denoted as Σ_y, is used.

$$\varphi(\boldsymbol{\theta}) = \sum_{e=1}^{n_{exp}} \sum_{f=1}^{n_{sp_e}} \left[\mathbf{y}(\boldsymbol{\xi}_{e,f}) - \hat{\mathbf{y}}(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})\right]^T \cdot \Sigma_y^{-1} \cdot \left[\mathbf{y}(\boldsymbol{\xi}_{e,f}) - \hat{\mathbf{y}}(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})\right] \quad (1.6)$$

Σ_y is not always known, it can be estimated but it requires a lot of repetitive data. The maximum-likelihood-based criterion makes it possible to overcome this difficulty. For more details about the different model parameter estimation criteria, refer to the review given by Bard (1974).

1.2.3.2 Model parameter estimation methods

Model parameters can be estimated conventionally in a one-step manner or sequentially over several steps. In one-step model parameter estimation method, the most used method, as the name suggests, all parameters are simultaneously estimated. However, there are several disadvantages of this method, for instance:

- if an incorrect model structure (e.g. incorrect reaction network or incorrect reaction rate expressions) is assumed, an erroneous overall model prediction is obtained, therefore, the origin of the model error is very difficult to determine;
- computational cost may become prohibitive if the number of candidate model structures or the model structure complexity (e.g. the number of reactions) increases.

Alternatively, in stepwise model parameter estimation method, also known as incremental estimation approach, the model parameter estimation problem is decomposed into a sequence of subproblems. To illustrate this approach, a three-step procedure used for an isothermal homogeneous reaction system is used (see Figure 1.2):

1. the reaction fluxes for the various species are estimated on the basis of balance equations and experimental data, e.g. concentration measurements;
2. given reaction stoichiometries (elucidated using TFA technique), the reaction rates are calculated from the reaction fluxes previously determined without knowledge of reaction kinetic laws;
3. the kinetic laws, i.e., the dependencies of the reaction rates on concentrations, are constructed by selecting the best model structure from a set of model candidates.

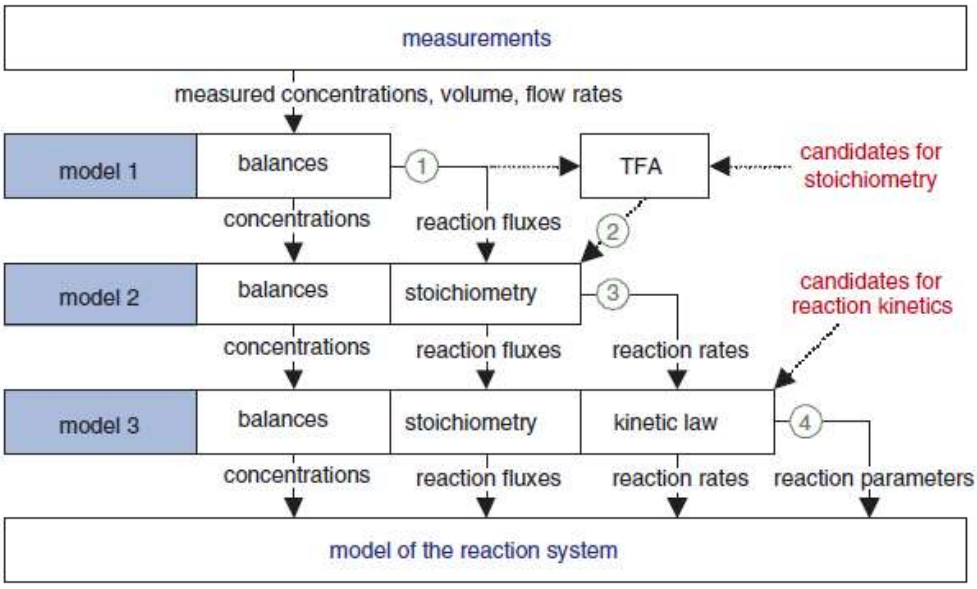


Figure 1.2: Schematic of stepwise estimation (Brendel et al., 2006).

The incremental approach is shown to be both efficient and flexible and has been successfully used for a homogeneous reaction system (Brendel et al., 2006) and a heterogeneous reaction system (Michalik et al., 2009). However, data-rich measurement techniques, such as IR spectroscopy and Raman spectroscopy, are always required.

1.3 Model evaluation

Following model development (see Section 1.2), it is of primary importance to statistically evaluate the adequacy of the model structure and the accuracy of the model parameters.

1.3.1 Model adequacy evaluation

In order to evaluate model structure adequacy, namely, whether the model can explain the experimental data in a satisfactory manner, a lack-of-fit test is usually used. Assuming that the measurement errors follow a zero mean normal distribution, according to whether Σ_y is known or not, different statistical tests are used:

- If Σ_y is known, a χ^2-test (chi-squared test), as outlined for instance in Donckels et al. (2009) and Violet (2016), is used. The SSR divided by Σ_y, known as the Weighted SSR (WSSR) (Eq. (1.6)), characterizing the difference between measured and predicted responses, is a sample of the χ^2-distribution with $(\sum_{e=1}^{n_{exp}} n_{sp_e} - n_u)$ degrees of freedom (n_u is the number of model parameters). A model is considered inadequate, if its WSSR is significantly higher than its associated χ^2-value.
- If Σ_y is not known, a F-test (Fisher test), as outlined for instance in Franceschini and Macchietto (2008), is used. The WSSR, which has to be firstly estimated through repeating measurements, is a sample of the F-distribution with $(\sum_{e=1}^{n_{exp}} n_{sp_e} - n_u - n_m)$ and $(n_m - 1)$ degrees of freedom (n_m is the number of repeated measurements). A model is considered inadequate if the estimated WSSR, i.e. the ratio between its SSR and the estimated variance, is significantly higher than its associated F-value.

1.3.2 Model accuracy evaluation

In order to evaluate model parameter accuracy or precision, e.g. whether the confidence interval of each parameter is accurate enough, the $n_u \times n_u$–dimensional parameter variance-covariance matrix, denoted as $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$, quantifying the uncertainty on the parameter estimates, is used and can be approximated using the Fisher Information Matrix M_F (Franceschini and Macchietto, 2008):

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_e, \cdots, \boldsymbol{\xi}_{n_{exp}}\right) = \left[\sum_{e=1}^{n_{exp}} M_F\left(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_e\right)\right]^{-1} \quad (1.7)$$

with:

$$M_F(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_e) = \sum_{r=1}^{n_{resp}} \sum_{s=1}^{n_{resp}} \sigma_{r,s} \mathbf{Q}_{r,e}{}^T \mathbf{Q}_{s,e} \quad (1.8)$$

where n_{resp} is the number of measurable variables, $\sigma_{r,s}$ is the inverse of the covariance between the rth and s^{th} measurable variables, $\mathbf{Q}_{r,e}$ represents the $n_{sp_e} \times n_u$–dimensional dynamic sensitivity matrix of the rth measurable variable evaluated at different sampling times of the eth experiment and is written as:

$$\mathbf{Q}_{r,e} = \begin{bmatrix} \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_1} & \cdots & \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_{n_u}} \leftarrow t_{e,1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_1} & \cdots & \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_{n_u}} \leftarrow t_{e,n_{sp_e}} \end{bmatrix} \quad (1.9)$$

From $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$ (Eq. (1.7)), the confidence interval of the uth parameter is calculated by:

$$\Delta\hat{\theta}_u = t\sqrt{\Sigma_{\hat{\theta}_{uu}}} \quad (1.10)$$

where t is the t-value tested by reference to a t-distribution with the degree of freedom corresponding to the total number of samples minus the number of model parameters, $\Sigma_{\hat{\theta}_{uu}}$ is the variance of the uth parameter, i.e., the uth diagonal element of $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$.

1.4 Post-evaluation

Following model evaluation in terms of (adequacy of) structures and (accuracy of) parameters (see Section 1.3), there are three potential problems to face: (i) if several rival models (constructed based on different reaction networks, stoichiometries, or even reaction rate expressions) are available, how to identify efficiently the model with the most adequate structure? (ii) if model parameters are not accurate enough, how to improve efficiently model parameter precision? (iii) how to validate the predictive capacity of the model within the operating space? The so-called post-evaluation module aims at solving such problems.

1.4.1 Model identification

If several model structures are proposed, the model with the best fit can be identified through two manners:

- Direct identification: Identify directly the model with the least criterion, which can be:

  - the parameter estimation criterion (see Section 1.2.3.1), which is used to characterize the fitness degree between model predictions and experimental responses;
  - the criterion, in which both model adequacy (usually quantified by parameter estimation criterion) and model structure complexity (equivalent for instance to the number of reactions in the candidate reaction network) are taken into consideration. For more details about such criteria, refer to the reviews given by Crampin et al. (2004);
- Stepwise identification: Design a Data Acquisition Experiment for Model Discrimination (DAEMD) and leave apart the infeasible models after experimental implementation followed by model parameter estimation until only one model remains or until the designed DAEMD has no more discrimination potential.

The optimal set of operating conditions for model discrimination can be identified within the operating space by maximizing the criteria constructed based on the divergence between different model outputs. The first design criterion used for discrimination of two rival models was proposed by Hunter and Reiner (1965) and represented as follows:

$$\varphi_{mn}(\boldsymbol{\xi}) = \sum_{s=1}^{n_{sp}} \Delta\hat{\mathbf{y}}_{mn}\left(\xi_s, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\theta}}_n\right)^T \cdot \Delta\hat{\mathbf{y}}_{mn}\left(\xi_s, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\theta}}_n\right) \quad (1.11)$$

where:

$$\Delta\hat{\mathbf{y}}_{mn}\left(\xi_s, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\theta}}_n\right) = \hat{\mathbf{y}}_m\left(\xi_s, \hat{\boldsymbol{\theta}}_m\right) - \hat{\mathbf{y}}_n\left(\xi_s, \hat{\boldsymbol{\theta}}_n\right) \quad (1.12)$$

represents the difference between the outputs predicted by the models m and n for the sth sample in the experiment $\boldsymbol{\xi}$. Note that this notation will be simplified to $\Delta\hat{\mathbf{y}}_{mn}(\xi_s)$ in the following.

It is important to point out that this criterion does not integrate the uncertainty on the measurements, nor on the model predictions. Therefore, (i) when the measurement errors of the designed experiment are actually large, discrimination may not be possible; (ii) when the uncertainty on the model predictions is large, the predicted difference in the model predictions may be less pronounced than expected and the value of the designed experiment with regard to model discrimination may ultimately be limited (Donckels et al., 2012).

In the view of the two potential situations mentioned above, it is necessary to improve this criterion by weighting the difference between model predictions according to Eq. (1.13).

$$\varphi_{mn}(\boldsymbol{\xi}) = \sum_{s=1}^{n_{sp}} \Delta\hat{\mathbf{y}}_{mn}\left(\xi_s, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\theta}}_n\right)^T \cdot \mathbf{W}^{-1} \cdot \Delta\hat{\mathbf{y}}_{mn}\left(\xi_s, \hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{\theta}}_n\right) \quad (1.13)$$

where W is the weighting matrix quantifying the uncertainty on either only the measurements (Hunter and Reiner, 1965), or both the measurements and the model predictions (Buzzi-Ferraris and Forzati, 1983; Schwaab et al., 2008; Donckels et al., 2009).

The weighting matrix quantifying the uncertainty on only the measurements is Σ_y, meanwhile, the matrix quantifying the uncertainty on both the measurements and the model predictions is defined as:

$$\mathrm{W_{mn}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}_\mathrm{m}, \widehat{\boldsymbol{\theta}}_\mathrm{n}\right) = 2 \cdot \Sigma_y(\xi_\mathrm{s}) + \boldsymbol{\Sigma}_{\hat{\boldsymbol{y}}_\mathrm{m}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}_\mathrm{m}\right) + \boldsymbol{\Sigma}_{\hat{\boldsymbol{y}}_\mathrm{n}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}_\mathrm{n}\right) \quad (1.14)$$

Here, $\boldsymbol{\Sigma}_{\hat{\boldsymbol{y}}}$ represents the $n_{resp} \times n_{resp}$–dimensional model prediction variance-covariance matrix, and can be calculated by propagating the uncertainty on the parameter estimates (see Section 1.3.2) according to Eq. (1.15).

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{y}}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}\right) = \mathbf{G}^\mathrm{T} \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} \mathbf{G} \quad (1.15)$$

where G is the Jacobian matrix representing the sensitivities of the model predictions $\hat{y}\left(\xi_\mathrm{s}, \hat{\theta}\right)$ to changes in the parameters written as follow:

$$\mathbf{G} = \begin{bmatrix} \dfrac{\partial \hat{y}_1\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}\right)}{\partial \hat{\theta}_1} & \cdots & \dfrac{\partial \hat{y}_{n_{resp}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}\right)}{\partial \hat{\theta}_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \hat{y}_1\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}\right)}{\partial \hat{\theta}_{n_u}} & \cdots & \dfrac{\partial \hat{y}_{n_{resp}}\left(\xi_\mathrm{s}, \widehat{\boldsymbol{\theta}}\right)}{\partial \hat{\theta}_{n_u}} \end{bmatrix} \quad (1.16)$$

Note that $\Sigma_{\hat{\theta}}$ can be estimated not only from the information of the experiments already performed according to Eq. (1.7), but also from the information of both the experiments already performed and the designed experiment to be performed according to Eq. (1.17).

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_{n_{exp}}, \boldsymbol{\xi}_{new}\right) = \left[\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_{n_{exp}}\right)^{-1} + \mathrm{M_F}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{new}\right)\right]^{-1} \quad (1.17)$$

Thus, the improved criteria are summarized in Table 1.2.

Table 1.2: Overview of the improved criteria for discrimination of two models in literature.

N°	Information taken into consideration for improvement			Reference
	Measurement error	Model prediction error	Information content of the designed experiment	
1	×			Hunter and Reiner (1965)
2	×	×		Buzzi-Ferraris and Forzati (1983)
3	×	×	×	Schwaab et al. (2008); Donckels et al. (2009)

All the criteria described above were designed for discrimination of two rival models. When more than two rival models are available, the global criterion, usually the highest pairwise criterion (Buzzi-Ferraris et al., 1990; Schwaab et al., 2006), is used and can be formalized as:

$$\varphi(\boldsymbol{\xi}) = \max_{m,n} \varphi_{mn}(\boldsymbol{\xi}) \quad (1.18)$$

1.4.2 Model refining

Model-based experimental design is an interesting tool to improve model parameter precision. Improving model parameter precision is equivalent in mathematical terms to decreasing the so-called size of the parameter variance-covariance matrix. Therefore, the criterion for experimental design can be defined as:

$$\varphi(\boldsymbol{\xi}_{new}) = h\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_{n_{exp}}, \boldsymbol{\xi}_{new}\right)\right) \quad (1.19)$$

where h is the so-called size (optimality) function of the parameter variance-covariance matrix calculated based on the best currently available parameter estimates from all the existing experiments and the designed new experiment.

The most common so-called size (optimality) functions are the determinant (D-optimality), the largest eigenvalue (E-optimality) and the trace (A-optimality), whose geometrical interpretations can be intuitively illustrated by Figure 1.3 for a two-parameter case.

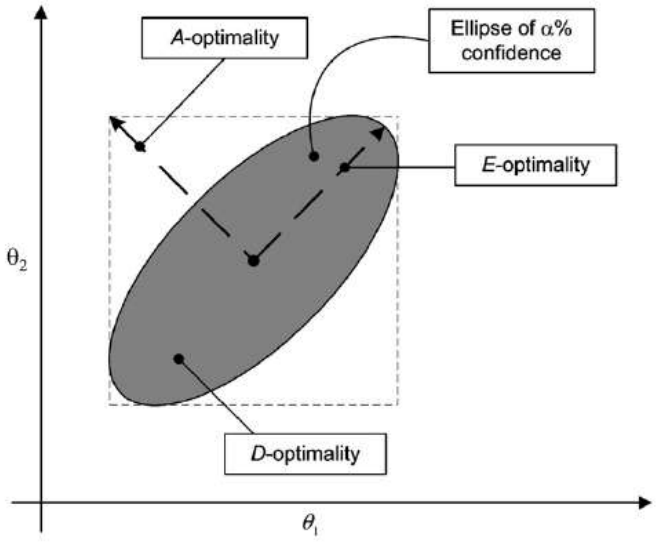


Figure 1.3: Geometrical interpretation of the standard criteria for the experimental design (Franceschini and Macchietto, 2008).

Each of these criteria has drawbacks, for instance: (i) the D-optimal criterion tends to give excessive importance to the more sensitive model parameters, which results in the fact that the less sensitive model parameters may also have the large confidence intervals (Pinto et al.,1990); (ii) the off-

diagonal elements of the information matrix are not included in the A-optimal criterion, which results in an appreciable loss of information, in particular in case of high cross-correlations between model parameters (De Pauw, 2005). A detail review about model-based experimental design used for parameter precision is given by Franceschini and Macchietto (2008).

The efficiency of model parameter precision improvement can be enhanced by using several types of reactors, which enlarges the experimental windows (Mathieu, 2013).

1.4.3 Model validation

The developed model, which has been successfully evaluated in terms of structure and parameters, can be used to describe only the experiments used for model parameter estimation. However, whether it possesses a satisfactory prediction capacity is uncertain and has to be verified. Therefore, validation experiment, usually for performance optimization, is designed and performed to check the model prediction quality.

As mentioned above, it is necessary to integrate a non-negative weight factor (or matrix) into the least square criterion due to the fact that the various measurable variables (i.e. the system outputs, Eq. (1.2)) may exhibit very different orders of magnitude (see Section 1.2.3.1). By analogy, the experimental design criterion for performance optimization, varying a lot from a project to another, should be defined as the sum of normalized performances to be optimized and represented by Eq. (1.20), since the various performances (i.e. the system outputs, Eq. (1.3)) may also exhibit very different orders of magnitude.

$$\varphi(\boldsymbol{\xi}_{new}) = \frac{1}{\sqrt{n_{perf}}} \sum_{p=1}^{n_{perf}} J_p^n(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{new}) \quad (1.20)$$

where n_{perf} is the number of performances to be optimized, J_p^n is the pth normalized performance defined as:

$$J_p^n(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) = \begin{cases} 1 & \text{if } \dfrac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) - o_p}{v_p - o_p} > 1 \\ \dfrac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) - o_p}{v_p - o_p} & \text{if } 0 \le \dfrac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) - o_p}{v_p - o_p} \le 1 \\ 0 & \text{if } \dfrac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) - o_p}{v_p - o_p} < 0 \end{cases} \quad (1.21)$$

where J_p is the function representing the pth performance, v_p and o_p are respectively the veto and objective values of the pth performance defined by the user.

1.5 Thesis context and objectives

Ideal model-development strategy should be both methodologically and practically systematic. A model-development strategy can be considered methodologically systematic only if it covers all three essential modules and if each module is systematic, specifically:

- Model structure development (see Section 1.2.1) and model parameter estimation (see Section 1.2.3) are two indispensable steps of model development module. Therefore, whether model development module is systematic depends on whether model structure analysis (see Section 1.2.2) is taken into consideration;
- In the systematic model evaluation module, both model (structure) adequacy and model (parameter) accuracy have to be evaluated;
- The so-called post-evaluation module is systematic only if it covers all the solutions for the three potential problems following model evaluation in terms of (adequacy of) structure(s) and (accuracy of) parameters.

Practical systematicness of a model-development strategy is intuitively reflected in two aspects:

1. Diversity of strategy application: a practically systematic model-development strategy should have the potential to :
   - be applied in any chemical reaction system, including not only the most widely studied homogeneous reaction systems but also the heterogeneous reaction systems;
   - construct simultaneously the model structures for different reactors, e.g. stirred-tank reactor, tubular reactor, fluidized bed reactor, fixed-bed reactor, etc.
2. Ease in strategy application: model-development strategy can be considered as only a methodological guideline, however, in practice, user expertise on reaction kinetics, modeling and statistical analysis are always required. Therefore, a practically systematic model-development strategy should also be easily applied by the user, concretely, it is preferred to be numerically implemented as a software (or codes), integrating all (or part) of calculations and allowing the user to concentrate to propose reasonable reaction networks and perform the experiments.

A survey on systematicness assessment of various model-development strategies reported in literature is given in Table 1.3. It can be observed that none of the reported model-development strategies is both methodologically and practically systematic.

In this context, it is necessary to: (i) develop a methodologically systematic model-development strategy, (ii) implement such a strategy as a software, which makes it practically systematic, (iii)

validate experimentally the feasibilities and generalities of the developed strategy and strategy-based software, corresponding to respectively the methodological, numerical, experimental objectives of the thesis.

Table 1.3: Systematicness assessment for various model-development strategies.

N°			1	2	3	4	5	6
Reference			Asprey and Macchietto (2000); Violet (2016)	Brendel et al. (2006)	Michalik et al. (2009)	Mathieu (2013)	Hii (2017)	Zhang et al. (2015)
Methodology	Model development	Model structure development [a]	U	T	U	U	O	S
		Model structure analysis	Yes	No	Yes	No	No	No
		Model parameter estimation [b]	D	S	S	D	D	D
	Model evaluation	Model structure evaluation	Yes	No	No	Yes	No	No
		Model parameter evaluation	Yes	No	No	Yes	No	No
	Post-evaluation	Model identification [c]	S	D	D	No	D	S
		Model refining	Yes	No	No	Yes	No	No
		Model validation	No	No	No	Yes	No	Yes
Practice	Application	Multiple reactors	No	No	No	Yes	No	No
		Application reaction systems [d]	O	O	B	O	O	O
Software implementation			No	No	No	No	Yes	No

[a] Model structure development, approximatively equivalent to reaction network development, can be achieved through four manners (see Section 1.2.1): U, reaction network is given by user; T, TFA technique is used to develop reaction network; S, reaction network is proposed within reaction supernetwork; O, reaction network is elucidated through evolutionary algorithms.

[b] Model parameters can be estimated through two methods (see Section 1.2.3.2): D, direct method; S, stepwise method.

[c] The model with the most adequate model structure can be identified through two methods (see Section 1.4.1): D, direct method; S, stepwise method.

[d] Application reaction systems can be classified into two categories: O, only homogeneous reaction systems; B, both homogeneous and heterogeneous reaction systems.

References

Achenie, L.K.E., Biegler, L.T., 1990. A superstructure based approach to chemical reactor network synthesis. Computers & Chemical Engineering, 14(1), 23-40.

Amrhein, M., Srinivasan, B., Bonvin, D., 1999. Target factor analysis of reaction data: use of data pre-treatment and reaction-invariant relationships. Chemical Engineering Science, 54(5), 579-591.

Asprey, S.P., Macchietto, S., 2000. Statistical tools for optimal dynamic model building. Computers & Chemical Engineering, 24(2-7), 1261-1267.

Bard, Y., 1974. Nonlinear Parameter Estimation. Academic Press, New York.

Bonvin, D., Rippin, D.W.T., 1990. Target factor analysis for the identification of stoichiometric models. Chemical Engineering Science, 45(12), 3417-3426.

Brendel, M., Bonvin, D., Marquardt, W., 2006. Incremental identification of kinetic models for homogeneous reaction systems. Chemical Engineering Science, 61(16), 5404-5420.

Burnham, S.C., Searson, D.P., Willis, M.J., Wright, A.R., 2008. Inference of chemical reaction networks. Chemical Engineering Science, 63(4), 862-873.

Buzzi-Ferraris, G., Forzati, P., 1983. A new sequential experimental design procedure for discriminating among rival models. Chemical Engineering Science, 38(2), 225–232.

Buzzi-Ferraris, G., Forzatti, P., Canu, P., 1990. An improved version of a sequential design criterion for discriminating among rival multiresponse models. Chemical Engineering Science, 45(2), 477-481.

Crampin, E.J., Schnell, S., McSharry, P.E., 2004. Mathematical and computational techniques to deduce complex biochemical reaction mechanisms. Progress in Biophysics and Molecular Biology, 86(1), 77-112.

De Pauw, D., 2005. Thesis, Optimal experimental design for calibration of bioprocess models: a validated software toolbox. Ghent University, Ghent, Belgium.

Donckels, B.M.R., De Pauw, D.J.W., De Baets, B., Maertens, J., Vanrolleghem, P.A., 2009. An anticipatory approach to optimal experimental design for model discrimination. Chemometrics and Intelligent Laboratory Systems, 95(1), 53–63.

Donckels, B.M.R., De Pauw, D.J.W., Vanrolleghem, P.A., De Baets, B., 2012. Performance assessment of the anticipatory approach to optimal experimental design for model discrimination. Chemometrics and Intelligent Laboratory Systems, 110(1), 20-31.

Fotopoulos, J., Georgakis, C., Stenger, H.G., 1994. Structured target factor analysis for the stoichiometric modeling of batch reactors. American Control Conference, 495-499.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science, 63(19), 4846-4872.

Grewal, M., Glover, K., 1976. Identifiability of linear and nonlinear dynamical systems. IEEE Transactions on automatic control, 21(6), 833-837.

Hii, C.J.K., 2017. Thesis, Elucidation of chemical reaction networks through genetic algorithm. Newcastle University, Newcastle, United Kingdom.

Hunter, W.G., Reiner, A.M., 1965. Designs for discriminating between two rival models. Technometrics, 7(3), 307–323.

Koza, J.R., Mydlowec, W., Lanza, G., Yu, J., Keane, M.A., 2007. Automatic computational discovery of chemical reaction networks using genetic programming. In Computational Discovery of Scientific Knowledge, 205-227. Springer, Berlin, Heidelberg.

Mathieu, F., 2013. Thesis, Stratégie d'intensification des procédés. Université de Lorraine, Nancy, France.

Miao, H., Xia, X., Perelson, A.S., Wu, H., 2011. On identifiability of nonlinear ODE models and applications in viral dynamics. SIAM review, 53(1), 3-39.

Michalik, C., Brendel, M., Marquardt, W., 2009. Incremental identification of fluid multi-phase reaction systems. AIChE journal, 55(4), 1009-1022.

Ollivier, F., 1990. Thesis, Le problème de l'identifiabilité structurelle globale: approche théorique, méthodes effectives et bornes de complexité. Ecole Polytechnique, Palaiseau, France.

Pinto, J.C., Lobao, M.W., Monteiro, J.L., 1990. Sequential experimental design for parameter estimation: a different approach. Chemical Engineering Science, 45(4), 883-892.

Searson, D.P., Willis, M.J., Wright, A.R., 2012. Reverse Engineering Chemical Reaction Networks from Time Series Data. In M. Dehmer, K. Varmuza, & D. Bonchev, Statistical Modelling of Molecular Descriptors in QSAR/QSPR (pp. 327-348). John Wiley & Sons Incorporated.

Schwaab, M., Silva, F.M., Queipo, C.A., Barreto Jr, A.G., Nele, M., Pinto, J.C., 2006. A new approach for sequential experimental design for model discrimination. Chemical Engineering Science, 61(17), 5791-5806.

Schwaab, M., Monteiro, J.L., Pinto, J.C., 2008. Sequential experimental design for model discrimination: Taking into account the posterior covariance matrix of differences between model predictions. Chemical Engineering Science, 63(9), 2408-2419.

Pohjanpalo, H., 1978. System identifiability based on the power series expansion of the solution. Mathematical Biosciences, 41(1-2), 21-33.

Vajda, S., Rabitz, H., 1989. State isomorphism approach to global identifiability of nonlinear systems. IEEE Transactions on Automatic Control, 34(2), 220-223.

Violet, L., 2016. Thesis, Stratégie expérimentales optimales pour la discrimination de modèles stoechio-cinétiques. Université de Toulouse, Toulouse, France.

Walter, E., Pronzato, L., 1994. Identification de modèles paramétriques à partir de données expérimentales. Masson.

Zhang, W., Binns, M., Theodoropoulos, C., Kim, J.K., Smith, R., 2015. Model Building Methodology for Complex Reaction Systems. Industrial & Engineering Chemistry Research, 54(16), 4603-4615.

Chapter 2. Systematic strategy for model development

In the previous chapter, a literature survey on the existing model-development strategies has been presented. It can be seen that each strategy has application limitations. None of them can be applied to address the common problem that scientists and engineers have to face, namely, given a homogeneous or heterogeneous catalytic reaction system and several potential reactors, how to develop reasonable reactor models and to identify the optimal reactor with its associated optimal operating conditions? In this context, a methodologically systematic model-development strategy based on model-based experimental design for simultaneous model development and reactor performance optimization has been developed.

In this chapter, firstly, the descriptive presentation of the methodologically systematic five-module model-development strategy is given. Subsequently, the systematic two-step procedure for model structure development and the systematic three-step sequential procedure for model parameter development are described. Then, model evaluations and model-based experimental designs scattered in various modules for different purposes, are presented together. Finally, the optimization algorithm used for model parameter estimation and model-based experimental design is presented.

Considering that this strategy is developed by assimilating the model-development strategies reported in literature, the descriptions of the common points involved in this strategy and those reported in literature, such as model structure analysis, experimental design criteria, etc., are not given in this chapter, but the references to the corresponding sections in Chapter1 are introduced in the text.

2.1 Strategy description

Figure 2.1 presents a complete overview of the proposed methodologically systematic model-development strategy. Before presenting the details of each step, one should note that the complex steps involved in the complete strategy can be grouped into five modules with different objectives (shown in Figure 2.2):

- Model development module (see Section 2.1.1): One or several sets of Models Under Test (MUTs) are developed to describe the reaction system under investigation in different reactors. The number of Sets of MUTs (SMUTs) depends on the number of candidate reaction system descriptions (see Section 2.2.1); the number of MUTs contained in one SMUT depends on the number of reactors taken into consideration.

- Initial data acquisition module (see Section 2.1.1): Initial data acquisition experiments, namely, Preliminary Experiments (PEs), are designed and performed. Experimental data used for model parameter estimation, including the operating conditions and the corresponding measured responses, is initialized and will be updated after the experiments carried out in the subsequent modules.
- Model identification module (see Section 2.1.2): the objective of this module is to identify the most suitable SMUT in terms of adequacy of reaction system description used for model structure development.
- Model validation module (see Section 2.1.3): the objective is to check experimentally the model prediction quality.
- Model refining module (see Section 2.1.4): the objective is to improve the model prediction precision by increasing model parameter precision.

Inter-module and intra-module workflows will be described in this section. Note that all the steps mentioned below refer to Figure 2.1.

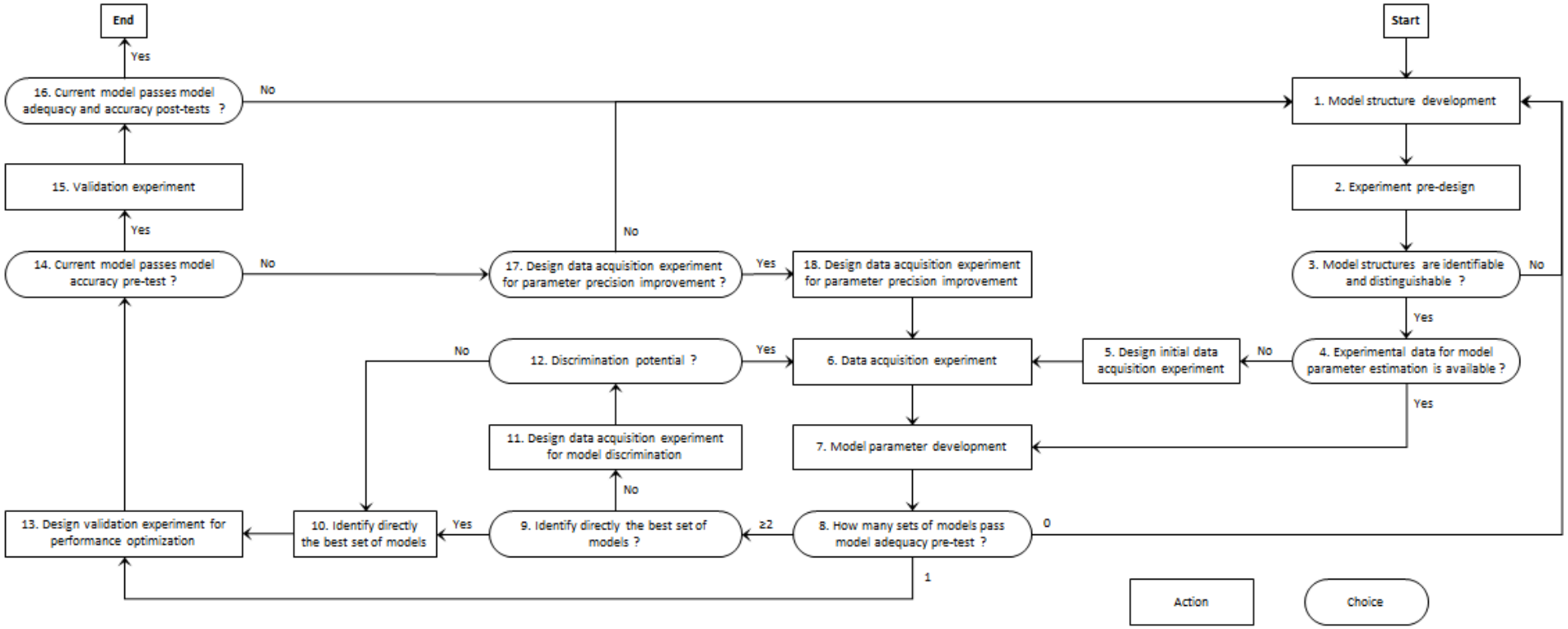


Figure 2.1: Systematic model-development strategy based on model-based experimental design.

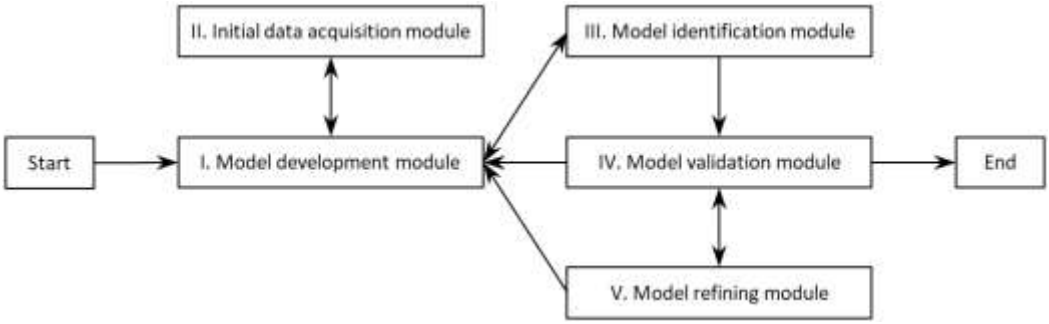


Figure 2.2: General procedure for model development.

2.1.1 Model development module and initial data acquisition module

Model development module and initial data acquisition module (illustrated in Figure 2.3) alternate and are described together.

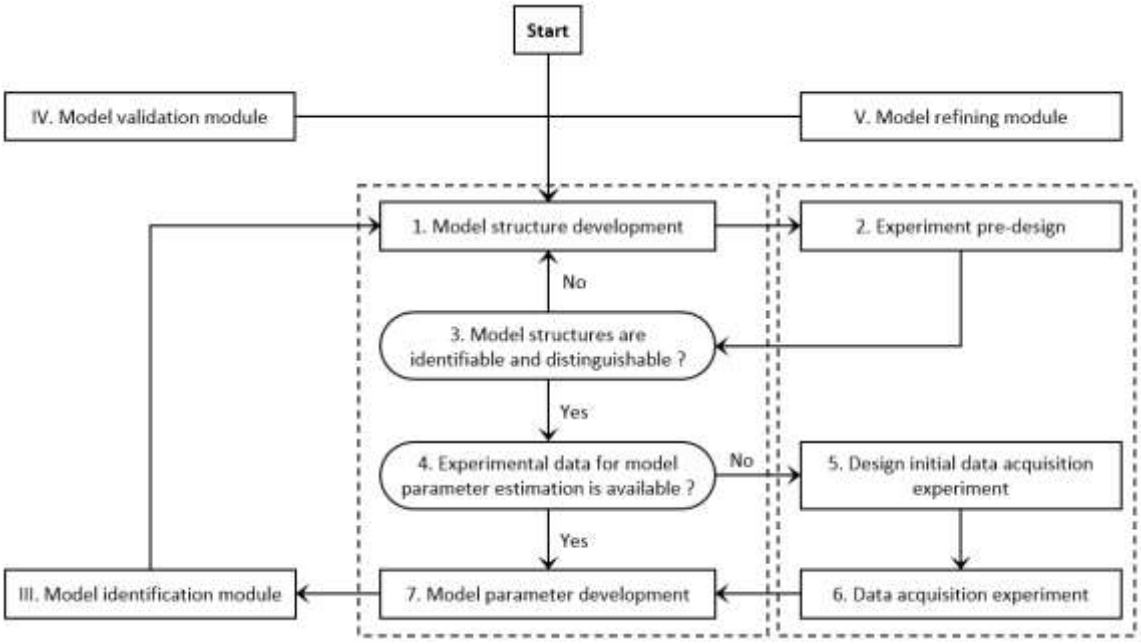


Figure 2.3: Model development module (in the left dotted box) and initial data acquisition module (in the right dotted box).

The first steps consist in developing one or several candidate sets of model structures (step 1) (see Section 2.2), defining the operating space and a set of variables to be actually measured (step 2). Given these information, model structure analysis is performed for each set of model structures to evaluate whether its parameters are identifiable and whether it can be distinguished from the others (step 3) (see Section 1.2.2). The three first steps are repeated until candidate sets of model structures are proven identifiable and distinguishable. Then, one should verify whether one of the following situations happens (step 4):

1. initial model structure development;

2. following model structure re-development due to the failure of model adequacy evaluation carried out at model identification, validation and refining modules, the parameters involved in the new model structures are not identifiable using current experimental data.

If one of these situations happens, PEs have to be designed and performed (steps 5-6). Finally, one or several SMUTs are developed after model parameter development (step 7) (see Section 2.3).

2.1.2 Model identification module

Following the model development module, the model identification module (shown in Figure 2.4) is performed.

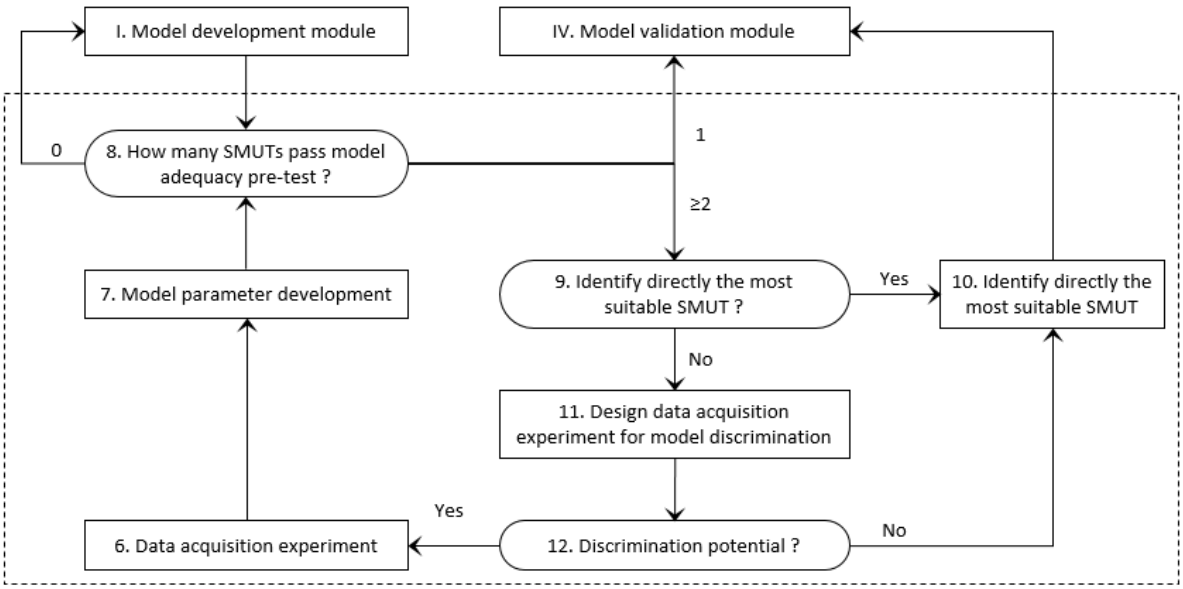


Figure 2.4: Model identification module (in dotted box).

Firstly, model adequacy pre-test is carried out for each SMUT to investigate whether it can be used to explain the observed data of currently available experiments (step 8) (see Section 2.4). According to the number of SMUTs validating model adequacy pre-test, different steps have to be performed: (i) if no SMUT validates model adequacy pre-test, the user should come back to step 1 (model development module); (ii) if only one SMUT validates model adequacy pre-test, the user should continue to step 13 (model validation module); (iii) if at least two SMUTs validate model adequacy pre-test, the most suitable SMUT can be identified through two methods:

- direct identification method (steps 9-10): the SMUT with the minimal parameter estimation criterion is directly identify (see Section 2.3.3);
- stepwise identification method (iterative steps 9-11-12-6-…-6-…-8/12-10): a series of Data Acquisition Experiments for Model Discrimination (DAEMD), which can be carried out in one run, and in which the responses predicted by the SMUTs are the most different between each

other, is designed (see Section 2.5). After the DAEMD, at least one SMUT can be removed, and the iterative steps continue until only one SMUT passes model adequacy pre-test (step 8) or until the DAEMD has no more discrimination potential (step 12). For the later situation, direct identification method is used.

The default method is the direct one, but if the SMUTs possess the equal or similar parameter estimation criteria, the stepwise method seems better to be used.

2.1.3 Model validation module

Model validation module (shown in Figure 2.5) has two possible inputs:

1. output of model identification module, i.e. the SMUT passing model adequacy pre-test;
2. output of model refining module (see Section 2.1.4), i.e. the SMUT passing both model adequacy and accuracy pre-tests.

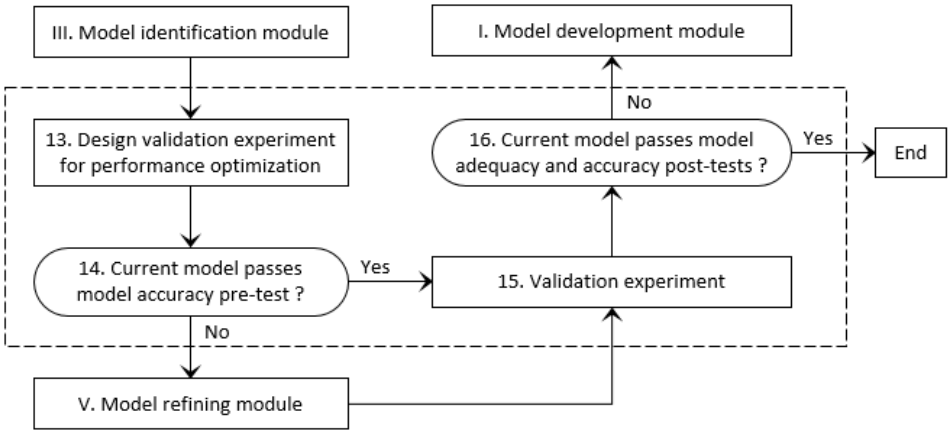


Figure 2.5: Model validation module (in dotted box).

After the Validation Experiment for Performance Optimization (VEPO) is designed (step 13) (see Section 2.5), model accuracy pre-test is carried out to evaluate whether the prediction errors of the MUT in the current SMUT corresponding to the optimal reactor (hereinafter referred to as the current model) for the VEPO to be soon performed are greater than the measurement errors (step 14) (see Section 2.4). If the current model passes model accuracy pre-test, after the VEPO (step 15), model adequacy and accuracy post-tests are carried out (step 16) (see Section 2.4). Otherwise, model refining module has to be performed.

If the current model passes both model adequacy and accuracy post-tests, the current SMUT (structures and parameters) is validated and the optimal reactor with its associated optimal operating conditions for the given synthesis can be properly identified. Otherwise, model development module has to be performed.

Various modules are repeated until a set of adequate and accurate models, that fit all experiments, is identified.

2.1.4 Model refining module

After a failure of model accuracy pre-test carried out in model validation module, model refining module (iterative steps 17-18-6-7-8-13-14-…-6-…-17/8/14) (shown in Figure 2.6) is performed.

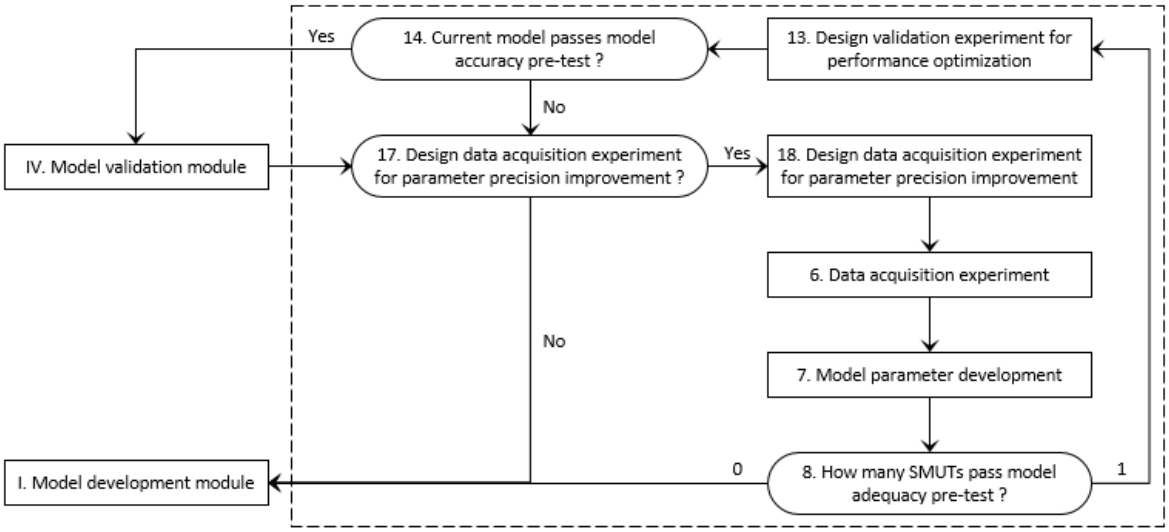


Figure 2.6: Model refining module (in dotted box).

At the beginning of each iteration, a manual intervention is required (step 17), namely, the user has to evaluate on the basis of his expertise whether the design of a series of Data Acquisition Experiments for Parameter Precision Improvement (DAEPPI) (step 18) (see Section 2.5) as well as its subsequent steps is necessary, since there is currently no quantified criterion. For instance, after several iterations, if model parameter precision increases no more or insignificantly, the design of DAEPPI is not necessary.

2.2 Model structure development

The general model structure for any reaction system can be represented by the following set of differential and algebraic equations:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \boldsymbol{\xi}, \boldsymbol{\theta}, t) \quad (2.1)$$

$$\mathbf{x}\left(t^{int}\right) = \mathbf{x}^{int} \quad (2.2)$$

$$\hat{\mathbf{y}}(t) = \mathbf{x}(t)M_T \quad (2.3)$$

where:

- $\mathbf{x}(t)$ is an n_{sv}-dimensional vector of time-dependent state variables (sv), such as: the species concentrations, the phase volumes and the system temperature,
- $\boldsymbol{\xi}$ is an $\sum_{e=1}^{n_{exp}} n_{sp_e} \times n_{oc}$-dimensional array of operating conditions (oc), such as: the sampling times, the initial conditions, the flow rate, the duration and the composition of each feed, n_{exp} and n_{sp_e} represent respectively the number of experimental runs and the number of samples in the eth experiment,
- $\boldsymbol{\theta}$ is an n_u-dimensional vector of model parameters,
- t is the time, between 0 and τ, the experiment duration,
- $\mathbf{f}$ is an n_{sv}-dimensional set of non-linear ordinary differential equations describing $\dot{\mathbf{x}}(t)$, the first derivatives of $\mathbf{x}(t)$ with respect to time,
- $\mathbf{x}^{int}$ is the initial state for $t=t^{int}=0$,
- $\hat{\mathbf{y}}(t)$ is an n_{resp}-dimensional vector of measurable variables, that can be transformed from $\mathbf{x}(t)$ by M_T, the $n_{sv} \times n_{resp}$-dimensional transformation matrix.

In order to develop the specific model structure for a given reaction system in a given reactor, one should propose a reasonable reaction system description and indicate the reactor type.

2.2.1 Reaction system description

Reaction system description is composed of two elements, namely, the reaction network and the rate law for each item involved in the reaction network.

2.2.1.1 Reaction supernetwork and reaction network

In this work, reaction network development is a sequential procedure composed of two steps: (i) development of a reaction supernetwork; (ii) extraction of the reaction network(s) from the reaction supernetwork developed before.

A conventional reaction supernetwork (see Section 1.2.1.2) contains only chemical reactions and is restricted to homogeneous reaction systems. In this work, in order to expand the application range of reaction supernetwork to heterogeneous reaction systems, it is designed to contain also mass transfers, adsorptions and desorptions, and can be developed through two methods:

- theoretical method: the optimizer automatically generates all possible chemical reactions, mass transfers, adsorptions and desorptions, and then, eliminates the ones meeting the predefined elimination criteria. For instance, the chemical reactions which do not satisfy the mass balance

(Burnham et al., 2008) or the atomic balance (Zhang et al., 2015), the mass transfers in which the involved species are (assumed to be) not soluble in the debit or credit phase, should be eliminated;

- empirical method: the user elucidates all possible chemical reactions, mass transfers, adsorptions and desorptions on the basis of the literature and his expertise.

In this reaction supernetwork, mass transfers, adsorptions and desorptions are described in the same manner as chemical reactions. For instance, mass transfer of species i between phases k and l is represented by Eq. (2.4), in which the reactant is the species in the debit phase, the product is the species in the credit phase; adsorption of species i in the phase k on the vacant active site (*) of the catalyst is represented by Eq. (2.5), in which the reactants are the species in the phase k and the vacant active site of the catalyst, the product is the adsorbed species.

$$i_k \leftrightarrow i_l \quad (2.4)$$

$$i_k + * \leftrightarrow i^* \quad (2.5)$$

Furthermore, each item involved in the reaction supernetwork is assumed to be irreversible. Therefore, a reversible chemical reaction has to be represented by two chemical reactions, a forward one and a backward one. Meanwhile, mass transfer between two phases is assumed to proceed always from one phase to another phase, which is mechanistically unreasonable. The sign of the mass-transfer rate, which can be positive or negative, is used to indicate the mass-transfer direction. Mass transfer proceeds in the opposite direction, if the calculated mass-transfer rate is negative. By analogy, adsorption/desorption is assumed to proceed always in the adsorption direction, the desorption direction is taken, when the apparent adsorption rate is negative.

The reaction network, a subset of items extracted from the reaction supernetwork, can be also proposed by the user or the optimizer. To illustrate this procedure, refer to the selected validation reaction systems, namely, NaOH-catalyzed ethanolysis of sunflower oil (see Sections 4.2.1.5 and 4.2.1.6) and epoxidation of sunflower oil by performic acid generated in situ (see Sections 5.2.1.5 and 5.2.1.6).

2.2.1.2 Rate law expression

Once the reaction network is available, a rate law for each involved item has to be defined. Rate law expressions have two forms: default and user-defined. For what concerns default reaction rate expressions, those for different chemical reaction types are given below:

- In a homogeneous catalytic or non-catalytic reaction system, the reaction rate of the chemical reaction j in the phase k is expressed according to a power law:

$$r_{j,k} = k_{j,k} \prod_{i}^{n_{s_k}} c_{i,k}^{\gamma_{i,j,k}} \quad (2.6)$$

Where $k_{j,k}$ is the reaction rate constant of the chemical reaction j in the phase k, n_{s_k} is the number of species in the phase k, $c_{i,k}$ is the concentration of the species i in the phase k, $\gamma_{i,j,k}$ is the kinetic order of the species i for the chemical reaction j in the phase k.

- In a heterogeneous catalytic reaction system,
  - Reaction rate of decomposition of adsorbed reactant A is expressed:

$$r = k\theta_A c_S \quad (2.7)$$

    where c_S is the concentration of all active sites (occupied or not) of the catalyst, and θ_A is the fraction of the sites occupied by reactant A.

  - Reaction rate between two adsorbed reactants A and B according to Langmuir–Hinshelwood mechanism:

$$r = k\theta_A \theta_B c_S^{\,2} \quad (2.8)$$

  - Reaction rate between adsorbed reactant A and reactant B in the phase k according to Eley–Rideal mechanism:

$$r = k\theta_A c_S c_{B,k} \quad (2.9)$$

Mass-transfer rate of Eq. (2.4) is by default represented by Eq. (2.10) on the basis of the two-film theory (Whitman, 1923). For additional mass-transfer rate expressions, refer to the book (Taylor and Krishna, 1993).

$$J_{i,k\to l} = K_{i,k}\left(c_{i,k} - m_{i,k|l} c_{i,l}\right) \quad (2.10)$$

where $K_{i,k}$ is the overall mass-transfer coefficient of the species i, based on the debit phase k, $m_{i,k|l}$ is the partition coefficient of the species i between the phases k and l.

The default apparent adsorption rate of Eq. (2.5) is written as:

$$L_i = k_i^a c_{i,k}\left(1 - \sum_{i}^{n_{s_k}} \theta_i\right) c_S - k_i^d \theta_i c_S \quad (2.11)$$

Where k_i^a and k_i^d are respectively the rate constants for respectively surface adsorption and desorption of species i.

2.2.2 Reactor type

Once the reaction system description is available, non-linear ordinary differential equations describing the variation of each time-dependent state variable taken into consideration has to be specified, which depends on the reactor type. For instance, given a homogeneous catalytic reaction system and a Stirred-Tank Reactor (STR), which are concerned in this work, the ordinary differential equations describing the variations of the concentration of the species i in the phase k, the volume of the phase k and the system temperature are represented by Eqs. (2.12)-(2.14). The deduction process of these equations is given in Appendix 1.

$$\frac{dc_{i,k}}{dt} = \frac{1}{V_k}\left(F_{i,k}^{in} - F_{i,k}^{out} - c_{i,k}\frac{dV_k}{dt}\right) + \sum_{j}^{n_{r_k}} \nu_{i,j,k} r_{j,k} - \sum_{l}^{n_p - 1} J_{i,k \to l \neq k} \quad (2.12)$$

$$\frac{dV_k}{dt} = \sum_{i}^{n_{s_k}} v_i \left(F_{i,k}^{in} - F_{i,k}^{out} + V_k \sum_{j}^{n_{r_k}} \nu_{i,j,k} r_{j,k} - V_k \sum_{l}^{n_p - 1} J_{i,k \to l \neq k} \right) \quad (2.13)$$

$$\frac{dT}{dt} = \frac{\sum_{k}^{n_p} \sum_{i}^{n_{s_k}} F_{i,k}^{in} M_i Cp_i \left(T^{in} - T\right) + US\left(T^{exch} - T\right) - \sum_{k}^{n_p} \sum_{j}^{n_{r_k}} V_k \Delta h_{j,k} r_{j,k}}{\sum_{k}^{n_p} \sum_{i}^{n_{s_k}} V_k c_{i,k} M_i Cp_i + m_r Cp_r} \quad (2.14)$$

where:

- $F_{i,k}^{in}$ and $F_{i,k}^{out}$ are respectively the molar inflow and the outflow rates of the species i in the phase k, with the following constraints:
  - for a Batch Stirred-Tank Reactor (BSTR), there is neither reactant addition ($\sum_{k}^{n_p} \sum_{i}^{n_{s_k}} v_i F_{i,k}^{in}$=0, $F_{i,k}^{in}$=0) nor product removal ($\sum_{k}^{n_p} \sum_{i}^{n_{s_k}} v_i F_{i,k}^{out}$=0, $F_{i,k}^{out}$=0);
  - for a Semi-Batch Stirred-Tank Reactor (SBSTR), there is either reactant addition or product removal;
  - for a Continuous Stirred-Tank Reactor (CSTR), there are both reactant addition and product removal, and their volumetric flow rates are identical,
- n_{r_k} is the number of chemical reactions occurring in the phase k,
- $\nu_{i,j,k}$ is the stoichiometric coefficient of the species i for the chemical reaction j in the phase k,
- n_p is the number of phases,
- v_i is the molar volume of the species i,
- M_i is the molar mass of the species i,
- Cp_i is the heat capacity of the species i,
- T^{in} is the feed temperature,

- U is the global heat-transfer coefficient,
- S is the heat-transfer area,
- T^{exch} is the temperature of the fluid flowing in the jacket,
- $\Delta h_{j,k}$ is the molar enthalpy of the reaction j in the phase k,
- m_r is the empty reactor mass,
- Cp_r is the heat capacity of the reactor manufacturing material.

2.3 Model parameter development

Model parameter development is a sequential procedure composed of three steps: (i) model parameter reformulation, (ii) model parameter pre-assignment, (iii) model parameter estimation. Each step will be described below.

2.3.1 Model parameter reformulation

Before parameter estimation, one should first indicate the type of each involved parameter from independent parameter, temperature-dependent parameter and multi-dependent parameter, and then, reformulate the dependent parameters as a function of the independent parameters.

For example, for the model structure constructed above and used for describing the homogeneous catalytic reaction system in the STR-type reactor, the involved parameters are:

- reaction rate constants, kinetic orders (and molar reaction enthalpies, if necessary) for the chemical reactions,
- overall mass-transfer coefficients and distribution coefficients for the mass transfers.

Among them, the chemical reaction rate constants can be reformulated by Eq. (2.15) according to the Arrhenius law when it is considered as a temperature-dependent parameter.

$$k_{j,k} = A_{j,k} e^{-\frac{Ea_{j,k}}{RT}} \quad (2.15)$$

Where $A_{j,k}$ and $Ea_{j,k}$ are the independent kinetic parameters (pre-exponential factor and energy of activation respectively). Considering $A_{j,k}$ and $Ea_{j,k}$ may exhibit very different orders of magnitude, they can be further reformulated by Eqs. (2.16)-(2.17):

$$A_{j,k} = 10^{\frac{T_{max} \log k_{j,k,T_{max}} - T_{min} \log k_{j,k,T_{min}}}{T_{max} - T_{min}}} \quad (2.16)$$

$$Ea_{j,k} = \frac{R T_{max} T_{min} \left(\log k_{j,k,T_{max}} - \log k_{j,k,T_{min}}\right) \ln 10}{T_{max} - T_{min}} \quad (2.17)$$

where $\log k_{j,k,T_{max}}$ and $\log k_{j,k,T_{min}}$ are the reaction rate constant logarithms at respectively the maximal and minimal possible reaction temperatures, namely, the independent parameters to be actually estimated instead of the temperature-dependent parameter $k_{j,k}$.

Molar reaction enthalpy can be reformulated by Eq. (2.18) as a function of the enthalpy at the predefined reference temperature, usually the room temperature.

$$\Delta h_{j,k} = \Delta h_{j,k,ref} + (T - T_{ref}) \sum_{i} \nu_{i,j,k} M_i Cp_i \quad (2.18)$$

2.3.2 Model parameter pre-assignment

One effective way to decrease the computation time of model parameter estimation is to reduce the number of model parameters to be estimated by pre-assigning reasonable values to some model parameters.

For instance, for the same model structure illustrated for model parameter reformulation, among the involved parameters, the kinetic orders for the Reactant (R), the Solvent (S), the Catalyst (C) and the Product (P) can be by default taken equal to its stoichiometric coefficient, 0, 1 and 0 respectively ($\gamma \sim k c_R^{\nu_R} c_S^{0} c_C^{1} c_P^{0}$); the molar reaction enthalpies and the partition coefficients can be reasonably defined according to the literature, therefore, the parameters to be actually estimated are only the reaction rate constants and the overall mass-transfer coefficients.

2.3.3 Model parameter estimation

After model parameter reformulation and pre-assignment, the values of the model parameters are estimated by Eq. (2.19), in which the criterion is built based on Maximum Likelihood Estimator. For more model parameter estimation criteria, refer to Section 1.2.3.1.

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{\sum_{e=1}^{n_{exp}} n_{sp_e}} \sum_{e=1}^{n_{exp}} \frac{n_{sp_e}}{2} \sum_{r=1}^{n_{resp}} \ln\left\{ \frac{2\pi}{n_{sp_e}} \sum_{f=1}^{n_{sp_e}} \left[y_r(\boldsymbol{\xi}_{e,f}) - \hat{y}_r(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta}) \right]^2 \right\} \quad (2.19)$$

where $y_r(\boldsymbol{\xi}_{e,f})$ is the rth measured response obtained under the operating conditions $\boldsymbol{\xi}_{e,f}$, which is the fth set of the operating conditions of the eth experiment, $\hat{y}_r(\boldsymbol{\xi}_{e,f}, \boldsymbol{\theta})$ is the corresponding model prediction.

2.4 Model evaluation

At the steps 8, 14 and 16 of the strategy, model is evaluated in terms of structure adequacy and prediction accuracy.

Model structure adequacy can be, not only qualitatively evaluated by using the minimized value of parameter estimation criterion once the parameters are identified or by comparing graphically measured and predicted responses, but also quantitatively evaluated using χ^2 statistic distribution (see Section 1.3.1). Specifically, according to whether model parameters are updated after experiments, model adequacy tests can be classified into two categories:

- model adequacy pre-test implemented at the step 8 (i.e. after data acquisition experiments followed by model parameter estimation) to evaluate for each SMUT whether it allows fitting all available experiments;
- model adequacy post-test implemented at the step 16 to verify whether the current model describes adequately the VEPO performed at the step 15.

In order to evaluate model prediction accuracy, which is practically equivalent to verify for each measurable variable whether its model prediction error is less than its corresponding measurement error, model prediction variance-covariance matrix, as its name suggests, quantifying the uncertainty on the model predictions, is used and can be calculated by propagating the uncertainty on the parameter estimates according to Eq. (1.15). Specifically, model accuracy tests can be also classified into two categories:

- model accuracy pre-test implemented at the step 14 to *evaluate* whether the current model *has the potential to predict* accurately the VEPO to be soon performed at step 15;
- model accuracy post-test implemented at the step 16 to *verify* whether the current model *has predicted* accurately the VEPO performed at the step 15.

2.5 Model-based experimental design

At the steps 11, 13 and 18 of the strategy, in order to identify the optimal reactor with its associated optimal operating conditions for respectively the Data Acquisition Experiments for Model Discrimination (DAEMD), the Validation Experiment for Performance Optimization (VEPO) and the Data Acquisition Experiments for Parameter Precision Improvement (DAEPPI), model-based experimental design is performed, specifically:

1. the optimal operating conditions of each considered reactor have to be identified by minimizing or maximizing the specific experimental design criterion according to Eq. (2.20).

$$\boldsymbol{\xi}_{R_q}^{opt} = \arg \begin{cases} \max\limits_{\boldsymbol{\xi}_{R_q} \in \Xi_{R_q}} \varphi_{DAEMD}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}\right) \\ \min\limits_{\boldsymbol{\xi}_{R_q} \in \Xi_{R_q}} \varphi_{VEPO}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}\right) \\ \min\limits_{\boldsymbol{\xi}_{R_q} \in \Xi_{R_q}} \varphi_{DAEPPI}\left(\widehat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_e, \cdots, \boldsymbol{\xi}_{n_{exp}}, \boldsymbol{\xi}_{R_q}\right) \end{cases} \quad (2.20)$$

where:

- Ξ_{R_q} represents the research field of operating conditions for the qth reactor;
- φ_{DAEMD}, criterion for DAEMD, is defined as the highest pairwise criterion and represented by Eq. (1.18), in which the pairwise criterion proposed by Buzzi-Ferraris and Forzati (1983) is used by default. For more details about other pairwise criteria, refer to the Section 1.4.1;
- φ_{VEPO}, criterion for VEPO, is defined as the sum of normalized performances and represented by Eq. (1.20), in which the veto and objective values of each performance have to be specified by the user according to the project. For more information, refer to the Section 1.4.3;
- φ_{DAEPPI}, criterion for DAEPPI, is defined as the so-called size of the parameter variance-covariance matrix and represented by Eq. (1.19), in which the so-called size function is defined as the geometric mean of the diagonal elements. Therefore, the mathematical meaning of the criterion is the average variance of model parameters. For more details about other so-called size (optimality) functions, refer to the Section 1.4.2;

2. the optimal reactor with the minimal or maximal criterion can be easily identified.

2.6 Optimization algorithm

Both model parameter estimation and model-based experimental design for the DAEMD, the VEPO and the DAEPPI are optimization problems. To find the optimal solution, the use of an optimization algorithm is required. In this work, a hybrid optimization algorithm combining global optimization and local search (Balland et al., 2000; Mouhab et al., 2008), is used: concrete optimization algorithms as well as the corresponding Matlab functions are shown in Table 2.1.

Table 2.1: Concrete optimization algorithms and the corresponding Matlab functions.

Algorithm type	Algorithm	Matlab function [a]
Global optimization	Genetic algorithm	ga
	Particle swarm optimization	pso [b]
Local search	Interior-point algorithm [c]	fmincon

[a] The systematic model-development strategy is implemented as software using MATLAB R2014a. Software description is given in Section 3.1.

b Matlab implementation of particle swarm optimization is achieved by Chen (2018). pso has the same syntax as ga.

c fmincon proposes four optimization algorithms: 'interior-point' (default), 'trust-region-reflective', 'sqp' and 'active-set'. The default algorithm is used.

2.7 Conclusion

A methodologically systematic model-development strategy, consisting of initial data acquisition, model development, model identification, model validation and model refining modules, has been developed. Its systematicness is illustrated by the following features:

1. a systematic three-step model development procedure, including model structure development, model structure analysis and model parameter development, is applied;
2. model evaluation is performed in terms of both (adequacy of) structures and (accuracy of) parameters;
3. model-based experimental design is used for different purposes, i.e. model identification, validation and refining;
4. the developed strategy can be applied not only to the homogeneous reaction systems but also to the heterogeneous ones, thanks to the use of the reaction supernetwork containing not only chemical reactions but also mass transfers, adsorptions and desorptions;
5. multiple reactor models can be simultaneously developed.

Before validating experimentally the feasibility and generality of the developed strategy, necessary preparation works are required. In the next chapter, the following preparations will be presented.

1. Development of a strategy-based computer-aided software used for facilitating user's work;
2. Screening of the validation reaction systems used to demonstrate the feasibility and generality of the strategy (as well as the strategy-based software).

References

Balland, L., Estel, L., Cosmao, J.M, Mouhab, N., 2000. A genetic algorithm with decimal coding for the estimation of kinetic and energetic parameters. Chemometrics and Intelligent Laboratory Systems, 50(1), 121-135.

Burnham, S.C., Searson, D.P., Willis, M.J., Wright, A.R., 2008. Inference of chemical reaction networks. Chemical Engineering Science, 63(4), 862-873.

Chen, S., 2018. Constrained Particle Swarm Optimization (2009-2018). MATLAB File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/25986.

Mouhab, N., Balland, L., Talouba, I.B., Cosmao, J.M., 2008. Study of a chemical reaction in heterogeneous liquid–liquid medium producing a surfactant and a cosolvent. Chemical Engineering and Processing: Process Intensification, 47(3), 363-369.

Taylor, R., Krishna, R., 1993. Multicomponent mass transfer. Vol. 2. John Wiley & Sons.

Whitman, W. G., 1923. A Preliminary experimental confirmation of the two-film theory of gas absorption. Chemical and Metallurgical Engineering, 29, 146-148.

Zhang, W., Binns, M., Theodoropoulos, C., Kim, J.K., Smith, R., 2015. Model Building Methodology for Complex Reaction Systems. Industrial & Engineering Chemistry Research, 54(16), 4603-4615.

Chapter 3. Preparation works for strategy application

In the previous chapter, a methodologically systematic model-development strategy has been developed. In the practical application of such a strategy, the user has to perform different numerical and experimental steps including design and implementation of the experiments for different purposes, model development and evaluation. The involved numerical steps can be realized by the aid of homemade codes or commercial computer-aided softwares. Currently available softwares, namely, model parameter estimation solvers, such as ReactOp®, ProSimBatch®, μKE (Metaxas et al., 2010), etc., provide only the model-development functions, i.e. to construct automatically the model structure for a given reaction system description and estimate the model parameters using the experimental data. However, model-based experimental design functions are not covered by such softwares. Therefore, the user has to write codes to solve the real problems, and these codes have to be rewritten when a new reaction system is studied. In this context, implementation of the strategy in a software, which integrates the essential numerical steps and further makes it practically systematic, appears necessary.

In this chapter, preparation works for strategy application are presented. Firstly, the developed strategy-based software, facilitating significantly the user work, is presented. Secondly, the screening procedure of the validation reaction systems, used to demonstrate the feasibilities and generalities of the developed strategy and strategy-based software, is described.

3.1 Software description

The initial version of the strategy-based software, providing user-friendly interfaces and integrating model development, model evaluation, model-based experimental design for model refining and performance optimization, has been developed using MATLAB R2014a. In order to reduce software development difficulty, the following simplifications have been applied without losing the fundamental functions:

- The reaction supernetwork is developed through empirical method (see Section 2.2.1.1);
- The reaction network under test is proposed by the user on the basis of his expertise (see Section 2.2.1.1);
- The kinetic orders are not considered as the model parameters to be estimated (see Section 2.3.2);
- Model structure analysis is not integrated into the software;
- Model adequacy is qualitatively evaluated (see Section 2.4);

- If several Sets of Models Under Test (SMUTs) are available, the most suitable SMUT is directly identified (see Section 2.1.2).

Before software application, the user has to ensure that the case study is in the application range of the software in terms of reaction system type and reactor configuration. Specifically, the most common reaction systems for syntheses of fine and pharmaceutical chemicals, namely, liquid-phase reaction systems, including liquid, liquid-liquid and liquid-liquid-liquid reaction systems, are taken into consideration; reactor configuration should be either Stirred-Tank Reactor (STR) or Tubular Reactor (TR). Then, the user has to strictly follow the instructions given by the software. Software workflow is shown in Figure 3.1. Each module will be described in the following sections and illustrated while applying the strategy to the experimental case studies (see Sections 4.2 and 5.2).

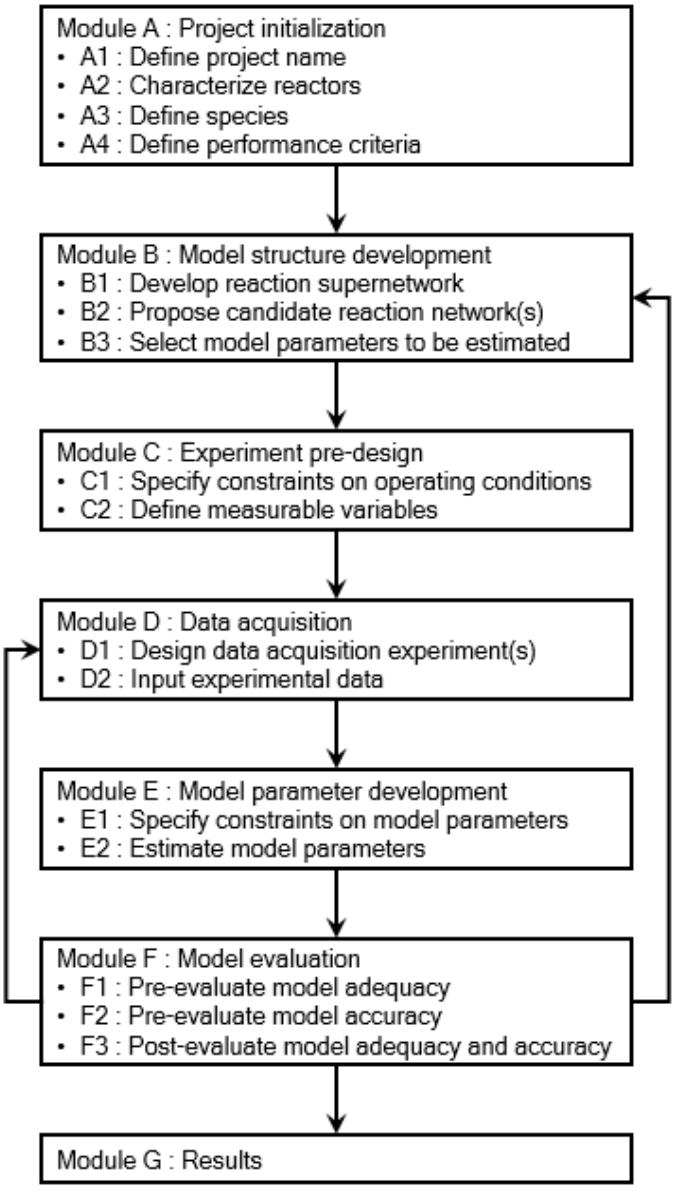


Figure 3.1: Software workflow.

3.1.1 Module A: Project initialization

The software requires the user to enter the basic information about the case study.

3.1.1.1 Submodule A1: Define project name

The user is asked to give a name to the case study. A file with this name is generated, and all the subsequent simulation results will be stored in it.

3.1.1.2 Submodule A2: Characterize reactors

The user has to:

1. indicate the number of considered reactors;
2. select for each reactor:
   - its configuration from STR and TR,
   - its process type from batch, semi-batch and continuous,
   - its thermal model from adiabatic, isothermal and isoperibolic;
3. define the dimensions for each reactor, as well as the overall heat-transfer coefficient and the heat-transfer surface for the reactor whose thermal mode is isoperibolic.

3.1.1.3 Submodule A3: Define species

The user has to:

1. indicate the number of all possible species and the number of all possible phases in the reaction system;
2. define for each species its molar mass, density (and specific heat capacity, if the thermal mode of at least one reactor is not isothermal), as well as its type from solvent, reactant, catalyst and product;
3. indicate for each species whether it is present in each possible phase.

Then, in the background, the following information is automatically generated:

- operating conditions in general form (illustrated in Table 3.1), that characterize all possible reactors. Note that the number of virtual feeds used for simulation is equal to that of all possible phases in the reaction system, not necessarily equal to that of actual feeds, which can be for instance illustrated by Figure 4.6;
- time-dependent state variables, from which the measurable variables are defined.

Table 3.1: Operating conditions in general form (× = needs to be defined).

Operating condition in general form	Symbol	Unit	BSTR	SBSTR	CSTR [a]	CTR
Experiment duration	τ	s	×	×	×	
Initial reaction temperature	T^{int}	K	×	×	×	
Initial load of reactants, solvents and catalysts	n_i^{int}	mol	×	×	×	
Feed temperature	T_j^{in}	K		×	×	×
Feed flow rate	F_j^{in}	m^3/s		×	×	×
Feed duration	t_j^{in}	s		×		
Concentration of reactants, solvents and catalysts in the feed	$c_{i,j}^{in}$	mol/m^3		×	×	×

[a] The behavior of the CSTR studied in this work is defined to be composed of two stages, i.e. transient stage followed by steady stage. The former one corresponds to the period from the initial loading to the formation of steady stage, which we generally refer to for the default CSTR.

3.1.1.4 Submodule A4: Define performance criteria

The user has to:

1. indicate the number of performance criteria;
2. define each performance criterion from the suggestion list including the following conventional performances:
   - residual concentration of one species,
   - total residual concentration of several species,
   - conversion of one reactant,
   - average conversion of several reactants,
   - selectivity of one product,
   - total selectivity of several products,
   - reaction duration,
   - maximal reaction temperature, for the reactors whose thermal mode is not isothermal;
3. define, for each performance, its objective and veto values.

Then, in the background, the corresponding experimental design criterion for performance optimization, constructed based on Eq. (1.20), is automatically generated.

3.1.2 Module B: Model structure development

The software requires the user to:

- enter a reaction supernetwork and indicate the type of each involved item, i.e. chemical reaction or mass transfer (B1);
- propose the reaction networks under test and define the corresponding kinetic orders (B2);

- select the Parameters to be Actually Estimated (PAEs) by indicating for each involved parameter (automatically generated according to the proposed reaction networks) its dependence from a suggestion list (B3). For instance, three suggested dependences for the reaction rate constant of the chemical reaction j are given below:
  - temperature-independence, which results that the reaction rate constant of the chemical reaction j itself is considered as the PAE;
  - temperature-dependence according to Eq. (2.15), which results that the pre-exponential factor and activation energy of the chemical reaction j are considered as the PAEs;
  - temperature-dependence according to Eqs. (2.15)-(2.17), which results that the reaction rate constant logarithms of the chemical reaction j at respectively the maximal and minimal possible reaction temperatures are considered as the PAEs.

Given two essential elements for model structure construction, namely, candidate reaction system descriptions and reactors specified in respectively this module and the submodule A2, the corresponding specific model structures are automatically derived in the background from the general model structures pre-integrated in the software, e.g. Eqs. (2.12)-(2.14) for the homogeneous catalytic reaction systems in STR.

3.1.3 Module C: Experimental pre-design

3.1.3.1 Submodule C1: Specify constraints on operating conditions

The user has to specify the bounds on operating conditions in general form (see Table 3.1) and the linear or nonlinear equality or inequality constraints among them. Note that the same operating conditions can be represented by the following three different manners:

- practical operating conditions, guiding the user for operating manipulations,
- practical operating variables, i.e. process impact factors,
- operating conditions in general form, which are accepted by the software.

It is important for the user to not confuse them. To illustrate the relationship among them, refer to Section 4.2.1.8, in which mutual transition of different representation manners of operating conditions for the case study 1 is given.

3.1.3.2 Submodule C2: Define measurable variables

The user has to define the measurable variables, which can be divided into two types:

- conventional measurable variables, i.e. measurable time-dependent state variables,

- unconventional measurable variables, i.e. measurable products of time-dependent state variables.

Considering that model structure analysis is not integrated into the present version of the software, the user has to ensure that, given the defined variables to be actually measured, all the candidate model structures constructed at the module B are identifiable and discriminable.

3.1.4 Module D: Data acquisition

The experimental data of the data acquisition experiment(s), including the operating conditions in general form and the measured responses, is acquired. For the preliminary experiments, the user has not only to input the experimental data into the software but also to define the measurement errors (on the basis of accuracy of measurement instruments and his expertise). For the data acquisition experiments for parameter precision improvement designed by the software, the user only has to input the measured responses into the software.

3.1.5 Module E: Model parameter development

The software first requires the user to specify reasonably the bounds on model parameters and the potential linear or nonlinear inequality constraints among them (E1), and then estimates the values and the confidence intervals of the model parameters from the currently available experimental data (E2).

For parameter estimation, if at least one parameter estimate reaches its bound, new parameter bounds will be constructed by bringing adaptive modifications to the current parameter bounds. Parameter estimation will be repeated until a set of model parameters, none of which reaching its bounds, is identified.

3.1.6 Module F: Model evaluation

Model evaluation is performed in this module. According to the model evaluation results, the various modules (shown in Table 3.2) will be repeated until a set of adequate and accurate models allowing fitting all experiments is identified.

Table 3.2: Submodules following and followed by various model-evaluation submodules.

Model evaluation in terms of adequacy and accuracy	Following	Followed by	
		If evaluation succeeds	If evaluation fails
Model adequacy pre-evaluation (F1)	E2	F2	B1 or B2
Model accuracy pre-evaluation (F2)	F1	F3	D1
Model adequacy and accuracy post-evaluations (F3)	F2	end	B1 or B2

The software is developed based on the strategy, each software (sub)module can be mapped to one or several steps of the strategy (see Section 2.1). Specific description about model-evaluation module, namely, repetition of the strategy description, is not given. For more details, refer to step 8 of the strategy for the submodule F1, steps 13-14 for the submodule F2, step 16 for the submodule F3.

3.1.7 Module G: Results

The set of adequate and accurate models (structures and parameters) allowing fitting all experiments comprising the data acquisition experiments and the validation experiments within an acceptable tolerance, as well as the optimal reactor with its associated optimal operating conditions for the given liquid-phase reaction system are displayed in this module.

3.2 Screening of validation reaction systems

In order to validate experimentally the feasibility and the generality of the strategy (as well as the practicability of the strategy-based software), at least two validation reaction systems with different characteristics, such as thermal mode, number of possible phases, are required. In this section, the screening procedure of validation reaction systems is described.

3.2.1 Screening criteria

There are no unified screening criteria, a compromise among the following three criteria is taken:

- Reaction system complexity: During the reaction, at any time, liquid-phase reaction system has only one uniformity, either homogeneity or heterogeneity. As the reaction progresses, if the reaction system is always homogeneous or heterogeneous, it is considered as a simple reaction system; if the homogeneous reaction system evolves into the heterogeneous one, or the heterogeneous reaction system evolves into the homogeneous one, it is considered as a complex reaction system. A simple reaction system is expected, since only one reaction system description (see Section 2.2.1), rather than several reaction system descriptions in serial connection (e.g. see Table 3.5), is required for model structure development.
- Reaction mechanism maturity: An immature reaction mechanism, i.e. still under debate, is expected.
- Experimental implementation difficulty: Experiments, easily, practically and economically available in terms of experimental reagents, set-up, conditions and protocol, are expected.

3.2.2 Candidate reaction systems

Eight candidate reaction systems (shown in Table 3.3) are described below. For the validation reaction systems finally selected (see Section 3.2.3), surveys on the reported reaction networks as well as the corresponding retro-developed reaction supernetworks are given. Hence, an independent section about literature review is not introduced while applying the strategy to the experimental case studies in Chapter 4 and Chapter 5.

Table 3.3: Candidate reaction systems.

Number of phases	Candidate reaction system
One	Acid perhydrolysis of carboxylic acid [a]
	First step of lorcaserin synthesis
Two	Saponification of methyl benzoate by soda
	Homogeneous alkali-catalyzed alcoholysis of vegetable oil [a]
	Epoxidation of vegetable oil by percarboxylic acid generated in situ [a]
	Liquid-liquid phase transfer catalysis [a]
Three	Liquid-liquid-liquid phase transfer catalysis [a]
From two to one	Synthesis of glycerol carbonate from glycerol and dimethyl carbonate

[a] A set of several reaction systems that can be summarized into one category, not a concrete reaction system.

3.2.2.1 Reaction system 1: Acid perhydrolysis of carboxylic acid

PerCarboxylic Acids (PCAs) are powerful and environmental friendly oxidizing agents used in chemical processing, synthesis and bleaching. Among PCAs, PerFormic Acid (PFA) (Zhao et al., 2007; Zhao et al., 2008) and PerAcetic Acid (PAA) (Filippis et al., 2009; Sun et al., 2011) are the most attractive ones. Generally, PCAs are prepared by reacting Carboxylic Acids (CAs) with Hydrogen Peroxide (HP) in the presence of homogeneous acid catalysts such as sulfuric acid and phosphoric acid. In this reaction system, the involved chemical reactions consist of perhydrolysis of CA, hydrolysis of PCA, probably decomposition of PCA and HP.

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Mature reaction mechanism,
- No particular requirements on experimental implementation.

3.2.2.2 Reaction system 2: First step of lorcaserin synthesis

Obesity is bothering more and more people. One can lose weight through physical exercise and appropriate diet. For the latter one, one can take a weight-loss drug, which reduces appetite by activating a type of serotonin receptor known as the 5-HT_{2C} receptor in the hypothalamus, which is

known to control appetite (Shukla et al., 2015). (1R)-8-chloro-1-methyl-2,3,4,5-tetrahydro-1H-3-benzazepine, also known as lorcaserin, is the active pharmaceutical ingredient of weight-loss drug, and can be synthesized through two steps (Grom et al., 2016):

1. 1-(2-bromoethyl)-4-chlorobenzene reacts with allylamine to produce N-(4-chlorophenethyl)prop-2-en-1-aminium chloride as intermediate, impurities such as 4-chlorostyrene, N,N-bis(4-chlorophenethyl)prop-2-en-1-amine being produced through parallel reactions. Allylamine is used as solvent and reagent concomitantly, whereas the reactions take place in a one-phase system.
2. electrophilic addition followed by electrophilic substitution of the intermediate, catalyzed by $AlCl_3$, generates lorcaserin.

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Immature reaction mechanism,
- Procedure for sample analysis is complex, multiple analytical instruments have to be used, such as nuclear magnetic resonance, Fourier transform infrared spectroscopy and gas chromatography.

3.2.2.3 Reaction system 3: Saponification of methyl benzoate by soda

In a heterogeneous liquid-liquid reaction system, the global transformation rate depends on both chemical reaction and mass transfer. If the reaction produces surfactant, this one may influence the mass transfer in two opposing manners: an increase of the interfacial area and a decrease of the overall mass-transfer coefficient. If the reaction also produces cosolvent, it will increase the solubility of the reagents. The total effect on the mass transfer will thus depend on the relative importance of the variation of these three parameters as reaction progresses. To illustrate this kind of reaction producing simultaneously a surfactant and a cosolvent, saponification of methyl benzoate by soda, producing sodium benzoate (surfactant) and methanol (cosolvent) is studied (Mouhab et al., 2008).

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Mature reaction mechanism,
- No particular requirements on experimental implementation.

3.2.2.4 Reaction system 4: Homogeneous alkali-catalyzed alcoholysis of vegetable oil

The diesel fuel consumption has been increasing continuously over the past decades and will continue in the future. Due to the depletion of petroleum resources and the environmental consequences of exhaust gases, alternative fuels for diesel engines are becoming significantly important. The alternatives to diesel fuel must be technically and environmentally acceptable, economically competitive and readily available (Stamenković et al., 2008). Biodiesel, which is composed mainly of fatty acid alkyl esters, and produced from renewable sources such as vegetable oils, animal fats, and even waste cooking oils by transesterification with an alcohol (alcoholysis), is an attractive alternative due to its environmental benefits: less greenhouse effect, less pollution of air, water and soil and less health risk, compared to diesel fuels.

Transesterification of triglycerides with alcohols in the presence of catalysts is the most common process for producing biodiesel. In the commercial production of biodiesel, the most commonly used triglyceride feedstocks, alcohols and catalysts are respectively vegetable oils such as rapeseed (also called colza), soybean, palm and sunflower oils; low molecular weight alcohols such as methanol and ethanol; homogeneous base catalysts such as potassium/sodium hydroxide/alkoxides. The most important operating variables that influence the alcoholysis reaction are the type and the amount of catalyst, the alcohol-to-oil molar ratio, the reaction temperature, the stirring intensity and the purity of reactants.

The intrinsic reaction mechanism of the homogeneous alkali-catalyzed alcoholysis of vegetable oil is complicated because of the following reasons:

- Vegetable oil, a complex mixture, consists mainly of triglycerides of four fatty acids, namely, linoleic, oleic, stearic and palmitic acids. The rest is composed by triglycerides of other fatty acids, such as linolenic and palmitoleic acids, diglycerides and monoglycerides of these fatty acids, free fatty acids, water and traces of sterols and vitamin E (Komers et al., 2002).
- Different reactions take place in the reaction system, such as alcoholysis of glycerides, saponification of glycerides and fatty acid alkyl esters, neutralization of free fatty acids (Eze et al., 2014);
- Alcoholysis reaction is heterogeneous during the whole course. In the studies of the alcoholysis process rate, three regimes are well-recognized: an initial mass transfer-controlled regime (slow), followed by a chemically-controlled regime (fast), and a final regime, close to equilibrium (slow) (Stamenković et al., 2008).

In order to reduce the modeling difficulty, practically, the following assumptions are usually introduced:

- Vegetable oil consists of only one triglyceride, where the three fatty acids are identical, and potentially one free fatty acid;
- The reaction system is considered controlled by: (i) one regime: heterogeneous regime (Richard et al., 2013) or pseudo-homogeneous regime (Darnoko and Cheryan, 2000; Komers et al., 2002; Berchmans et al., 2013; Likozar and Levec, 2014b; Eze et al., 2014; Reyero et al., 2015); (ii) two regimes: heterogeneous regime followed by pseudo-homogeneous regime (Likozar and Levec, 2014a; Likozar et al., 2016).

Table 3.4: Reaction supernetwork for the homogeneous alkali-catalyzed alcoholysis of vegetable oil.

N°	Reaction	Comments
1	$A(I) + AMH(I) \rightarrow AMA(I) + W(I)$	Hydroxide–alkoxide equilibrium reactions
2	$AMA(I) + W(I) \rightarrow A(I) + AMH(I)$	
3	$TG(I) + A(I) \xrightarrow{Cat.} FAAE(I) + DG(I)$	Transesterification of glycerides
4	$DG(I) + A(I) \xrightarrow{Cat.} FAAE(I) + MG(I)$	
5	$MG(I) + A(I) \xrightarrow{Cat.} FAAE(I) + G(I)$	
6	$FAAE(I) + DG(I) \xrightarrow{Cat.} TG(I) + A(I)$	
7	$FAAE(I) + MG(I) \xrightarrow{Cat.} DG(I) + A(I)$	
8	$FAAE(I) + G(I) \xrightarrow{Cat.} MG(I) + A(I)$	
9	$TG(I) + AMH(I) \rightarrow AMSFA(I) + DG(I)$	Saponification of glycerides and FAAE
10	$DG(I) + AMH(I) \rightarrow AMSFA(I) + MG(I)$	
11	$MG(I) + AMH(I) \rightarrow AMSFA(I) + G(I)$	
12	$FAAE(I) + AMH(I) \rightarrow AMSFA(I) + A(I)$	
13	$FFA(I) + AMH(I) \rightarrow AMSFA(I) + W(I)$	Neutralization of FFA
14	$FFA(I) + AMA(I) \rightarrow AMSFA(I) + A(I)$	
15	$TG(I) \rightarrow TG(II)$	Mass transfer
16	$DG(I) \rightarrow DG(II)$	
17	$MG(I) \rightarrow MG(II)$	
18	$FAAE(I) \rightarrow FAAE(II)$	
19	$A(I) \rightarrow A(II)$	
20	$G(I) \rightarrow G(II)$	
21	$FFA(I) \rightarrow FFA(II)$	
22	$W(I) \rightarrow W(II)$	

On the basis of the above assumption about the composition of the vegetable oil, in the reaction system of the homogeneous alkali-catalyzed alcoholysis of vegetable oil, all possible species are Alcohol (A), Alkali-Metal Hydroxide (AMH), Alkali-Metal Alkoxide (AMA), Water (W), Tri-Glyceride (TG), Di-Glyceride (DG), Mono-Glyceride (MG), Free Fatty Acid (FFA), Fatty Acid

Alkyl Ester (FAAE), Glycerol (G) and Alkali-Metal Salt of Fatty Acid (AMSFA). The corresponding reaction supernetwork, comprising 14 chemical reactions (N°1-14) and 8 mass transfers (N°15-22) in which the species can be transferred between the polar phase (I) and the nonpolar phase (II), has been developed through empirical method (see Table 3.4). It covers all the reaction networks reported in the literature (see Table 3.5).

Table 3.5: Survey on the reported reaction networks for the homogeneous alkali-catalyzed alcoholysis of vegetable oil.

Reaction system type	Reaction stage	Reactions taken into consideration	References
Simple		3, 5, 7	Darnoko and Cheryan, 2000
		3-8	Likozar and Levec, 2014b
		3-12	Komers et al., 2002
		3-13	Berchmans et al., 2013
		1-12	Reyero et al., 2015
		1-14	Eze et al., 2014
		3-8, 15-17	Richard et al., 2013
Complex	1	3, 5, 7, 15	Stamenković et al., 2008
	2	3, 5, 7	
	3	3-8	
	1	3-8, 15-20	Likozar and Levec, 2014a; Likozar et al., 2016
	2	3-8	

The results of this reaction system assessment in terms of screening criteria are:

- Theoretically, reaction system is complex, however in fact, it can be considered as a simple one under reasonable hypotheses,
- Immature reaction mechanism,
- No particular requirements on experimental implementation.

3.2.2.5 Reaction system 5: Epoxidation of vegetable oil by percarboxylic acid generated in situ

Epoxidized Vegetable Oils (EVOs) are mainly used as plasticizers and stabilizers for PVC resins. The EVO demand is expected to increase very much in the future, since phthalates, conventional plasticizers for PVC resins, have been banned in many countries for their negative effects on health.

Epoxidation of vegetable oil by PerCarboxylic Acid (PCA) generated in situ is the most common process for producing EVOs. In this reaction system, the involved reactions consist of:

- Synthesis of PCA from Carboxylic Acid (CA) and Hydrogen Peroxide (HP) in the presence of homogeneous acid catalyst (see Section 3.2.2.1);
- Epoxidation of vegetable oil, i.e. Double-Bond-containing compound (DB), by PCA resulting in the formation of Epoxide (Ep);

- Epoxide ring-opening reactions by nucleophilic agents, such as CA, HP, PCA, water (W), resulting in the formation of hydroxyl-containing decomposition compound (OH), whose hydroxyl group can react further with Ep to give internal ethers or oligomeric ethers, such as Dimer (Dim) and Trimer (Trim) (Petrović et al., 2002);
- Nucleophilic rearrangement of Ep resulting in Ketone (Ket) (Rangarajan et al., 1995; Petrović et al., 2002);
- Formation of decomposition compound by water (OH_W) from decomposition compound by CA (OH_{CA}) by acid hydrolysis (Janković and Sinadinovic-Fiser, 2004).

Epoxide ring-opening reactions can take place at the aqueous-oil interface and in the oil phase. Their mechanisms are different: at the aqueous-oil interface, the reaction can be considered as acid-catalyzed nucleophilic substitution composed of two steps: the attack of Ep by proton resulting in the formation of carbocation, then, the attack of carbocation by nucleophilic agents resulting in the formation of OH, whereas in the oil phase, the reaction involves the addition of nucleophilic agents to Ep via a hydrogen bond (Campanella and Baltanas, 2006).

In view of the information available in the literature as well as the assumption about the composition of the vegetable oil (see Section 3.2.2.4), all possible species in this reaction system are: CA, HP, PCA, W, catalyst, DB, Ep, OH_{CA}, decomposition compound by HP (OH_{HP}), decomposition compound by PCA (OH_{PCA}), OH_W, Ket, Dim, Trim, carbon dioxide (CO_2) and oxygen (O_2). The corresponding reaction supernetwork, comprising 25 chemical reactions (N°1-25) and 4 mass transfers (N°26-29) in which the species can be transferred between the aqueous phase (I) and the oil phase (II), has been developed through empirical method (see Table 3.6). It covers all the reaction networks reported in the literature (see Table 3.7).

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Immature reaction mechanism particularly about epoxide ring-opening reactions,
- No particular requirements on experimental implementation.

Table 3.6: Reaction supernetwork for the epoxidation of vegetable oil by percarboxylic acid generated in situ.

N°	Apparent Reaction	Comments
1	$CA(I) + HP(I) \xrightarrow{H^+(I)} PCA(I) + W(I)$	Acid perhydrolysis of CA
2	$PCA(I) + W(I) \xrightarrow{H^+(I)} CA(I) + HP(I)$	Acid hydrolysis and perhydrolysis of PCA
3	$PCA(I) + HP(I) \xrightarrow{H^+(I)} CA(I) + W(I) + O_2(g)$	
4	$PCA(I) \xrightarrow{H^+(I)} CA(I) + 1/2O_2(g)$	Decomposition of PCA and HP
5	$PCA(I) \rightarrow W(I) + CO_2(g)$	
6	$PCA(I) \rightarrow 2/3CA(I) + 1/3W(I) + 1/3O_2(g) + 1/3CO_2(g)$	
7	$HP(I) \xrightarrow{H^+(I)} W(I) + 1/2O_2(g)$	
8	$DB(II) + PCA(II) \rightarrow Ep(II) + CA(II)$	Epoxidation of DB
9	$Ep(II) + CA(I) \xrightarrow{H^+(I)} OH_{CA}(II)$	Epoxide ring-opening
10	$Ep(II) + HP(I) \xrightarrow{H^+(I)} OH_{HP}(II)$	
11	$Ep(II) + PCA(I) \xrightarrow{H^+(I)} OH_{PCA}(II)$	
12	$Ep(II) + W(I) \xrightarrow{H^+(I)} OH_{W}(II)$	
13 [a]	$OH(II) \xrightarrow{H^+(I)} OH(II)$	
14 [b]	$Ep(II) + OH(II) \xrightarrow{H^+(I)} Dim(II)$	
15 [c]	$OH(II) + OH(II) \xrightarrow{H^+(I)} Dim(II)$	
16 [a]	$Dim(II) \xrightarrow{H^+(I)} Dim(II)$	
17 [b]	$Ep(II) + Dim(II) \xrightarrow{H^+(I)} Trim(II)$	
18 [c]	$OH(II) + Dim(II) \xrightarrow{H^+(I)} Trim(II)$	
19 [a]	$Trim(II) \xrightarrow{H^+(I)} Trim(II)$	
20	$Ep(II) \xrightarrow{H^+(I)} Ket(II)$	
21	$Ep(II) + CA(II) \rightarrow OH_{CA}(II)$	
22	$Ep(II) + HP(II) \rightarrow OH_{HP}(II)$	
23	$Ep(II) + PCA(II) \rightarrow OH_{PCA}(II)$	
24	$Ep(II) + W(II) \rightarrow OH_{W}(II)$	
25	$OH_{CA}(II) + W(I) \xrightarrow{H^+(I)} OH_{W}(II) + CA(I)$	Acid hydrolysis of OH_{CA}
26	$CA(I) \rightarrow CA(II)$	Mass transfer
27	$HP(I) \rightarrow HP(II)$	
28	$PCA(I) \rightarrow PCA(II)$	
29	$W(I) \rightarrow W(II)$	

[a] The undecomposed epoxide group reacts with the decomposed epoxide group, i.e. the hydroxyl group, in the same molecule to give internal ether.

[b] Ep reacts with OH (including Dim) to produce oligomeric ether.

[c] The hydroxyl group in OH reacts with the undecomposed epoxide group in OH (including Dim) to produce oligomeric ether.

Table 3.7: Survey on the reported reaction networks for the epoxidation of vegetable oil by percarboxylic acid generated in situ.

N° Reaction network	Chemical reactions and mass transfers taken into consideration				References
	Essential reaction	Decomposition	ring-opening	Mass transfer	
1	1, 2, 8	-	21	26, 28	Rangarajan et al., 1995
2	1, 2, 8	7	9-12	26-29	Santacesaria et al., 2011
3	1, 2, 8	7	9-12	26, 28	Santacesaria et al., 2012
4	1, 2, 8	5	12	26, 28	Leveneur et al., 2014
5	1, 2, 8	7	12, 14, 17	-	De Haro et al., 2016
6	1, 2, 8	6	9-12	26, 28	Wu et al., 2016
7	1, 2, 8	4, 5	9, 11, 12	26, 28	Zheng et al., 2016
8	1, 2, 8	6, 7	9-12	26-29	Moreno et al., 2017

3.2.2.6 Reaction system 6: Liquid-liquid phase transfer catalysis

Phase Transfer Catalysis (PTC) is an attractive technique for synthesis of organic chemicals from two reactants existing in two different phases, which normally cannot react with each other due to their low mutual solubility in other phase and hence their low interaction (Maity et al., 2009). Liquid-Liquid PTC (LLPTC) is the most conventional, because of the mild operating conditions and the use of cheap solvents, and can be used for synthesis of the organic chemicals, e.g. phenyl benzoate (Yang and Huang, 2006a), dibenzyl sulfide and benzyl mercaptan (Maity et al., 2009), thioester (Simion et al., 2010), distyryl derivates (Zhao et al., 2015a). However, its drawbacks are also obvious, for instance, the homogeneous catalyst is normally not reused and a large quantity of water is required to wash the organic phase (Zhao et al., 2015b).

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Mature reaction mechanism, but difficult for understanding,
- Usually, multiple analytical instruments are required.

3.2.2.7 Reaction system 7: Liquid-liquid-liquid phase transfer catalysis

A Liquid-Liquid-Liquid PTC (LLLPTC) system contains a separate catalyst or catalytic intermediate rich phase located between the aqueous and organic phases, and can be used for synthesis of organic chemicals, e.g. benzyl salicylate (Yang and Li, 2006b), benzyl thiocyanate (Yadav and Sowbna, 2012a), mandelic acid (Yadav and Sowbna, 2012b), stilbene (Zhao et al., 2015b). The advantages of LLLPTC over LLPTC are: (i) enhanced reaction rate, (ii) milder operating conditions, (iii) easier catalyst recovery and reuse, (iv) increased yield and selectivity. The reaction mechanisms of LLLPTC are similar and can be derived from the most basic one

consisting of four steps: (i) synthesis of catalytic intermediate from aqueous reactant and phase-transfer catalyst in the aqueous phase; (ii) transfer of catalytic intermediate and organic reactant from respectively the aqueous and organic phases to the third liquid phase; (iii) synthesis of desired product from catalytic intermediate and organic reactant in the third liquid phase; (iv) transfer of regenerated phase-transfer catalyst and desired product from the third liquid phase to respectively the aqueous and organic phases.

The results of this reaction system assessment in terms of screening criteria are:

- Simple reaction system,
- Mature reaction mechanism,
- No particular requirements on experimental implementation.

3.2.2.8 Reaction system 8: Synthesis of glycerol carbonate from glycerol and dimethyl carbonate

Following the significant development of the biodiesel industry, the price of its by-product, glycerol, has lowered. Thus, currently, research on the synthesis of high added value chemicals derived from glycerol has gained more attention. Glycerol Carbonate (GC), an important glycerol derivative, has proven useful in many applications due to its high boiling point, low toxicity and good biodegradability (Ochoa-Gómez et al., 2009), and can be synthesized by phosgenation, direct carboxylation and transesterification of glycerol or by glycerolysis of urea. Among these synthesis routes, transesterification is the simplest one. In the reaction system of GC synthesis from glycerol and dimethyl carbonate by transesterification using homogeneous basic catalysts, e.g. potassium carbonate (Esteban et al., 2015a) and potassium methoxide (Esteban et al., 2015b), initially, glycerol and dimethyl carbonate have very low miscibility with each other, which results in an emulsion-like liquid–liquid biphasic system. Nevertheless, as the reaction takes place and generates products, the system evolves into a single phase liquid.

The results of this reaction system assessment in terms of screening criteria are:

- Complex reaction system, distinct kinetic models are required;
- Immature reaction mechanism;
- Focused beam reflectance measurement for online monitoring of the droplet size is required.

3.2.3 Validation reaction systems

The results of reaction system assessment in terms of screening criteria are summarized in Table 3.8. It can been seen that the 4th and 5th reaction systems, related to the valorization of vegetable oils,

namely, homogeneous alkali-catalyzed alcoholysis of vegetable oil and epoxidation of vegetable oil by PCA generated in situ, validate all criteria and can be selected as the validation ones.

Table 3.8: Result of candidate reaction system assessment in terms of screening criteria (+ = expected; ± = neutral; − = unexpected).

N° reaction system	1	2	3	4	5	6	7	8
Reaction system complexity	+	+	+	+	+	+	+	−
Reaction system description maturity	−	+	−	+	+	±	−	+
Experimental implementation difficulty	+	−	+	+	+	±	+	±
Overall assessment result	−	−	−	+	+	±	−	−

Both reaction systems have the following characteristics:

1. Simple reaction system, i.e. one reaction system description is required;
2. Main reaction mechanisms are clear and fairly representative, one is a substitution, the other is an oxidation. Meanwhile, side reaction mechanisms are still under debate and have to be explored;
3. Experiments are easily, practically and economically available in terms of experimental reagents, set-up, conditions and protocol, specifically,
   - involved reagents are very common and cheap;
   - operating conditions are mild, e.g. reaction temperature between 20°C and 85°C;
   - the experimental set-up (see Figure 4.1) in our laboratory will allow to carry out the experiments;
   - one analytical instrument is used for sample analysis, e.g. gas or liquid chromatograph for alcoholysis, automatic titrator for epoxidation.

The reactants and catalyst for each synthesis should be selected to obtain the concrete reaction systems. The same vegetable oil is used for both syntheses. Sunflower oil is selected as oil feedstock among the common edible oils in supermarkets. In the following sections, the concrete reaction systems will be finally obtained after choosing the reactants and catalysts used.

3.2.3.1 Case study 1: NaOH-catalyzed ethanolysis of sunflower oil

A concrete reaction system can be obtained after choosing:

- the alcohol between the two most common alcohols, i.e. methanol and ethanol;
- the base catalyst among sodium hydroxide, sodium ethoxide, potassium hydroxide and potassium ethoxide.

Methanol, which is mainly produced by oxidation processes of methane, the main constituent of natural gas, is the most frequently used alcohol for the synthesis of biodiesel. Ethanol, particularly

bioethanol, which can be obtained from renewable sources, such as sugar cane and sugar beet, is preferable to methanol due to its superior dissolving power for vegetable oils, low toxicity and its renewable origin. Compared with the fatty acid methyl esters, fatty acid ethyl esters exhibit some improved fuel properties, such as higher stability toward oxidation, better lubricity, lower iodine value and cloud and pour points thus improving the performance in cold weather. In the view of the above facts (Richard et al., 2013; Reyero et al., 2015), ethanol is selected as alcohol feedstock. From the view of economy and practice, NaOH is selected as catalyst. Therefore, the selected alcoholysis system is NaOH-catalyzed ethanolysis of sunflower oil.

3.2.3.2 Case study 2: Epoxidation of sunflower oil by performic acid generated in situ

A concrete reaction system can be obtained after choosing:

- the PCA between the two most common PCAs, i.e. PFA and PAA;
- the homogeneous acid catalyst between the reported common acids, i.e. sulfuric acid and phosphoric acid.

It is necessary to introduce the difference between the use of PFA and PAA. As shown by the literature review, epoxidation by PAA, slower and less exothermic than by PFA, was studied essentially. The reaction network of epoxidation by PFA is more complicated in the view of the following facts:

- PFA is less stable than PAA, the decomposition of PFA cannot be neglected, and furthermore, both the spontaneous decomposition and the radical-introduced decomposition contribute to the decomposition of PFA (Sun et al., 2011);
- FA and PFA, stronger nucleophilic agents, favor epoxide ring-opening reactions.

In the view of the above facts, epoxidation of sunflower oil by PFA generated in situ is selected. The most common acid, namely, sulfuric acid, is used to catalyze FA perhydrolysis.

3.3 Conclusion

Two necessary preparation works for strategy application and validation have been presented. Specifically, (i) based on the systematic model-development strategy, a computer-aided software, making model development easier and faster, has been developed; (ii) Eight candidate reaction systems have been assessed in terms of reaction system complexity, reaction mechanism maturity and experimental implementation difficulty. Two reaction systems, related to the valorization of sunflower oil, namely, NaOH-catalyzed ethanolysis of sunflower oil and epoxidation of sunflower

oil by PFA generated in situ, have been selected as the validation case studies. Hence, in the next chapter, strategy application for the first case study will be presented.

Reference

Berchmans, H.J., Morishita, K., Takarada, T., 2013. Kinetic study of hydroxide-catalyzed methanolysis of Jatropha curcas–waste food oil mixture for biodiesel production. Fuel, 104, 46-52.

Campanella, A., Baltanás, M.A., 2006. Degradation of the oxirane ring of epoxidized vegetable oils in liquid–liquid heterogeneous reaction systems. Chemical Engineering Journal, 118(3), 141-152.

Darnoko, D., Cheryan, M., 2000. Kinetics of palm oil transesterification in a batch reactor. Journal of the American Oil Chemists' Society, 77(12), 1263-1267.

De Haro, J.C., Izarra, I., Rodríguez, J.F., Pérez, Á., Carmona, M., 2016. Modelling the epoxidation reaction of grape seed oil by peracetic acid. Journal of cleaner production, 138, 70-76.

Esteban, J., Fuente, E., Blanco, A., Ladero, M., Garcia-Ochoa, F., 2015a. Phenomenological kinetic model of the synthesis of glycerol carbonate assisted by focused beam reflectance measurements. Chemical Engineering Journal, 260, 434-443.

Esteban, J., Domínguez, E., Ladero, M., Garcia-Ochoa, F., 2015b. Kinetics of the production of glycerol carbonate by transesterification of glycerol with dimethyl and ethylene carbonate using potassium methoxide, a highly active catalyst. Fuel Processing Technology, 138, 243-251.

Eze, V.C., Phan, A.N., Harvey, A.P., 2014. A more robust model of the biodiesel reaction, allowing identification of process conditions for significantly enhanced rate and water tolerance. Bioresource technology, 156, 222-231.

Filippis, P.D., Scarsella, M., Verdone, N., 2009. Peroxyformic acid formation: a kinetic study. Industrial & Engineering Chemistry Research, 48(3), 1372-1375.

Grom, M., Stavber, G., Drnovšek, P., Likozar, B., 2016. Modelling chemical kinetics of a complex reaction network of active pharmaceutical ingredient (API) synthesis with process optimization for benzazepine heterocyclic compound. Chemical Engineering Journal, 283, 703-716.

Janković, M.R., Sinadinović-Fišer, S.V., 2004. Kinetic models of reaction systems for the in situ epoxidation of unsaturated fatty acid esters and triglycerides. Hemijska industrija, 58(12), 569-576.

Komers, K., Skopal, F., Stloukal, R., Machek, J., 2002. Kinetics and mechanism of the KOH—catalyzed methanolysis of rapeseed oil for biodiesel production. European Journal of Lipid Science and Technology, 104(11), 728-737.

Leveneur, S., Zheng, J., Taouk, B., Burel, F., Wärnå, J., Salmi, T., 2014. Interaction of thermal and kinetic parameters for a liquid–liquid reaction system: Application to vegetable oils epoxidation by peroxycarboxylic acid. Journal of the Taiwan Institute of Chemical Engineers, 45(4), 1449-1458.

Likozar, B., Levec, J., 2014a. Effect of process conditions on equilibrium, reaction kinetics and mass transfer for triglyceride transesterification to biodiesel: experimental and modeling based on fatty acid composition. Fuel Processing Technology, 122, 30-41.

Likozar, B., Levec, J., 2014b. Transesterification of canola, palm, peanut, soybean and sunflower oil with methanol, ethanol, isopropanol, butanol and tert-butanol to biodiesel: Modelling of chemical equilibrium, reaction kinetics and mass transfer based on fatty acid composition. Applied Energy, 123, 108-120.

Likozar, B., Pohar, A., Levec, J., 2016. Transesterification of oil to biodiesel in a continuous tubular reactor with static mixers: modelling reaction kinetics, mass transfer, scale-up and optimization considering fatty acid composition. Fuel Processing Technology, 142, 326-336.

Metaxas, K., Thybaut, J.W., Morra, G., Farrusseng, D., Mirodatos, C., Marin, G.B., 2010. A microkinetic vision on high-throughput catalyst formulation and optimization: development of an appropriate software tool. Topics in Catalysis, 53(1-2), 64-76.

Maity, S.K., Sen, S., Pradhan, N.C., 2009. A new mechanistic model for liquid–liquid phase transfer catalysis: Reaction of benzyl chloride with aqueous ammonium sulfide. Chemical Engineering Science, 64(21), 4365-4374.

Moreno, V.C., Russo, V., Tesser, R., Di Serio, M., Salzano, E., 2017. Thermal risk in semi-batch reactors: The epoxidation of soybean oil. Process Safety and Environmental Protection, 109, 529-537.

Mouhab, N., Balland, L., Talouba, I.B., Cosmao, J.M., 2008. Study of a chemical reaction in heterogeneous liquid–liquid medium producing a surfactant and a cosolvent. Chemical Engineering and Processing: Process Intensification, 47(3), 363-369.

Ochoa-Gómez, J.R., Gómez-Jiménez-Aberasturi, O., Maestro-Madurga, B., Pesquera-Rodríguez, A., Ramírez-López, C., Lorenzo-Ibarreta, L., Villarán-Velasco, M.C., 2009. Synthesis of glycerol carbonate from glycerol and dimethyl carbonate by transesterification: catalyst screening and reaction optimization. Applied Catalysis A: General, 366(2), 315-324.

Petrović, Z.S., Zlatanić, A., Lava, C.C., Sinadinović-Fišer, S., 2002. Epoxidation of soybean oil in toluene with peroxoacetic and peroxoformic acids—kinetics and side reactions. European Journal of Lipid Science and Technology, 104(5), 293-299.

Rangarajan, B., Havey, A., Grulke, E.A., Culnan, P.D., 1995. Kinetic parameters of a two-phase model forin situ epoxidation of soybean oil. Journal of the American Oil Chemists' Society, 72(10), 1161-1169.

Reyero, I., Arzamendi, G., Zabala, S., Gandía, L.M., 2015. Kinetics of the NaOH-catalyzed transesterification of sunflower oil with ethanol to produce biodiesel. Fuel Processing Technology, 129, 147-155.

Richard, R., Thiebaud-Roux, S., Prat, L., 2013. Modelling the kinetics of transesterification reaction of sunflower oil with ethanol in microreactors. Chemical Engineering Science, 87, 258-269.

Santacesaria, E., Tesser, R., Di Serio, M., Turco, R., Russo, V., Verde, D., 2011. A biphasic model describing soybean oil epoxidation with H2O2 in a fed-batch reactor. Chemical engineering journal, 173(1), 198-209.

Santacesaria, E., Renken, A., Russo, V., Turco, R., Tesser, R., Di Serio, M., 2012. Biphasic model describing soybean oil epoxidation with H2O2 in continuous reactors. Industrial & Engineering Chemistry Research, 51(26), 8760-8767.

Shukla, A.P., Kumar, R.B., Aronne, L.J., 2015. Lorcaserin Hcl for the treatment of obesity. Expert opinion on pharmacotherapy, 16(16), 2531-2538.

Simion, C., Hashimoto, I., Mitoma, Y., Simion, A.M., Egashira, N., 2010. Rapid and convenient thioester synthesis under phase-transfer catalysis conditions. Phosphorus, Sulfur, and Silicon, 185(12), 2480-2488.

Stamenković, O.S., Todorović, Z.B., Lazić, M.L., Veljković, V.B., Skala, D.U., 2008. Kinetics of sunflower oil methanolysis at low temperatures. Bioresource Technology, 99(5), 1131-1140.

Sun, X., Zhao, X., Du, W., Liu, D., 2011. Kinetics of formic acid-autocatalyzed preparation of performic acid in aqueous phase. Chinese Journal of Chemical Engineering, 19(6), 964-971.

Wu, Z., Nie, Y., Chen, W., Wu, L., Chen, P., Lu, M., Ji, J., 2016. Mass transfer and reaction kinetics of soybean oil epoxidation in a formic acid-autocatalyzed reaction system. The Canadian Journal of Chemical Engineering, 94(8), 1576-1582.

Yadav, G.D., Sowbna, P.R., 2012a. Modeling of microwave irradiated liquid–liquid–liquid (MILLL) phase transfer catalyzed green synthesis of benzyl thiocyanate. Chemical Engineering Journal, 179, 221-230.

Yadav, G.D., Sowbna, P.R., 2012b. Process intensification and waste minimization in liquid–liquid–liquid phase transfer catalyzed selective synthesis of mandelic acid. Chemical Engineering Research and Design, 90(9), 1281-1291.

Yang, H.M., Huang, C.C., 2006a. Kinetics for benzoylation of sodium phenoxide by liquid–liquid phase-transfer catalysis. Applied Catalysis A: General, 299, 258-265.

Yang, H.M., Li, C.C., 2006b. Kinetics for synthesizing benzyl salicylate by third-liquid phase-transfer catalysis. Journal of Molecular Catalysis A: Chemical, 246(1-2), 255-262.

Zhao, X., Zhang, T., Zhou, Y., Liu, D., 2007. Preparation of peracetic acid from hydrogen peroxide: Part I: Kinetics for peracetic acid synthesis and hydrolysis. Journal of Molecular Catalysis A: Chemical, 271(1-2), 246-252.

Zhao, X., Cheng, K., Hao, J., Liu, D., 2008. Preparation of peracetic acid from hydrogen peroxide, part II: Kinetics for spontaneous decomposition of peracetic acid in the liquid phase. Journal of Molecular Catalysis A: Chemical, 284(1-2), 58-68.

Zhao, Q., Sun, J., Li, F., He, J., Liu, B., 2015a. Mechanism and kinetics of Horner–Wadsworth–Emmons reaction in liquid–liquid phase-transfer catalytic system. Journal of Molecular Catalysis A: Chemical, 400, 111-120.

Zhao, Q., Sun, J., Liu, B., He, J., 2015b. Novel kinetics model for third-liquid phase-transfer catalysis system of the “complex” carbanion: Competitive role between catalytic cycles. Chemical Engineering Journal, 280, 782-795.

Zheng, J.L., Wärnå, J., Salmi, T., Burel, F., Taouk, B., Leveneur, S., 2016. Kinetic modeling strategy for an exothermic multiphase reactor system: Application to vegetable oils epoxidation using P rileschajew method. AIChE Journal, 62(3), 726-741.

Chapter 4. Case study 1: NaOH-catalyzed ethanolysis of sunflower oil

In the previous chapter, necessary preparation works for strategy application have been presented: a strategy-based software has been developed; two validation reaction systems have been selected. In this chapter, the feasibility of the strategy and the practicability of the software are demonstrated by the first case study, i.e. NaOH-catalyzed ethanolysis of sunflower oil, which has been published together with the presentation of the strategy and the strategy-based software (Jiang et al., 2019).

This chapter starts with the presentation of the experimental aspects, focuses on the chronological application of the strategy to the ethanolysis case study and ends with strategy feasibility and software practicability analyses. Literature review about reaction system is not included in this chapter, since it has been given in Section 3.2.2.4, while screening the validation reaction systems.

4.1 Materials and methods

In this section, firstly, the materials and the experimental set-up used for the ethanolysis experiments are reported. Then, the experimental protocols and the analytical methods will be described.

4.1.1 Materials

Commercial sunflower oil is purchased from Auchan (Lobau), a local supermarket. Its fatty acid composition (% by weight) is measured as follow: 11% saturated acids, 27% mono-unsaturated acids, 62% poly-unsaturated acids. The other reagents used in this case study are shown in Table 4.1.

Table 4.1: Information about the reagents used for the case study 1.

Reagent	Grade	Supplied by	Used for
Ethanol	≥99.8%	Sigma-Aldrich	Synthesis
Sodium hydroxide	≥99.3%, pellets	Fisher Chemical	
Hydrochloric acid	Analytical, 37% solution in water	Sigma-Aldrich	Treating the samples, i.e. quenching the reactions
MonoC19 [a]	>99.9%	Nu-Chek Prep	Analyzing the compositions of glycerides in the treated light phase
DiC38 [b]	>99.9%	Nu-Chek Prep	
TriC57 [c]	>99.9%	Nu-Chek Prep	
Pyridine	>99%	Sigma-Aldrich	
Heptane	>99%	Sigma-Aldrich	
MSTFA [d]	for GC derivatization, ≥98.5%	Sigma-Aldrich	
Ethyl nonadecanoate	≥99.5%	Sigma-Aldrich	Analyzing the compositions of esters in the treated light phase
Toluene	Analytical quality	Sigma-Aldrich	

[a] Monononadecanoin; [b] Dinonadecanoin; [c] Trinonadecanoin; [d] N-Methyl-N-(trimethylsilyl)trifluoroacetamide.

4.1.2 Set-up

The used experimental set-up (shown in Figure 4.1) includes two reactor configurations, i.e. Stirred-Tank Reactor (STR) and Tubular Reactor (TR).

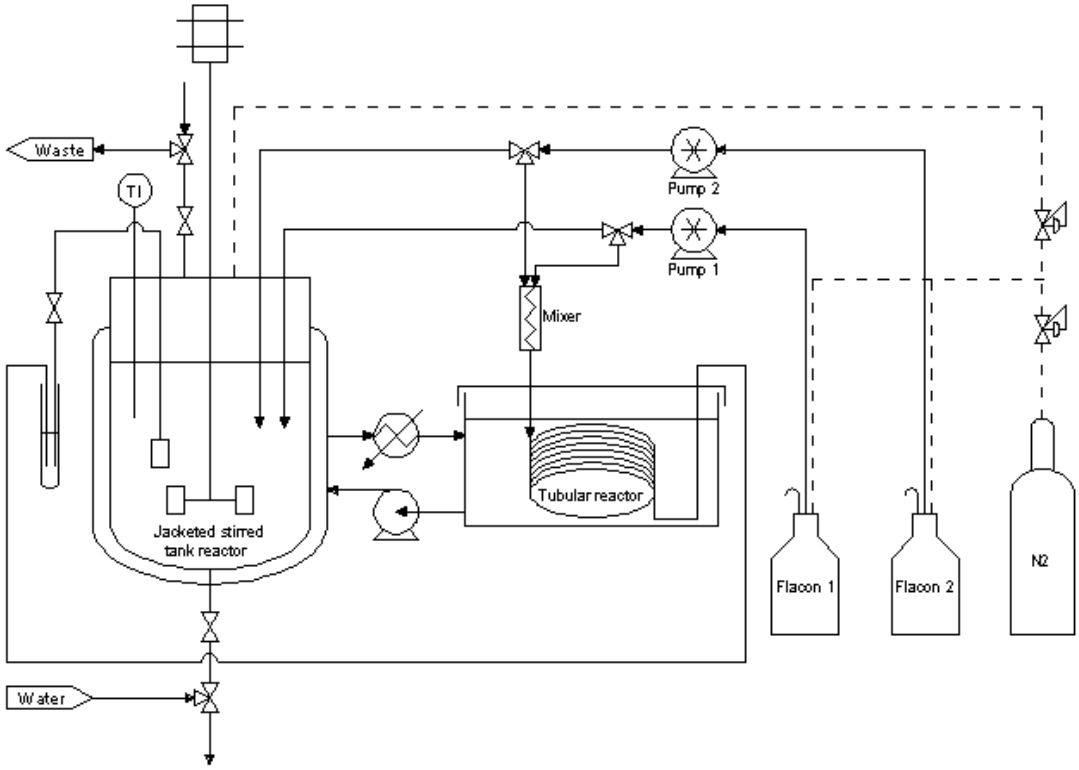


Figure 4.1: Experimental set-up.

The STR is equipped with:

- a Rushton-type impeller centrally placed at 30 mm from the bottom, the impeller diameter and the blade width are 34 mm and 6 mm respectively;
- a double jacket filled with thermofluid (Huber SilOil M40) circulating from a thermostat (Thermo Scientific Phoenix II), in which the TR is immersed to maintain the desired temperature;
- a sampling suite composed of a sampling joint and a nitrogen bubbling system, which can be also used to provide an inert gas or for high-pressure reaction operation;
- two measurement joints, through which various probes can be inserted into the reactor for measuring temperature or pH of the reaction system.

The useful information about the reactors, which will be required by the software, is reported in Table 4.2. Reactant solution is stored in flacon (500 mL) and fed into reactor by programmable pump (Postnova Analytics, 0–40 mL/min), which is usually used for pumping the solutions with low viscosity, such as water, ethanol, etc.

Table 4.2: Characteristics for each reactor.

Reactor configuration	STR	TR
Possible process type	Batch, Semi-Batch, Continuous	Continuous
Internal diameter [cm]	10	0.16
Height/Length [a] [cm]	10	635
Thermal capacity [b] [J/K]	660	– [c]
Heat-transfer surface [m^2]	0.0314	– [c]
Overall heat-transfer coefficient [$W/m^2/K$]	302	– [c]

[a] Height for STR, length for TR.

[b] Reactor thermal capacity is the product of the empty reactor mass and the heat capacity of the reactor manufacturing material.

[c] The temperature is homogeneous along the reactor (Mathieu, 2013), i.e. the experiments in the TR are always under isothermal mode, therefore, there is no need to define its thermal capacity, heat-transfer surface and the overall heat-transfer coefficient.

4.1.3 Experimental protocols

For this case study, experiments are carried out in Batch Stirred-Tank Reactor (BSTR) and Continuous Tubular Reactor (CTR) under isothermal mode.

For the experiments in the BSTR, oil is charged into the reactor and heated to the desired temperature, which is then maintained at the desired temperature. The mechanical stirring (550 rpm) is turned on during oil heating. As soon as ethanol with dissolved sodium hydroxide, which is heated separately to the desired temperature, is added to the reactor, the synthesis is timed. Samples (2 mL) are removed from the reaction mixture at the designed sampling times.

For the experiments in the CTR, the preheated solution 1 (i.e. mixture of ethanol and sodium hydroxide) and solution 2 (i.e. oil) are fed into the reactor through a T-mixer by pumps. Samples (2 mL) are collected at the outlet.

After being immediately quenched by adding 2 mL hydrochloric acid solution (3.7%) and vigorously shaken for 1 minute, the samples form two phases, the light phase is withdrawn and treated by TurboVap LV evaporator in order to remove ethanol from the solution. The composition of the treated light phase is then analyzed (see Section 4.1.4).

4.1.4 Analytical methods

4.1.4.1 Mass percentages of esters in the treated light phase

The Mass Percentages (MPs) of esters in the treated light phase are determined by gas chromatography analysis according to the European Standard EN 14103:2011. The analysis procedure is described as follow:

1. Sample preparation: Accurately weigh approximately 10 mg sample and 10 mg Internal Standard (IS), i.e. ethyl nonadecanoate, in a 2 mL GC vial, then add 1.5 mL toluene to dissolve them.
2. Sample analysis: 1 µL sample solution is analyzed in Agilent 7820A Chromatograph according to the conditions described in Table 4.3.

Table 4.3: Conditions for gas chromatography analysis of esters in the treated light phase.

Condition	Setting
Column	Agilent 19091N-113: 260 °C, 30 m × 320 µm × 0.25 µm
Carrier gas	H_2
Evolution of oven temperature	60 °C (hold for 2 min), then to 200 °C at 10 °C/min, then to 240 °C at 5 °C/min (hold for 7 min)
Detector temperature	250 °C

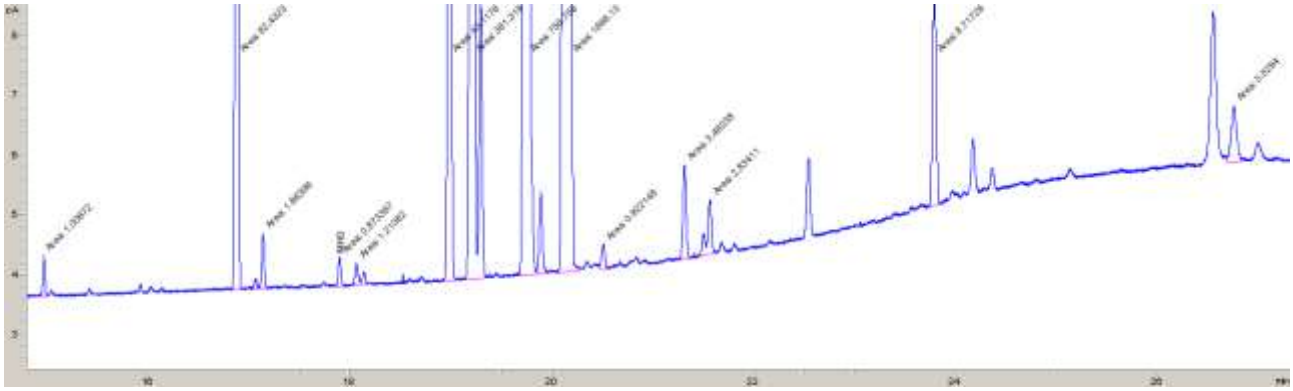


Figure 4.2: Chromatographic analysis of fatty acid ethyl esters in the 3rd sample of the 4th PE.

An example chromatogram of esters is shown in Figure 4.2, it can be observed that in the sample:

1. 13 Fatty Acid Ethyl Esters (FAEEs) (shown in Table 4.4) are identified;
2. among identified mono-unsaturated FAEEs, compared with the amount (proportional to peak area) of ethyl oleate, those of other mono-unsaturated FAEEs can be neglected;
3. among all identified poly-unsaturated FAEEs, compared with the amount of ethyl linoleate, that of ethyl linolenate can be neglected;
4. among all identified saturated FAEEs, compared with the amounts of ethyl palmitate and ethyl stearate, those of other saturated FAEEs can be neglected.

Table 4.4: Identified fatty acid ethyl esters in sample.

N° Peak	Retention time [min]	Peak identification	Fatty acid [a]
1	14.968	ethyl myristate	C14:0
2	16.889	ethyl palmitate	C16:0
3	17.150	ethyl palmitoleate	C16:1
4	17.907	ethyl heptadecanoate	C17:0
5	18.076	ethyl margaroleate	C17:1
6	18.997	ethyl stearate	C18:0
7	19.226	ethyl oleate	C18:1
8	19.768	ethyl linoleate	C18:2
9	20.191	ethyl nonadecanoate (IS)	C19:0
10	20.519	ethyl linolenate	C18:3
11	21.326	ethyl arachidate	C20:0
12	21.578	ethyl eicosonate	C20:1
13	23.798	ethyl behenate	C22:0
14	26.769	ethyl lignocerate	C24:0

[a] Fatty acid is represented in terms of 'C:D', C represents the total number of carbon atoms in fatty acid; D represents the number of double (unsaturated) bonds in it.

On the basis of these observations, reasonable assumptions will be made to simplify oil composition (see Section 4.2.1.3). According to the chromatogram, the Mass Percentage (MP) of each FAEE in the treated light phase can be calculated using Eq. (4.1).

$$MP_{FAEE} = \frac{A_{FAEE}}{A_{IS}} \times \frac{m_{IS}}{m_{sample}} \times 100\% \quad (4.1)$$

where A and m represent respectively the peak area and the added mass.

4.1.4.2 Mass percentages of glycerides in the treated light phase

The MPs of glycerides in the treated light phase are determined by gas chromatography analysis according to the European Standard EN 14105:2011. The analysis procedure is described as follow:

1. Sample preparation: (i) Accurately weigh approximately 10 mg sample in a 2 mL GC vial, then place it in the preparation tray of Agilent 7693 Autosampler; (ii) Add 10 µL solution of ISs, i.e. MonoC19, DiC38 and TriC57, in pyridine (2.5 mg/mL) and 10 µL MSTFA to sample, then mix sample for 5 minutes and wait for 30 minutes; (iii) Add 800 µL heptane to sample, then mix sample for 1 minute. The 1st step is manually performed, whereas the last two steps are performed by Agilent 7693 Autosampler.
2. Sample analysis: 1 µL sample solution is analyzed in Agilent 7890A Chromatograph according to the conditions described in Table 4.5.

Table 4.5: Conditions for gas chromatography analysis of glycerides in the treated light phase.

Condition	Setting
Column	Agilent 123-5711: 400 °C, 15 m × 320 µm × 0.1 µm
Carrier gas	H_2
Evolution of oven temperature	50 °C (hold for 1 min), then to 180 °C at 15 °C/min, then to 230 °C at 7 °C/min, then to 370 °C at 10 °C/min (hold for 15 min)
Detector temperature	380 °C

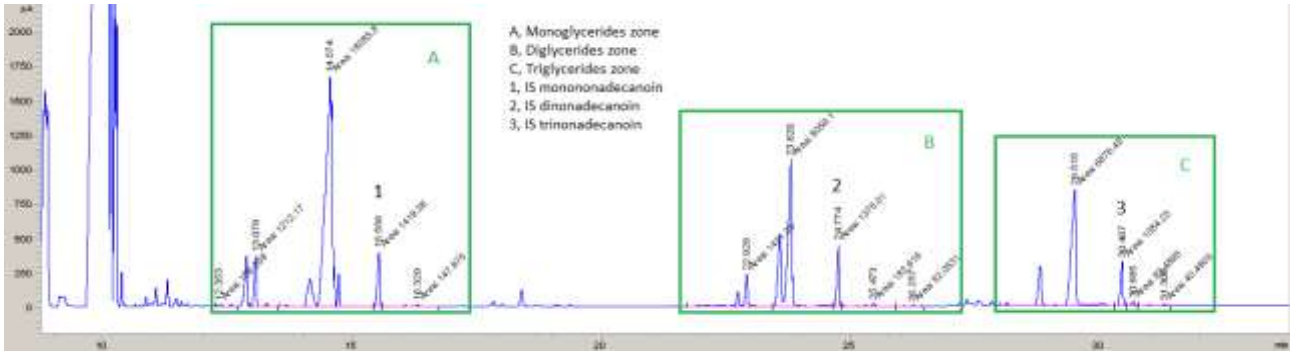


Figure 4.3: Chromatographic analysis of glycerides in the 3rd sample of the 4th PE.

An example chromatogram of glycerides is shown in Figure 4.3, it can be observed that:

1. four peak zones are clearly identified, corresponding respectively to FAEEs (retention time around 9-10 minutes), Mono-Glycerides (MGs), Di-Glycerides (DGs) and Tri-Glycerides (TGs);
2. no glyceride peak can be distinctively identified due to the superposition of peaks.

In the view of the 2nd observation, in this work, glyceride composition of sample is represented in terms of total MGs, DGs and TGs, instead of a specific glyceride. The total MPs of MGs, DGs and TGs in the treated light phase can be calculated using respectively Eqs. (4.2)-(4.4).

$$MP_{MGs} = \frac{A_{MGs}}{A_{MonoC19}} \times \frac{m_{MonoC19}}{m_{sample}} \times 100\% \quad (4.2)$$

$$MP_{DGs} = \frac{A_{DGs}}{A_{DiC38}} \times \frac{m_{DiC38}}{m_{sample}} \times 100\% \quad (4.3)$$

$$MP_{TGs} = \frac{A_{TGs}}{A_{TriC57}} \times \frac{m_{TriC57}}{m_{sample}} \times 100\% \quad (4.4)$$

4.2 Strategy application

This section presents the chronological application of the strategy to the ethanolysis case study (illustrated in Figure 4.4).

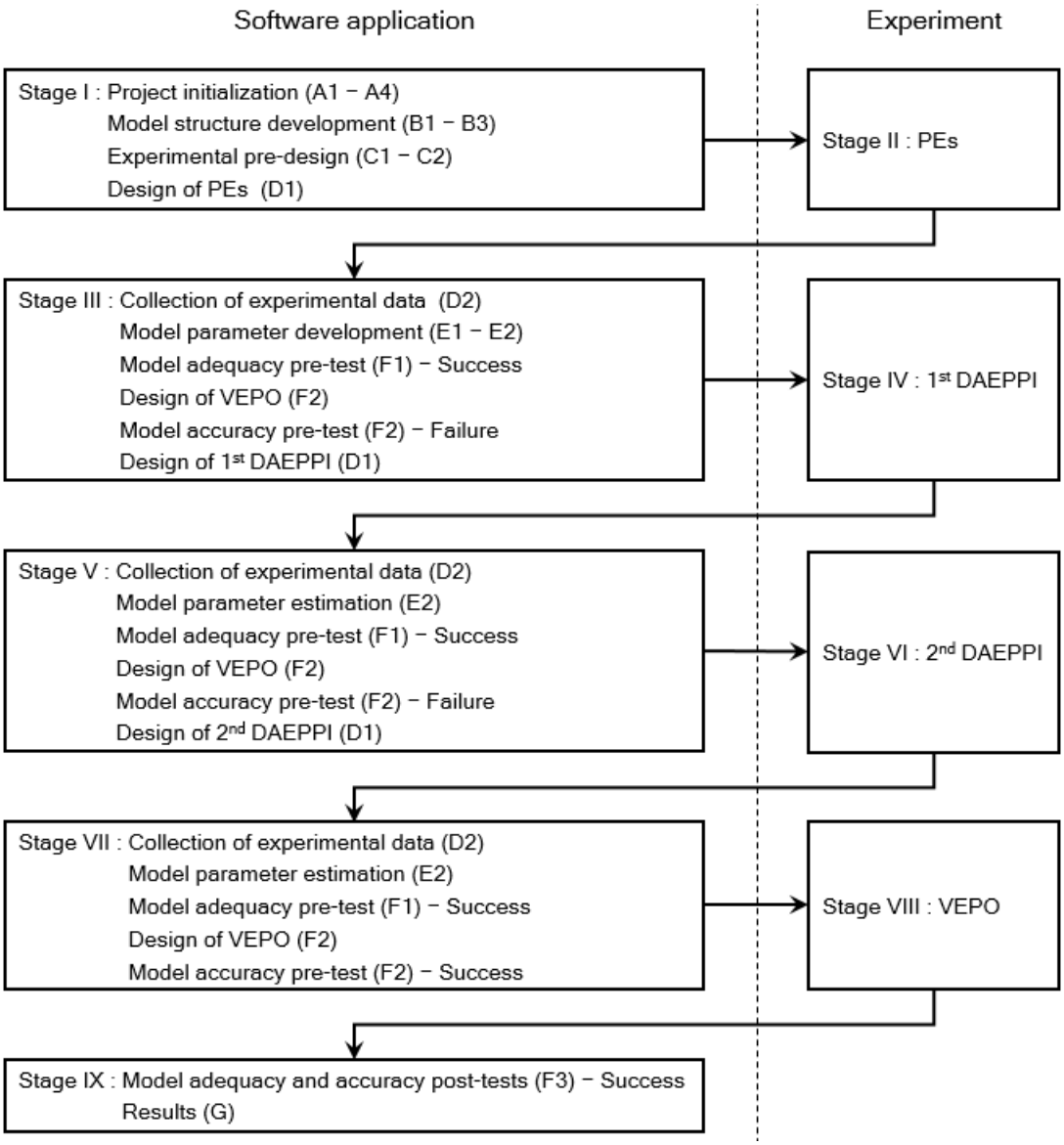


Figure 4.4: Operational sketch of strategy application for the ethanolysis case study.

4.2.1 Stage I: Software application

4.2.1.1 Submodule A1: Name project

Case study 1 is named after NaOH-catalyzed ethanolysis of sunflower oil. A file with this name is generated automatically, and all the simulation results will be stored in it.

4.2.1.2 Submodule A2: Characterize reactors

For this case study, all experiments are carried out under isothermal mode, therefore, the characteristics to be specified for the considered reactors are only their dimensions (shown in Table 4.2).

For the BSTR, in order to avoid that the user defines unreasonably the operating space (see Section 4.2.1.8), the dimensions are used to evaluate whether the initial load volume is greater than the maximum allowable volume of the reactor. Meanwhile, for the CTR, given the total flow rate, the dimensions are used to calculate the space time.

4.2.1.3 Submodule A3: Define species

Vegetable oil is a complex mixture (see Section 3.2.2.4), simplification of its composition appears necessary. For that, sunflower oil is assumed to consist only of the TGs where the three involved fatty acids are identical. Ethanolysis of such a TG results in one FAEE. In this work, the fact that 13 FAEEs are identified in the sample (see Table 4.4) means that sunflower oil is composed of 13 TGs, including:

- mono-unsaturated TGs: tri-palmitolein, tri-margarolein, tri-olein and tri-eicosonin;
- poly-unsaturated TGs: tri-linolein and tri-linolenin;
- saturated TGs: tri-myristin, tri-palmitin, tri-heptadecanoin, tri-stearin, tri-arachidin, tri-behenin and tri-lignocerin.

On the basis of the above assumption and the previous analysis result about chromatogram of esters (see Section 4.1.4.1), the following assumptions are further made:

1. all mono-unsaturated TGs are represented by tri-olein;
2. all poly-unsaturated TGs are represented by tri-linolein;
3. all saturated TGs are represented by tri-palmitin and tri-stearin, which are further represented by a pseudo-species. The pseudo-species is named after tri-palmitin, its molar mass and density are taken as the average values of those of tri-palmitin and tri-stearin.

Thus, sunflower oil is composed of only three TGs, i.e. Tri-Olein (TO), Tri-Linolein (TL) and Tri-Palmitin (TP), all possible species in the modeled reaction system are shown in Table 4.6. In the background, the operating conditions in general form (shown in Table 4.12) and the time-dependent state variables (see Figure 4.7) are automatically generated by the software.

Table 4.6: Information about all possible species in the reaction system for the case study 1.

Species	Symbol	Molar mass [g·mol^{-1}]	Density [kg·m^{-3}]	Species type	In phase I [a]?	In phase II [b]?
Ethanol	E	46.0682	789	Reactant	Yes	No
Sodium hydroxide	NaOH	39.9971	2130	Catalyst	Yes	No
Sodium ethoxide	NaOEt	68.0503	868	Product	Yes	No
Water	W	18.0153	998	Product	Yes	No
Tri-Olein	TO	885.4279	908	Reactant	Yes	Yes
Tri-Linolein	TL	879.3805	925	Reactant	Yes	Yes
Tri-Palmitin	TP	849.3958	868.6	Reactant	Yes	Yes
Di-Olein	DO	620.9831	978	Product	Yes	Yes
Di-Linolein	DL	616.9515	946	Product	Yes	Yes
Di-Palmitin	DP	596.9617	925	Product	Yes	Yes
Mono-Olein	MO	356.5383	1000	Product	Yes	Yes
Mono-Linolein	ML	354.5225	1000	Product	Yes	Yes
Mono-Palmitin	MP	344.5276	984.5	Product	Yes	Yes
Ethyl Oleate	EO	310.5130	870	Product	Yes	Yes
Ethyl Linoleate	EL	308.4972	876	Product	Yes	Yes
Ethyl Palmitate	EP	298.5023	853.5	Product	Yes	Yes
Sodium Oleate	SO	304.4419	870	Product	Yes	No
Sodium Linoleate	SL	302.4261	876	Product	Yes	No
Sodium Palmitate	SP	292.4312	853.5	Product	Yes	No
Glycerol	G	92.0935	1260	Product	Yes	No

[a] Phase I is the polar phase.

[b] Phase II is the non-polar phase.

4.2.1.4 Submodule A4: Define performance criteria

Table 4.7 specifies the performance criteria used for this case study. The objective values of the conversion of the considered TGs as well as the total selectivity of FAEEs comprising EO, EL and EP, are defined as 100% and 75% respectively, since the synthesis can generate three quarters of desired products (FAEEs) and one quarter of by-product (G) when the reactants (TGs) and the intermediates (DGs and MGs) are completely consumed. Figure 4.5 illustrates how to define the 4th performance criterion for the case study 1.

Table 4.7: Performance criteria for the case study 1.

Performance	Symbol	Unit	Objective value	Veto value
Conversion of TO	X_{TO}	%	100	0
Conversion of TL	X_{TL}	%	100	0
Conversion of TP	X_{TP}	%	100	0
Total selectivity of FAEEs	Y_{FAEEs}	%	75	0

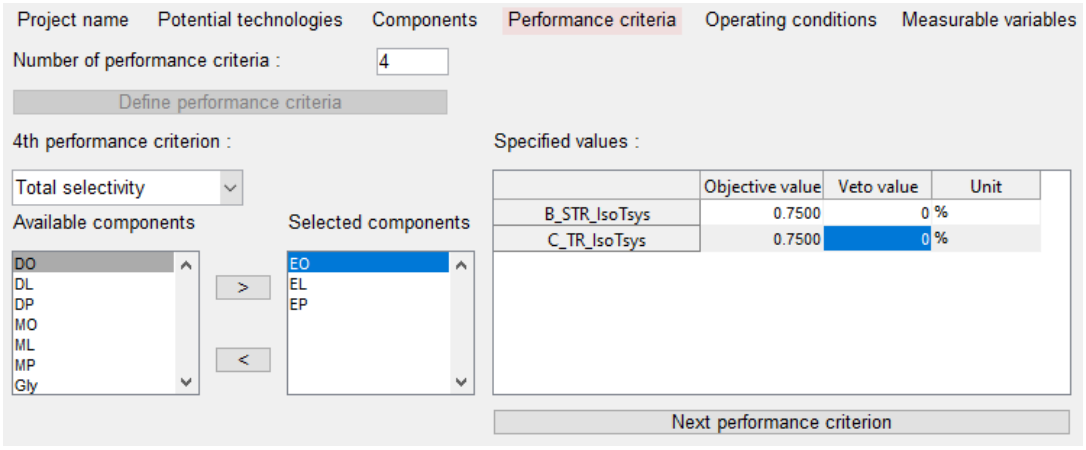


Figure 4.5: Interface for the definition of the 4th performance criterion.

4.2.1.5 Submodule B1: Develop reaction supernetwork

On the basis of the reaction supernetwork for the homogeneous alkali-catalyzed alcoholysis of vegetable oil (see Table 3.4), in which only one TG is considered, the reaction supernetwork for the NaOH-catalyzed ethanolysis of sunflower oil, in which three TGs have to be considered (see Section 4.2.1.3), is obtained by analogy (see Table 4.8).

Table 4.8: Reaction supernetwork for the case study 1.

N°	Reaction	N°	Reaction	N°	Reaction
1	$E(I) + NaOH(I) \rightarrow NaOEt(I) + W(I)$				
2	$NaOEt(I) + W(I) \rightarrow E(I) + NaOH(I)$				
3	$TO(I) + E(I) \xrightarrow{Cat.} EO(I) + DO(I)$	17	$TL(I) + E(I) \xrightarrow{Cat.} EL(I) + DL(I)$	31	$TP(I) + E(I) \xrightarrow{Cat.} EP(I) + DP(I)$
4	$DO(I) + E(I) \xrightarrow{Cat.} EO(I) + MO(I)$	18	$DL(I) + E(I) \xrightarrow{Cat.} EL(I) + ML(I)$	32	$DP(I) + E(I) \xrightarrow{Cat.} EP(I) + MP(I)$
5	$MO(I) + E(I) \xrightarrow{Cat.} EO(I) + G(I)$	19	$ML(I) + E(I) \xrightarrow{Cat.} EL(I) + G(I)$	33	$MP(I) + E(I) \xrightarrow{Cat.} EP(I) + G(I)$
6	$EO(I) + DO(I) \xrightarrow{Cat.} TO(I) + E(I)$	20	$EL(I) + DL(I) \xrightarrow{Cat.} TL(I) + E(I)$	34	$EP(I) + DP(I) \xrightarrow{Cat.} TP(I) + E(I)$
7	$EO(I) + MO(I) \xrightarrow{Cat.} DO(I) + E(I)$	21	$EL(I) + MO(I) \xrightarrow{Cat.} DL(I) + E(I)$	35	$EP(I) + MP(I) \xrightarrow{Cat.} DP(I) + E(I)$
8	$EO(I) + G(I) \xrightarrow{Cat.} MO(I) + E(I)$	22	$EL(I) + G(I) \xrightarrow{Cat.} ML(I) + E(I)$	36	$EP(I) + G(I) \xrightarrow{Cat.} MP(I) + E(I)$
9	$TO(I) + NaOH(I) \rightarrow SO(I) + DO(I)$	23	$TL(I) + NaOH(I) \rightarrow SL(I) + DL(I)$	37	$TP(I) + NaOH(I) \rightarrow SP(I) + DP(I)$
10	$DO(I) + NaOH(I) \rightarrow SO(I) + MO(I)$	24	$DL(I) + NaOH(I) \rightarrow SL(I) + ML(I)$	38	$DP(I) + NaOH(I) \rightarrow SP(I) + MP(I)$
11	$MO(I) + NaOH(I) \rightarrow SO(I) + G(I)$	25	$ML(I) + NaOH(I) \rightarrow SL(I) + G(I)$	39	$MP(I) + NaOH(I) \rightarrow SP(I) + G(I)$
12	$EO(I) + NaOH(I) \rightarrow SO(I) + E(I)$	26	$EL(I) + NaOH(I) \rightarrow SL(I) + E(I)$	40	$EP(I) + NaOH(I) \rightarrow SP(I) + E(I)$
13	$TO(I) \rightarrow TO(II)$	27	$TL(I) \rightarrow TL(II)$	41	$TP(I) \rightarrow TP(II)$
14	$DO(I) \rightarrow DO(II)$	28	$DL(I) \rightarrow DL(II)$	42	$DP(I) \rightarrow DP(II)$
15	$MO(I) \rightarrow MO(II)$	29	$ML(I) \rightarrow ML(II)$	43	$MP(I) \rightarrow MP(II)$
16	$EO(I) \rightarrow EO(II)$	30	$EL(I) \rightarrow EL(II)$	44	$EP(I) \rightarrow EP(II)$

It can be seen that in this reaction supernetwork:

1. neutralization reactions of Free Fatty Acids (FFAs) and mass transfers of FFAs between the polar phase (I) and the nonpolar phase (II) are not included due to the fact that FFAs are not included in sunflower oil (see Section 4.2.1.3);
2. mass transfers of ethanol, glycerol and water between the polar and nonpolar phases are not included, since they are defined to be present only in the polar phase (see Table 4.6).

4.2.1.6 Submodule B2: Propose candidate reaction networks

For this case study, two reaction networks under test (illustrated in Table 4.9) are initially proposed from the reaction supernetwork developed above. Then, the same subsequent procedure (from the submodule B3 to the submodule F1, see Figure 4.4) for both the RN1 and the RN2 will be carried out below. Concrete results obtained from the submodules, such as model parameter development and model adequacy pre-test, vary with the reaction network. For concision of this chapter, this procedure will be solely illustrated for the RN1, i.e. the best reaction network finally identified (see Section 4.2.3.4).

Table 4.9: Reaction networks under test for the case study 1.

N°	Reaction network	Reactions taken into consideration
1	Reduced pseudo-homogeneous reaction network (RN1)	3-8, 17-22, 31-36
2	Biphasic reaction network (RN2)	3-8, 13-22, 27-36, 41-44

4.2.1.7 Submodule B3: Select model parameters to be estimated

For the RN1, all involved chemical reactions are considered as elementary reactions, therefore, the default values (i.e. stoichiometric coefficients) are assigned to the kinetic orders (see Section 2.3.2). $\log k_{j,T_{max}}$ and $\log k_{j,T_{min}}$ for each involved chemical reaction are considered as the model parameters to be actually estimated for the model structures developed based on the RN1 (see Section 2.3.1). The minimal and maximal possible reaction temperatures correspond respectively to 303.15K and 351.15K (see Table 4.12).

4.2.1.8 Submodule C1: Specify constraints on operating conditions

The same operating conditions can be represented by three different manners: practical operating conditions, practical operating variables and operating conditions in general form. In this section, after specifying the constraints on the practical operating conditions for each considered reactor, those on the corresponding practical operating variables and operating conditions in general form are derived.

4.2.1.8.1 Practical operating conditions

Practical Operating Conditions (POCs) can be classified into two categories: Variable POCs (VPOCs) and Constant POCs (CPOCs). Given the experimental protocols, POCs used for characterizing the experiments in each considered reactor are identified. Subsequently, one has to:

1. evaluate qualitatively the identifiability of the experiment in each considered reactor. The experiment within a random practical operating space (i.e. bounds on the VPOCs) of the considered reactor, represented as $E(\xi)$ with $\xi \in \Xi$, is identifiable only if the proposition (4.5) is valid;

$$\forall(\xi,\xi') \in \Xi^2 \quad E(\xi) = E(\xi') \Rightarrow \xi = \xi' \quad (4.5)$$

2. specify the practical operating space of each considered reactor.

For this case study, it can be seen that given VPOCs of the CTR $\{m_{NaOH,1}, T_1, F_1, T_2, F_2\}$ (see Table 4.10), the experiment is identifiable, since Eq. (4.6) is always valid; whereas given those of the BSTR $\{m_{E,1}, m_{NaOH,1}, m_{oil,2}, T^{int}, \tau\}$, the experiment is not identifiable, in other words, there are infinite sets of VPOCs allowing the experiments in the BSTR to provide the same experimental responses, which is illustrated by Eq. (4.7).

$$(\xi_{CTR},\xi'_{CTR}) \in \Xi^2_{CTR} \quad E(\xi_{CTR}) = E(\xi'_{CTR}) \Rightarrow \begin{cases} m_{NaOH,1} = m'_{NaOH,1} \\ T_1 = T'_1 \\ F_1 = F'_1 \\ T_2 = T'_2 \\ F_2 = F'_2 \end{cases} \quad (4.6)$$

$$\forall(\xi_{BSTR},\xi'_{BSTR}) \in \Xi^2_{BSTR} \quad E(\xi_{BSTR}) = E(\xi'_{BSTR}) \Rightarrow \begin{cases} \frac{m_{E,1}}{m'_{E,1}} = \frac{m_{NaOH,1}}{m'_{NaOH,1}} = \frac{m_{oil,2}}{m'_{oil,2}} \\ T_{int} = T'_{int} \\ \tau = \tau' \end{cases} \quad (4.7)$$

In order that any experiment in the BSTR is identifiable, the VPOC, i.e. mass of oil in solution 2 (initially loaded into the reactor), is removed from the set of VPOCs, and considered as CPOC, which allows that the proposition (4.5) is valid.

The bounds on the practical operating conditions for each considered reactor are specified and shown in Table 4.10. The experiments extracted within the practical operating space of each considered reactor are identifiable. The flow rate of oil for the experiments in the CTR is defined as constant due to the practical limit of the pump. As mentioned in Section 4.1.2, the pump used in this work usually applies to pump the solutions with low viscosity, such as water, ethanol, etc. Reliable and reproducible flow rates are promised for the whole pumping range; whereas for the solutions

with high viscosity, such as oil, such flow rates are promised only for several discrete points, including the value actually defined, 4 mL/min.

Table 4.10: Bounds on practical operating conditions for each considered reactor.

Practical operating condition	Symbol	Unit	BSTR	CTR
Mass of ethanol in solution 1	$m_{E,1}$	g	31.5 – 252	200
Mass of NaOH in solution 1	$m_{NaOH,1}$	g	0.2 – 3	0.0380 – 0.7605
Mass of oil in solution 2	$m_{oil,2}$	g	200	200
Initial reaction temperature	T^{int}	°C	30 – 70	
Temperature of feed 1	T_1	°C		30 – 70
Flow rate of feed 1	F_1	mL/min		2 – 20
Temperature of feed 2	T_2	°C		30 – 70
Flow rate of feed 2	F_2	mL/min		4
Experiment duration	τ	s	30 – 8000	

4.2.1.8.2 Practical operating variables

For this case study, the practical operating variables to be studied are: (i) ethanol-to-oil molar ratio, (ii) NaOH amount, (iii) reaction temperature, (iv) reaction duration, since other operating variables influencing the reaction system, such as the catalyst type (NaOH or NaOEt, generated from NaOH) and the mixing intensity (e.g. for the experiments in the BSTR, mixing speed is 550 rpm, since according to the ethanolysis researches of Marjanović et al. (2010), Likozar and Levec (2014b) and Reyero et al. (2015), under such mixing speed, the reaction system is not affected by mass transfer limitations, that the best reaction network identified is the RN1 confirms the fact that the reaction system is only limited by the kinetics; for those in the CTR, T-mixer is used), have been given, while designing the experimental protocols.

Given the composition of oil (see Section 4.1.1), the properties of reactants and catalyst (see Table 4.6), the bounds on practical operating variables (shown in Table 4.11) are derived from those on the practical operating conditions (see Table 4.10). Specifically, for the experiments in the BSTR, with fixed initial mass of oil, ethanol-to-oil molar ratio and NaOH amount are derived from initial masses of ethanol and NaOH; for the experiments in the CTR, with fixed flow rate of oil, those are derived from flow rate of solution 1 (i.e. mixture of ethanol and NaOH) and concentration of NaOH in solution 1.

Bounds on practical operating variables are logical and reasonable in terms of implementation practicability: for instance, experiments in the BSTR can possess longer reaction duration, whereas those in the CTR can possess higher ethanol-to-oil molar ratio.

Table 4.11: Bounds on practical operating variables for each considered reactor.

Practical operating variable	BSTR	CTR
Ethanol-to-oil molar ratio [-]	3 – 24	8.23 – 82.26
NaOH amount [g per 100 g of oil]	0.1 – 1.5	0.0082 – 1.6411
Reaction temperature [°C]	30 – 70	30 – 70
Reaction duration [s]	30 – 8000	31.90 – 127.91

4.2.1.8.3 Operating conditions in general form

By analogy, the corresponding operating conditions in general form, that actually constitute the inputs into the software, are derived and shown in Table 4.12. The number of feeds included in the operating conditions in general form is 2, which confirms that the number of virtual feeds used for simulation is equal to that of all possible phases in the reaction system (see Section 3.1.1.3). In order to simplify the specification of the operating conditions in general form and the modelling, the feed II, i.e. the non-polar feed, is assumed to be non-existent. Hence, while simulation with the model structures constructed based on the RN2, at the initial state, the loaded materials are automatically partitioned into the polar and non-polar phases according to the associated distribution coefficients.

Table 4.12: Bounds on operating conditions in general form for each considered reactor.

Operating condition in general form	Symbol	Unit	BSTR	CTR
Experiment duration	τ	s	30 - 8000	
Initial reaction temperature	T^{int}	K	303.15 - 343.15	
Initial load of E	n_E^{int}	mol	0.6837 – 5.4696	
Initial load of NaOH	n_{NaOH}^{int}	mol	0.0050 - 0.0750	
Initial load of TO	n_{TO}^{int}	mol	0.0610	
Initial load of TL	n_{TL}^{int}	mol	0.1410	
Initial load of TP	n_{TP}^{int}	mol	0.0259	
Temperature of feed I	T_I^{in}	K		303.15 - 343.15
Temperature of feed II	T_{II}^{in}	K		303.15 - 343.15
Flow rate of feed I	F_I^{in}	$m^3 \cdot s^{-1}$		10^{-7} – 4×10^{-7}
Flow rate of feed II	F_{II}^{in}	$m^3 \cdot s^{-1}$		0
Concentration of E in feed I	$c_{E,I}^{in}$	$mol \cdot m^{-3}$		5708.9 – 14272.3
Concentration of NaOH in feed I	$c_{NaOH,I}^{in}$	$mol \cdot m^{-3}$		1.25 – 62.5
Concentration of TO in feed I	$c_{TO,I}^{in}$	$mol \cdot m^{-3}$		46.4 – 185.8
Concentration of TL in feed I	$c_{TL,I}^{in}$	$mol \cdot m^{-3}$		107.4 – 429.5
Concentration of TP in feed I	$c_{TP,I}^{in}$	$mol \cdot m^{-3}$		19.7 – 78.9
Concentration of TO in feed II	$c_{TO,II}^{in}$	$mol \cdot m^{-3}$		0
Concentration of TL in feed II	$c_{TL,II}^{in}$	$mol \cdot m^{-3}$		0
Concentration of TP in feed II	$c_{TP,II}^{in}$	$mol \cdot m^{-3}$		0

For consistency, the corresponding operating conditions in general form for the CTR are subject not only to their bounds but also to four equality constraints based on the oil composition constraints (Eqs. (4.8)-(4.10)) and the volume balance (Eq. (4.11), ethanol-to-oil volumic ratio in the feed I is equal to the flow-rate ratio of actual feeds) by neglecting the contribution of NaOH to liquid volume, which is illustrated by Figure 4.6.

$$\frac{c_{TO,I}^{in}}{c_{TO,I}^{in} + c_{TL,I}^{in} + c_{TP,I}^{in}} = \chi_{TO} \quad (4.8)$$

$$\frac{c_{TL,I}^{in}}{c_{TO,I}^{in} + c_{TL,I}^{in} + c_{TP,I}^{in}} = \chi_{TL} \quad (4.9)$$

$$\frac{c_{TP,I}^{in}}{c_{TO,I}^{in} + c_{TL,I}^{in} + c_{TP,I}^{in}} = \chi_{TP} \quad (4.10)$$

$$\frac{v_{TO}c_{TO,I}^{in} + v_{TL}c_{TL,I}^{in} + v_{TP}c_{TP,I}^{in}}{v_E c_{E,I}^{in}} = \frac{F_2}{F_I^{in} - F_2} \quad (4.11)$$

where χ_{TO}, χ_{TL} and χ_{TP} are the molar fractions of TO, TL and TP in oil, F_2 is the practical flow rate of solution 2 (i.e. oil), equal to 4 mL/min (see Table 4.10).

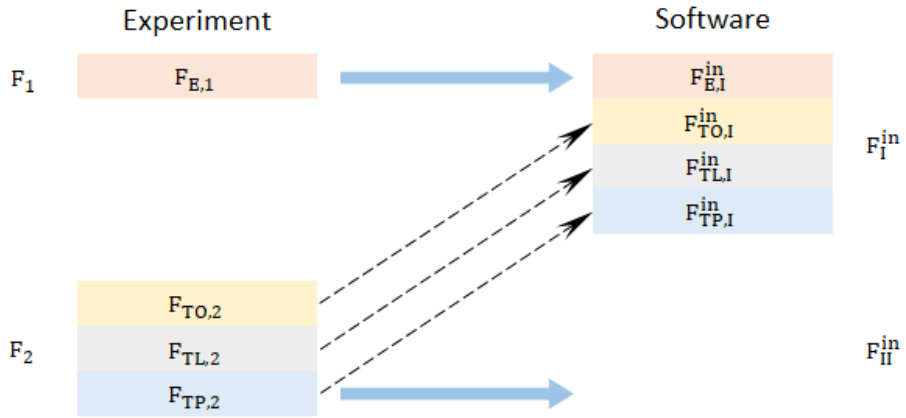


Figure 4.6: Illustration of the relationship between practical feeds and virtual feeds.

4.2.1.9 Submodule C2: Define measurable variables

The variables to be actually measured are the concentrations of EO, EL and EP as well as the total concentrations of MGs, DGs and TGs. Figure 4.7 illustrates how to define the measurable variables for this case study. The definition procedure involves four operations, specifically:

- operation 1 (in the red dotted box): select the conventional measurable variables from a specific suggestion list containing the time-dependent state variables (automatically generated after defining species in the submodule A3), and validate;

- operation 2 (in the blue dotted box): indicate whether there are Unconventional Measurable Variables (UMVs) that need to be defined, and validate;
- operation 3 (in the purple dotted box): indicate the number of UMVs, and validate;
- operation 4 (in the yellow dotted box): name UMVs, define the relationships between UMVs and the time-dependent state variables, for instance, the total concentration of MGs is equal to the sum of the concentrations of MO, ML and MP, and then validate.

Given these measurable variables, the parameters involved in the model structure constructed based on the RN1 are identifiable (Likozar and Levec, 2014b).

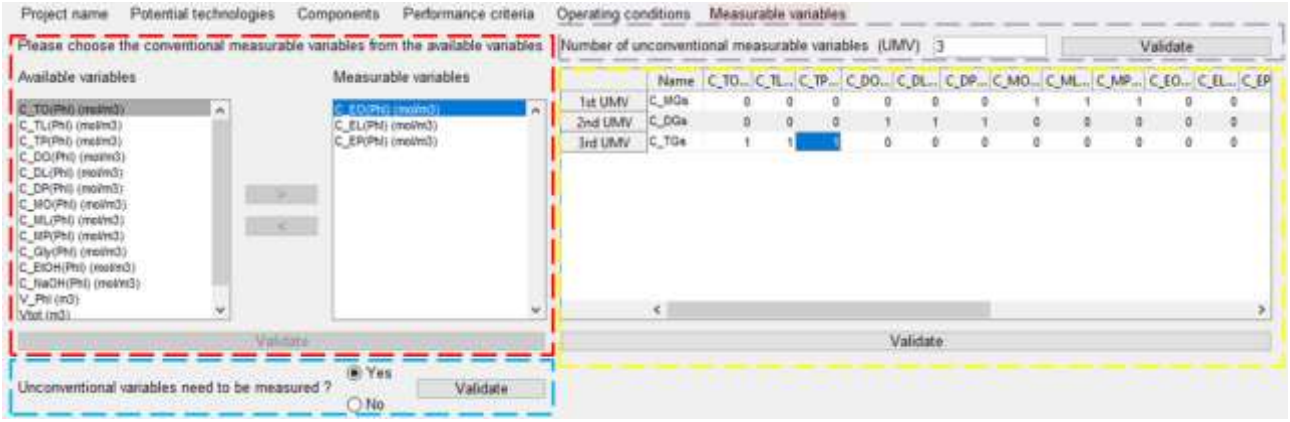


Figure 4.7: Interface for the selection of the measurable variables for the case study 1.

4.2.1.10 Submodule D1: Design data acquisition experiments

The design of the initial data acquisition experiments, i.e. the Preliminary Experiments (PEs), is performed by the user and composed of four steps:

1. Identify the experimental factors under test, i.e. reactor, ethanol-to-oil molar ratio, NaOH amount, reaction temperature;
2. Define the low and high levels for each considered experimental factor (see Table 4.13) within its associated bounds observed in the literature (Richard et al., 2013; Likozar and Levec, 2014b);

Table 4.13: Definition of the low and high levels for each considered factor.

Factor	Low level (–)	High level (+)
Reactor	BSTR	CTR
Ethanol-to-oil molar ratio [-]	6	12
NaOH amount [g per 100 g of oil]	0.5	1
Reaction temperature [°C]	45	60

3. Define a reference experiment and four associated derivative experiments (see Table 4.14) according to the following design criteria: (i) each experimental factor is assigned to only one of its associated high-level and low-level values; (ii) compared with the reference experiment, one experimental factor is changed in each derivative experiment;
4. In order to obtain more experimental information used for parameter estimation, several samples can be taken in one run. For instance, for the experiments in the BSTR, given the same initial loadings and reaction temperature, samples can be taken at different experiment durations, i.e. sampling times; for the experiments in the CTR, given the same feeds, samples can be taken at different reaction temperatures.

Table 4.14: Fractional factorial experimental design for case study 1.

Run	Reactor	Ethanol-to-oil molar ratio	NaOH amount	Reaction temperature
1 [a]	–	–	–	+
2	–	–	–	–
3	–	–	+	+
4	–	+	–	+
5	+	–	–	+

[a] Run 1 is the reference experiment.

Design result of PEs is shown in Table 4.15. It can be seen that, for the PEs in the BSTR, samples are taken at incremental time intervals, not uniform time intervals: since the reaction progress is rapid, necessary samples should be taken at the beginning.

Table 4.15: Designed operating variables of the PEs for case study 1.

Run	Reactor	Ethanol-to-oil molar ratio [-]	NaOH amount [g per 100 g of oil]	Reaction temperature [°C]	Sampling times [min]
1	BSTR	6	0.50	60	0.5, 1, 2, 4, 8, 16, 32, 64
2	BSTR	6	0.50	45	0.5, 1, 2, 4, 8, 16, 32, 64
3	BSTR	6	1.00	60	0.5, 1, 2, 4, 8, 16, 32, 64
4	BSTR	12	0.50	60	0.5, 1, 2, 4, 8, 16, 32, 64
5	CTR	6	0.50	60, 45	0.5 [a]

[a] Sample is taken at the outlet of the CTR, i.e. the sampling time is equal to the residence time, which depends on the total flow rate, given the dimensions of the CTR. The designed flow rates of feed 1 and feed 2 are respectively 19 mL/min and 7 mL/min.

4.2.2 Stage II: PEs

The PEs are performed according to the practical operating conditions derived from the designed operating variables obtained at the submodule D1 (see Section 4.2.1.10).

Figure 4.8 shows the evolution of experimental responses as a function of reaction time for the 1st PE. It can be seen that the chemical equilibrium is not reached at the last sampling time. Therefore,

for the subsequent PEs carried out in the BSTR, one more sample is required. Table 4.16 shows the practical operating variables of PEs.

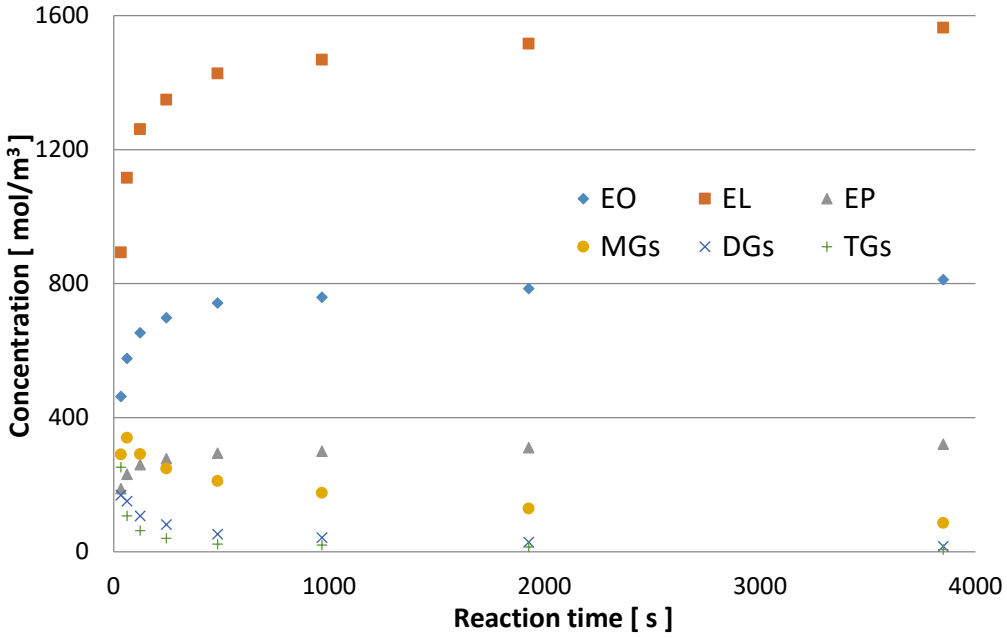


Figure 4.8: Evolution of the experimental responses for the 1st PE carried out at 60°C, NaOH concentration of 0.50 wt.% (per oil weight), ethanol-to-oil molar ratio of 6:1, using the BSTR

Table 4.16: Practical operating variables of the PEs for the case study 1.

Run	Reactor	Ethanol-to-oil molar ratio [-]	NaOH amount [g per 100 g of oil]	Reaction temperature [°C]	Sampling times [min]
1	BSTR	6	0.50	60	0.5, 1, 2, 4, 8, 16, 32, 64
2	BSTR	6	0.50	45	0.5, 1, 2, 4, 8, 16, 32, 64, 128 [a]
3	BSTR	6	1.00	60	0.5, 1, 2, 4, 8, 16, 32, 64, 128 [a]
4	BSTR	12	0.50	60	0.5, 1, 2, 4, 8, 16, 32, 64, 128 [a]
5	CTR	28 [b]	2.34 [b]	60, 45	1.18 [b]

[a] Additional sample;

[b] The practical ethanol-to-oil molar ratio, NaOH amount and sampling time have to be calculated using the practical flow rates of feed1 and feed 2 (4 and 6.85 mL/min respectively), which are different from the corresponding designed flow rates due to the limit of the pump.

4.2.3 Stage III: Software application

4.2.3.1 Submodule D2: Input experimental data

The operating conditions in general form, derived from the practical operating variables of the PEs (see Table 4.16), and the corresponding experimental responses (see Table A2.1), derived from the mass percentages of EO, EL, EP, and the total mass percentages of MGs, DGs, TGs, are input into the software.

The measurement error of each measurable variable, which will be used as the threshold to evaluate model prediction accuracy at the submodule F2 (see Sections 4.2.3.5, 4.2.5.4 and 4.2.7.4) and the submodule F3 (see Section 4.2.9.1), has to be defined here. For this case study, the threshold for the measurable variable of type concentration is set as 5 mol/m^3.

4.2.3.2 Submodule E1: Specify constraints on model parameters

In order that the estimates of A_j and Ea_j, derived from $\log k_{j,T_{max}}$ and $\log k_{j,T_{min}}$ according to Eqs. (2.16)-(2.17), are reasonable, neither too big or too small, referring to the study of Likozar and Levec, (2014b), the model parameters to be estimated are subjected to the bounds specified between -11 and -5 (covering all the parameter estimates in the study of Likozar and Levec, (2014b)), as well as the linear inequality constraints (Eqs. (4.12)-(4.13)).

$$\log k_{j,T_{max}} - \log k_{j,T_{min}} > 0.4 \quad (4.12)$$

$$\log k_{j,T_{max}} - \log k_{j,T_{min}} < 1.4 \quad (4.13)$$

4.2.3.3 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs. The result is shown in the 1st column of Table 4.17, in which the results of parameter estimations performed using different experimental data updated after the 1st Data Acquisition Experiments for Parameter Precision Improvement (DAEPPI) and the 2nd DAEPPI are also included and will be discussed later (see Sections 4.2.5.2 and 4.2.7.2).

Model parameters are obviously inaccurate, since the confidence intervals for some parameters (with gray shading) are very wide. Among these parameters, the confidence intervals for $\log(k_{3,Tmin})$, $\log(k_{3,Tmax})$, $\log(k_{31,Tmin})$, $\log(k_{31,Tmax})$, $\log(k_{6,Tmin})$, $\log(k_{6,Tmax})$, $\log(k_{34,Tmin})$, $\log(k_{24,Tmax})$, namely, the parameters to characterize respectively the reactions N°3, 31, 6, 34, dominating at the beginning of the synthesis, are especially wide. This confirms that the PEs do not provide enough information at the beginning of the synthesis.

Model adequacy and accuracy are directly reflected by:

- the parameter estimation criterion, which will be used for pre-evaluating model adequacy at the submodule F1 (see Sections 4.2.3.4, 4.2.5.3 and 4.2.7.3);
- the average variance of model parameters derived from the parameter variance-covariance matrix (Eq. (1.7)), which will be used to calculate the model prediction variance-covariance matrix (Eq. (1.15)) and further the confidence intervals associated to the model predictions, used

for pre-evaluating model accuracy at the submodule F2 (see Sections 4.2.3.5, 4.2.5.4 and 4.2.7.4).

Table 4.17: Evolution of parameter estimations for the case study 1.

N° parameter estimation (Strategy application stage)	1 (Stage III)	2 (Stage V)	3 (Stage VII)
Data used for parameter estimation	PEs	PEs + 1st DAEPPI	PEs + 1st DAEPPI + 2nd DAEPPI
Parameter estimation criterion (Eq. (2.19))	24.3357	24.0603	23.5021
Average variance of model parameters	1.6669	0.8850	0.0262
$\log(k_{3,Tmin})$	-5.9969 ± 15.8021	-6.0267 ± 35.8940	-6.0237 ± 0.0522
$\log(k_{3,Tmax})$	-5.3552 ± 65.5235	-5.3603 ± 60.9098	-5.3612 ± 0.1397
$\log(k_{4,Tmin})$	-7.3507 ± 1.7846	-7.3408 ± 3.2309	-7.3349 ± 0.6286
$\log(k_{4,Tmax})$	-6.2215 ± 3.3151	-6.2314 ± 3.6027	-6.2310 ± 0.7899
$\log(k_{5,Tmin})$	-8.1889 ± 0.1435	-8.1840 ± 0.0735	-8.1841 ± 0.0724
$\log(k_{5,Tmax})$	-7.5314 ± 0.0939	-7.5315 ± 0.0200	-7.5312 ± 0.0436
$\log(k_{17,Tmin})$	-6.9258 ± 3.5347	-6.9216 ± 0.5593	-6.9323 ± 0.0262
$\log(k_{17,Tmax})$	-6.1192 ± 5.7219	-6.0949 ± 0.3254	-6.1095 ± 0.2277
$\log(k_{18,Tmin})$	-7.1161 ± 1.4514	-7.1472 ± 0.4771	-7.1545 ± 0.0667
$\log(k_{18,Tmax})$	-6.6423 ± 2.8467	-6.6458 ± 0.2083	-6.6416 ± 0.1352
$\log(k_{19,Tmin})$	-7.1940 ± 5.2812	-7.1879 ± 0.2459	-7.1871 ± 0.3508
$\log(k_{19,Tmax})$	-6.0631 ± 9.8274	-6.0570 ± 0.2125	-6.0643 ± 0.3777
$\log(k_{31,Tmin})$	-6.2790 ± 284.1310	-6.3043 ± 64.0053	-6.3030 ± 0.0253
$\log(k_{31,Tmax})$	-5.4899 ± 515.8986	-5.5350 ± 70.2819	-5.5289 ± 0.7176
$\log(k_{32,Tmin})$	-7.2657 ± 26.4442	-7.2684 ± 8.3569	-7.2763 ± 0.6396
$\log(k_{32,Tmax})$	-6.3470 ± 46.5832	-6.3314 ± 8.3906	-6.3381 ± 0.9003
$\log(k_{33,Tmin})$	-8.0997 ± 1.4453	-8.1023 ± 0.5528	-8.1066 ± 0.4064
$\log(k_{33,Tmax})$	-6.7961 ± 2.0084	-6.8208 ± 0.8792	-6.8249 ± 0.3120
$\log(k_{6,Tmin})$	-6.5558 ± 13.0749	-6.5906 ± 47.3796	-6.6000 ± 0.1319
$\log(k_{6,Tmax})$	-6.0975 ± 63.4358	-6.1184 ± 72.6217	-6.1196 ± 0.3524
$\log(k_{7,Tmin})$	-7.1936 ± 2.2469	-7.1792 ± 2.9003	-7.1705 ± 0.6594
$\log(k_{7,Tmax})$	-5.8815 ± 2.4924	-5.8888 ± 3.1192	-5.8928 ± 0.9617
$\log(k_{8,Tmin})$	-9.9672 ± 7.3913	-9.9474 ± 0.5112	-9.9499 ± 1.8693
$\log(k_{8,Tmax})$	-8.9398 ± 3.8995	-8.9327 ± 3.1633	-8.9371 ± 2.9321
$\log(k_{20,Tmin})$	-5.8518 ± 5.1718	-5.8823 ± 1.2472	-5.8829 ± 0.1716
$\log(k_{20,Tmax})$	-5.3665 ± 8.9354	-5.3812 ± 0.5149	-5.3867 ± 0.5290
$\log(k_{21,Tmin})$	-8.4806 ± 1.4640	-8.4375 ± 0.1691	-8.4286 ± 0.0791
$\log(k_{21,Tmax})$	-7.6121 ± 2.6692	-7.5846 ± 0.2432	-7.5766 ± 0.2474
$\log(k_{22,Tmin})$	-7.1416 ± 5.3193	-7.1331 ± 0.3532	-7.1359 ± 0.3453
$\log(k_{22,Tmax})$	-6.5939 ± 9.9045	-6.5865 ± 0.3624	-6.5912 ± 0.3685
$\log(k_{34,Tmin})$	-5.9541 ± 255.6997	-5.9635 ± 64.0622	-5.9632 ± 1.0698
$\log(k_{34,Tmax})$	-5.4827 ± 466.5295	-5.4797 ± 76.2883	-5.4758 ± 0.5316
$\log(k_{35,Tmin})$	-6.1097 ± 19.9058	-6.1088 ± 9.7819	-6.1121 ± 0.7501
$\log(k_{35,Tmax})$	-5.6166 ± 37.2714	-5.6134 ± 8.0466	-5.6068 ± 0.3030
$\log(k_{36,Tmin})$	-9.1929 ± 4.8020	-9.1772 ± 1.9314	-9.1845 ± 2.1089
$\log(k_{36,Tmax})$	-8.7842 ± 6.2776	-8.7568 ± 3.8840	-8.7350 ± 2.8705

A new Set of Models Under Test (SMUT) is generated after parameter estimation due to the updated parameters. Therefore, in order to represent precisely the SMUT, its reaction network (used for model structure development) and the experiments (whose data are used for model parameter estimation) have to be mentioned. For instance, the SMUT[RN1,PEs] represents the SMUT whose model structures are constructed based on the reduced pseudo-homogeneous one (RN1) and parameters are the ones estimated from the experimental data of the PEs.

4.2.3.4 Submodule F1: Model adequacy pre-test

The SMUT[RN1,PEs] has been proven qualitatively feasible, since the experimental response points are reasonably scattered around or on the corresponding predicted curves for the BSTR (see Figure 4.9) and the CTR (see Figure 4.10).

The FAEEs (i.e. EO, EL and EP) show a great increase before the chemical equilibrium is reached due to the high production rate with the presence of excess of ethanol.

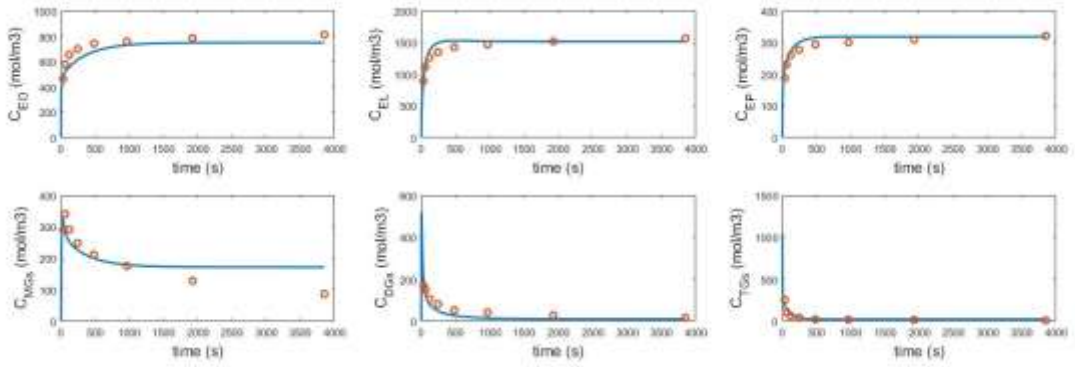


Figure 4.9: Evolution of the various measurable variables for the 1st PE carried out at 60°C, NaOH concentration of 0.50 wt.% (per oil weight), ethanol-to-oil molar ratio of 6:1, using BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).

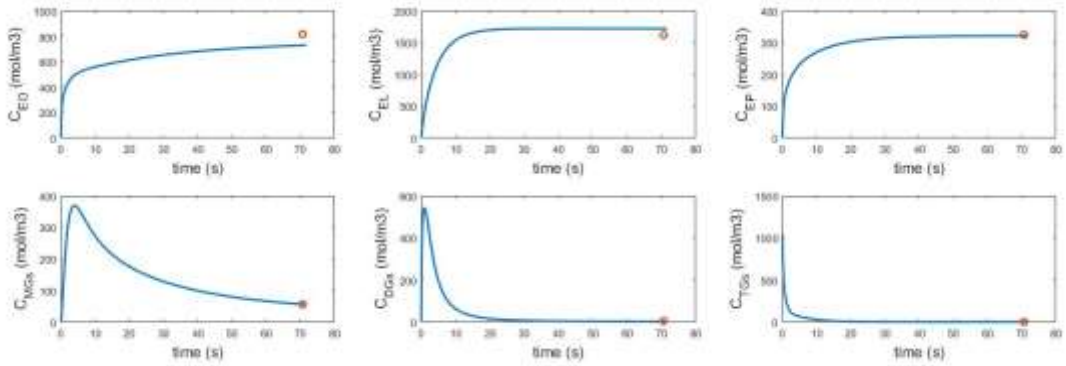


Figure 4.10: Evolution of the various measurable variables for the 1st sample of the 5th PE carried out at 60°C, NaOH concentration of 2.34 wt.% (per oil weight), ethanol-to-oil molar ratio of 28:1, using CTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).

We also observe a peak in the concentration profiles for intermediates (i.e. DG and MG), their production rates are larger than their consumption rates before the peak, and reversely after.

After model adequacy pre-test is implemented for the SMUT[RN1,PEs], the same test for the SMUT[RN2,PEs] is carried out. The SMUT[RN2,PEs] pass also model adequacy pre-test. The most suitable SMUT, the SMUT[RN1,PEs], is identified by comparing directly their parameter estimation criteria: the parameter estimation criterion (Eq. (2.19)) for the SMUT[RN2,PEs] (27.0638) is larger than that of the SMUT[RN1,PEs] (24.3357).

4.2.3.5 Submodule F2: Model accuracy pre-test

Model-based experimental design for the Validation Experiment for Performance Optimization (VEPO) is then performed on the basis of the SMUT[RN1,PEs], i.e. the most suitable SMUT identified at the submodule F1 (see Section 4.2.3.4). The result is shown in the 1st column of Table 4.18, in which the results of model-based experimental design for the VEPO performed based on the SMUTs updated after the 1st DAEPPI and the 2nd DAEPPI are also included and will be discussed later (see Sections 4.2.5.4 and 4.2.7.4).

The optimal reactor (i.e. the BSTR) and the optimal operating conditions (i.e. long reaction duration, near to maximal ethanol-to-oil molar ratio, near to maximal reaction temperature and high NaOH amount), are identified for the VEPO which should enable each performance to reach or approach its objective value.

Then, model accuracy pre-test is carried out to investigate whether the predicted error of each measurable variable for the VEPO is less than its corresponding threshold, i.e. measurement error, which is defined at the submodule D2 (see Section 4.2.3.1).

The SMUT[RN1,PEs] does not pass model accuracy pre-test, since the model prediction errors of C_{EO} and C_{MGs}, respectively 6.0422 mol/m^3 and 5.9789 mol/m^3, are larger than the threshold. Therefore, there are two possible choices for the next step: model structure development (submodule B1 or B2) or design of the DAEPPI (submodule D1). Considering that the SMUT[RN1,PEs] passes model adequacy pre-test, we decide to design and perform a DAEPPI and expect that the model parameter precision as well as the model prediction precision increases after the 1st DAEPPI so that the SMUT[RN1,PEs+1st DAEPPI] passes model accuracy pre-test.

Table 4.18: Results of model-based experimental design for the VEPO.

N° model-based experimental design (Strategy application stage)		1 (Stage III)	2 (Stage V)	3 (Stage VII)
SMUT used		SMUT1 [a]	SMUT2 [b]	SMUT3 [c]
VEPO	Optimal reactor	BSTR	BSTR	BSTR
	t [s]	7.8762×10^{3}	6.3560×10^{3}	6.4072×10^{3}
	T^{int} [K]	343.0719	343.0667	343.1467
	n_{TO}^{int} [mol]	0.0610	0.0610	0.0610
	n_{TL}^{int} [mol]	0.1410	0.1410	0.1410
	n_{TP}^{int} [mol]	0.0259	0.0259	0.0259
	n_{E}^{int} [mol]	5.4580	5.4448	5.4433
	n_{NaOH}^{int} [mol]	0.0584	0.0466	0.0434
Predicted optimal performances	Criterion for VEPO (Eq. (2.20))	0.0014	0.0014	0.0014
	X_{TO} [%]	100.00	100.00	100.00
	X_{TL} [%]	99.98	99.98	99.98
	X_{TP} [%]	100.00	100.00	100.00
	Y_{FAEEs} [%]	74.80	74.79	74.79
Model prediction errors	C_{EO} [mol/m^3]	6.0422	4.1454	2.4704
	C_{EL} [mol/m^3]	4.9064	1.1912	0.7597
	C_{EP} [mol/m^3]	1.2417	1.0422	0.6470
	C_{MGs} [mol/m^3]	5.9789	5.2151	0.8425
	C_{DGs} [mol/m^3]	0.8541	0.3703	0.1603
	C_{TGs} [mol/m^3]	0.5673	0.0664	0.0643

[a] SMUT1 = SMUT[RN1,PEs];
[b] SMUT2 = SMUT[RN1,PEs+1st DAEPPI];
[c] SMUT3 = SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI].

4.2.3.6 Submodule D1: Design data acquisition experiment

The information contained in a DAEPPI is proportional to the number of experiments in the DAEPPI. Meanwhile, it has been noticed that some sets of operating conditions can be repeated several times during one run (Mathieu, 2013), which reduces data acquisition efficiency. Therefore, for this case study, the number of experiments contained in one DAEPPI is artificially set to 2 to make a compromise between information quantity and efficiency of data acquisition.

After defining the number of experiments in the DAEPPI, model-based experimental design of the DAEPPI is performed. The result (shown in Table 4.19) indicates that the 1st DAEPPI should be carried out using the BSTR at long experiment durations and near to maximal ethanol-to-oil molar ratio (inhibiting reverse reactions). It is not easy to interpret quantitatively why the software proposes such operating conditions, since the interactions among the reactions defined in the RN1 are complex.

Table 4.19: Results of model-based experimental design for the 1st DAEPPI.

Criterion for DAEPPI (Eq. (2.20))	0.6591	
Optimal reactor	BSTR	
N° experiment	1	2
τ [s]	3.2396×10^{3}	4.1306×10^{3}
T^{int} [K]	335.3072	335.3072
n_{TO}^{int} [mol]	0.0610	0.0610
n_{TL}^{int} [mol]	0.1410	0.1410
n_{TP}^{int} [mol]	0.0259	0.0259
n_{E}^{int} [mol]	5.4324	5.4324
n_{NaOH}^{int} [mol]	0.0197	0.0197

For this case study, the reactions, whose parameters $\log k_{j,T_{max}}$ and $\log k_{j,T_{min}}$ are both lower than -7.5, are considered as the reactions with relatively long characteristic times. It is noticed that the reverse reactions contain more such reactions in comparison to the forward reactions (see Table 4.17). Data acquisition efficiency is related to the sensitivity coefficient, which is defined as the derivatives of model responses with respect to the model parameters (see Eq. (1.9)). The parameters involved in the reverse reactions with the long characteristic times are more sensitive than the parameters of the other reactions at the operating conditions identified above. This is confirmed by the fact that the average decline of the confidence intervals of such parameters (64.87%) is larger than that of all parameter confidence intervals (44.09%).

4.2.4 Stage IV: First DAEPPI

The 1st DAEPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at the submodule D1 (see Section 4.2.3.6).

4.2.5 Stage V: Software application

4.2.5.1 Submodule D2: Input experimental data

The experimental responses of the 1st DAEPPI (shown in Table A2.1) are input into the software.

4.2.5.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs and the 1st DAEPPI. The result is shown in 2nd column of Table 4.17. The current SMUT is updated to SMUT[RN1,PEs+1st DAEPPI]. It can be noticed that model parameter precision increases, since the average variance of model parameters decrease by 46.91%. Meanwhile, not every parameter confidence interval of the SMUT[RN1,PEs+1st DAEPPI] decreases in comparison to that of the SMUT[RN1,PEs] (see the 1st

column of Table 4.17), since the DAEPPI is designed to reduce the average variance of model parameters, not the variance of each parameter. For example, the parameters associated to reaction N°3 exhibit an increasing interval width: this can be explained by the fact that the 1st DEAPPI explored very long residence times whereas this reaction mainly expresses itself at very short residence times. As a result, the 1st DEAPPI does not bring any additional information for an accurate identification of these parameters.

4.2.5.3 Submodule F1: Model adequacy pre-test

The SMUT[RN1,PEs] passes model adequacy pre-test, but that does not mean that SMUT[RN1,PEs+1st DAEPPI], whose reaction network is the same as the SMUT[RN1,PEs], passes also the test. Therefore, model adequacy pre-test should be carried out for the SMUT [RN1,PEs+1st DAEPPI]. Here, the SMUT[RN1,PEs+1st DAEPPI] passes model adequacy pre-test, since its parameter estimation criterion (24.0603) is smaller than that of SMUT[RN1,PEs] (24.3357) (see Table 4.17).

4.2.5.4 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is performed on the basis of the SMUT[RN1,PEs+1st DAEPPI]. The result is shown in the 2nd column of Table 4.18. The same optimal reactor and similar optimal operating conditions as those of the 1st predicted VEPO are identified. Model accuracy pre-test fails again, since the model prediction error of C_{MGs} (5.2151 mol/m^3) is still larger than the threshold (5 mol/m^3). In view of the fact that the predicted error of each measurable variable decreases after the 1st DAEPPI, we decide to do one more DAEPPI and expect that the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI] passes model accuracy pre-test.

4.2.5.5 Submodule D1: Design data acquisition experiment

Model-based experimental design for the DAEPPI is performed. The result (shown in Table 4.20) indicates that the 2nd DAEPPI has to be carried out in the CTR (different from the reactor used for the 1st DAEPPI) under the minimal reaction temperature at the maximal and minimal ethanol-to-oil molar ratios. In this DAEPPI, the experiments are carried out at two extreme ethanol-to-oil molar ratios. We interpret this result by the fact that some parameters are more sensitive at one set of operating conditions, whereas the other parameters are more sensitive at another set of operating conditions, which allows obtaining maximum information on all parameters for improving their confidence intervals. In addition, the DAEPPI in the CTR will provide more information at the beginning of the synthesis, enabling to complete the available information for parameter estimation.

Table 4.20: Result of model-based experimental design for the 2nd DAEPPI.

Criterion for DAEPPI (see Eq. (2.20))	0.6752	
Optimal reactor	CTR	
N° experiment	1	2
T^{in} [K]	303.1500	303.1500
F^{in} [m^3/s]	4.0000×10^{-7}	1.0000×10^{-7}
c^{in}_{TO} [mol/m^3]	46.4457	185.6920
c^{in}_{TL} [mol/m^3]	107.3841	429.3261
c^{in}_{TP} [mol/m^3]	19.7342	78.8983
c^{in}_{E} [mol/m^3]	1.4265×10^{4}	5.7138×10^{3}
c^{in}_{NaOH} [mol/m^3]	2.6151	2.2506

4.2.6 Stage VI: Second DAEPPI

The 2nd DAEPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at D1 (see Section 4.2.5.5).

4.2.7 Stage VII: Software application

4.2.7.1 Submodule D2: Input experimental data

The experimental responses of the 2nd DAEPPI (shown in Table A2.1) are input into the software.

4.2.7.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs, the 1st DAEPPI and the 2nd DAEPPI. The current SMUT is updated to SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]. It has been noticed that model parameter precision increases by 98.43% in comparison to the SMUT[RN1,PEs]. The confidence interval for each parameter becomes narrow enough, the average and maximal relative error of model parameters being respectively 9.33% and 32.86% (see the 3rd column of Table 4.17).

4.2.7.3 Submodule F1: Model adequacy pre-test

The SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI] passes model adequacy pre-test, since its parameter estimation criterion (23.521) is smaller than the previous one (24.0603) (see Table 4.17).

4.2.7.4 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is performed on the basis of the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]. The result is shown in the 3rd column of Table 4.18. The predicted error of each measurable variable is less than its corresponding threshold, therefore,

the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI] passes model accuracy pre-test. The next step is to perform the VEPO experimentally.

4.2.8 Stage VIII: VEPO

The VEPO is performed according to the practical operating conditions derived from the operating conditions in general form obtained at the submodule F2 (see Section 4.2.7.4).

4.2.9 Stage IX: Software application

4.2.9.1 Submodule F3: Model adequacy and accuracy post-tests

Figure 4.11 presents the comparison between the experimental responses of the VEPO (shown in Table A2.1) and those predicted by the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]. The SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI] seems to be able to predict well all the measurable variables within an acceptable tolerance (even if the actual residual of total concentration of MGs is greater compared with others, it is also accepted), the SMUT[RN1,PE+1st DAEPPI+2nd DAEPPI] is therefore validated.

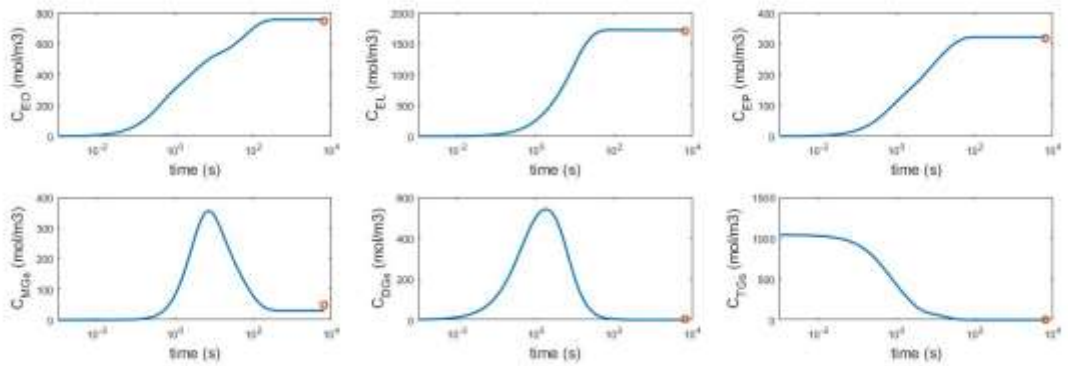


Figure 4.11: Concentration profiles for the VEPO using the BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]; singular points: experimental responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).

It can be seen that the optimal reaction duration seems to be 600 s, since the composition of the reaction system changes no more (i.e. the chemical equilibrium is reached) from 600 s. For this case study, the loss of the theoretically optimal reaction duration is due to the fact that the reaction duration is not considered as a performance criterion. Therefore, for the next case study, the reaction duration will be taken into consideration.

4.2.9.2 Module G: Results

The final results about the validated SMUT (i.e. the SMUT[RN1,PEs+1st DAEPPI+2nd DAEPPI]) and the optimal reactor (i.e. the BSTR) with its associated optimal operating conditions (see the 3rd column of Table 4.18) are displayed and stored in the file created at the submodule A1.

4.3 Strategy feasibility analysis

Table 4.21 shows the comparison of kinetic parameters obtained in this work and in the literature (illustrated by the ethanolysis of TP). The kinetic parameters we determined (derived from the parameter actually estimated) are of the same orders of magnitude as those obtained from the work of Likozar and Levec (2014b), except the pre-exponential factors for the reactions N° 33 and 36. Further assessment about the rationality of those kinetic parameters characterizing the reactions N° 33 and 36 is performed by evaluating whether such a kinetic parameter estimate in this work is comprised in its corresponding confidence interval derived from literature. The assessment result is shown in Table 4.22. It is observed that each parameter estimate is comprised in its corresponding confidence interval derived from literature, even if $\log(k_{36,Tmax})$ is very close to its lower bound. In view of the above facts, the estimation result of our work is deemed to be reasonable.

Table 4.21: Comparison of pre-exponential factors and activation energies obtained from this work and from literature.

N° Reaction	This work		Likozar and Levec (2014b)	
	Pre-exponential factor, A [$m^6/(mol^2 \cdot s)$]	Activation energy, Ea [J/mol]	Pre-exponential factor, A [$m^6/(mol^2 \cdot s)$]	Activation energy, Ea [J/mol]
31	2.29×10^{-1}	3.29×10^{4}	4.50×10^{-1}	5.50×10^{4}
32	3.87×10^{-1}	3.98×10^{4}	7.50×10^{-1}	4.92×10^{4}
33	1.86×10^{1}	5.44×10^{4}	2.42×10^{0}	4.80×10^{4}
34	4.00×10^{-3}	2.07×10^{4}	2.00×10^{-3}	3.80×10^{4}
35	3.84×10^{-3}	2.15×10^{4}	2.17×10^{-3}	3.10×10^{4}
36	1.27×10^{-6}	1.91×10^{4}	1.17×10^{-3}	3.42×10^{4}

Table 4.22: Check of parameter estimates in this work and parameter confidence intervals derived from literature.

N° Reaction	Estimate (this work)		Confidence interval (Likozar and Levec, 2014b)	
	$\log(k_{Tmin})$ [$m^6/(mol^2 \cdot s)$]	$\log(k_{Tmax})$ [$m^6/(mol^2 \cdot s)$]	$\log(k_{Tmin})$ [$m^6/(mol^2 \cdot s)$]	$\log(k_{Tmax})$ [$m^6/(mol^2 \cdot s)$]
33	-8.1066	-6.8249	[-8.6084 , -7.5659]	[-7.3733 , -6.4750]
36	-9.1845	-8.7350	[-9.6551 , -8.9242]	[-8.7352 , -8.1039]

4.4 Software practicability analysis

The practicability of the developed strategy-based software has been highlighted through the case study 1. The software provides not only user-friendly interfaces, such as Figures 4.5 and 4.7 used for defining respectively the performance criteria and the variables to be actually measured, but also powerful functionalities such as model development, model evaluation, model-based experimental design for model refining and performance optimization, even if several functions of the strategy

are not implemented (see Section 3.1). However, software application also exhibits some inconveniences:

- The user has to follow strictly the instructions proposed by the software. For instance, while specifying constraints on operating conditions in general form at the submodule C1 (see Sections 3.1.3.1 and 4.2.1.8), the user realizes that he has not reasonably defined one or several species properties (e.g. (i) the user defines wrongly the density of a reactant; (ii) a product has to present in both the organic and aqueous phases, whereas the user indicates that this product presents only in the organic phase) and wants to go back to the submodule A3 (see Sections 3.1.1.3 and 4.2.1.3) to redefine species proprieties. The initial version of the strategy-based software does not support this idea. For that, the user has to escape and restart the software.
- The experimental data is manually input into the software by the user, which is not very convenient, and loading an excel file containing the experimental data seems better.
- Model adequacy is qualitatively evaluated by the user according to the arguments provided by the software, such as the parameter estimation criteria and the predicted curves by SMUTs. Therefore, implementation of quantitative evaluation of model adequacy appears necessary, which allows human intervention to be unnecessary for model adequacy evaluation.

4.5 Conclusion

In this chapter, the feasibility of the strategy and the practicability of the software are firstly demonstrated by the case study 1, i.e. NaOH-catalyzed ethanolysis of sunflower oil. An adequate and accurate SMUT, whose model prediction quality has been verified experimentally, is developed after a series of experiments for different purposes:

- 5 PEs designed by the user, providing the basic information for the initial model development, as well as allowing to identify the best reaction network,
- 2 DAEPPIs designed by the strategy-based software, increasing model accuracy,
- 1 VEPO designed also by the strategy-based software, verifying the prediction quality.

In the next chapter, the feasibility and the generality of the strategy as well as the practicability of the software will be tested by the case study 2, a more complex one, in which heat-transfer phenomenon is taken into consideration, i.e. epoxidation of sunflower oil by performic acid generated in situ. For this case study, reaction duration will also be considered as a performance criterion.

References

EN 14103:2011. Fat and oil derivatives – fatty acid methyl esters (FAME) – determination of ester and linolenic acid methyl ester contents.

EN 14105:2011. Fat and oil derivatives – fatty acid methyl esters (FAME) – determination of free and total glycerol and mono-, di-, triglyceride contents.

Jiang, Z., Portha, J.F., Commenge, J.M., Falk, L., 2019. Development and implementation of systematic model-development strategy using model-based experimental design. Chemical Engineering Research and Design, 146, 290-310.

Likozar, B., Levec, J., 2014b. Transesterification of canola, palm, peanut, soybean and sunflower oil with methanol, ethanol, isopropanol, butanol and tert-butanol to biodiesel: Modelling of chemical equilibrium, reaction kinetics and mass transfer based on fatty acid composition. Applied Energy, 123, 108-120.

Marjanović, A.V., Stamenković, O.S., Todorović, Z.B., Lazić, M.L., Veljković, V.B., 2010. Kinetics of the base-catalyzed sunflower oil ethanolysis. Fuel, 89(3), 665-671.

Mathieu, F., 2013. Thesis, Stratégie d'intensification des procédés. Université de Lorraine.

Reyero, I., Arzamendi, G., Zabala, S., Gandía, L.M., 2015. Kinetics of the NaOH-catalyzed transesterification of sunflower oil with ethanol to produce biodiesel. Fuel Processing Technology, 129, 147-155.

Richard, R., Thiebaud-Roux, S., Prat, L., 2013. Modelling the kinetics of transesterification reaction of sunflower oil with ethanol in microreactors. Chemical Engineering Science, 87, 258-269.

Chapter 5. Case study 2: Epoxidation of sunflower oil by performic acid generated in situ

In the previous chapter, the feasibility of the strategy and the practicability of the strategy-based software have been demonstrated by the case study 1, i.e. NaOH-catalyzed ethanolysis of sunflower oil. In this chapter, those will be further demonstrated by the case study 2, i.e. epoxidation of sunflower oil by performic acid generated in situ. In addition, the generalities of the strategy and the strategy-based software will be also tested by the case study 2, which differs from the case study 1 in that chemical reactions take place in both involved phases and heat-transfer phenomenon has to be taken into consideration.

This chapter starts with the presentation of the experimental aspects, focuses on the chronological application of the strategy to the in-situ epoxidation case study and ends with strategy feasibility and software practicability analyses. Literature review about reaction system is not included in this chapter, since it has been given in Section 3.2.2.5, while screening the validation reaction systems.

5.1 Materials and methods

In this section, firstly, the materials used for the in-situ epoxidation experiments are reported. Then, the experimental protocols and analytical methods will be described. The experimental set-up used for the case study 2 is the same as that used for the case study 1. Therefore, its description is not repeated in this section. For more details about it, please refer to Section 4.1.2.

5.1.1 Materials

Commercial sunflower oil is purchased from Auchan (Lobau), a local supermarket. Its fatty acid composition (% by weight) is measured as follow: 11% saturated acids, 29% mono-unsaturated acids, 60% poly-unsaturated acids. The other reagents used in this case study are shown in Table 5.1.

Table 5.1: Information about the reagents used for the case study 2.

Reagent	Grade	Supplied by	Used for
Formic acid	≥98%	Sigma Aldrich	Synthesis
Hydrogen peroxide	35%	Merck	
Sulfuric acid	95.0-97.0%	Sigma Aldrich	
Ethyl acetate	≥99.0% (GC)	Sigma Aldrich	Treating the samples
Sodium bicarbonate			
Sodium hydroxide in aqueous solution	1 N	VWR DBH PROLABO CHEMICALS	Analyzing the concentration of hydrogen ion in the aqueous phase
Bromothymol blue			
Sulfuric acid			Analyzing the concentration of hydrogen peroxide in the aqueous phase
Ammonium cerium(IV) sulfate dihydrate	-	Sigma Aldrich	
1,10-Phenanthroline [a]	≥99.5%	Prolabo	
Iron(II) sulfate heptahydrate [a]	≥99.0%	Sigma Aldrich	
Ethanol [a]	absolute anhydrous	Carlo Erba Reagents	
Chloroform	99.0-99.4%	Merck	Analyzing the concentration of double bond in the oil phase
Hanus solution	c(IBr)=0.1N	Merck	
Potassium iodide	≥99.5%	Sigma Aldrich	
Sodium thiosulphate pentahydrate	99.7%	VWR DBH PROLABO CHEMICALS	
Starch	-	Sigma Aldrich	
Tetraethylammonium bromide (TEABr)	99%	Sigma Aldrich	Analyzing the concentration of epoxide in the oil phase
Acetic acid	≥99.8%	Sigma Aldrich	
Perchloric acid	70%	Fisher Chemical	
Crystal violet			

[a] Reagent used for the preparation of the ferroin solution.

5.1.2 Experimental protocols

For this case study, the experiments are carried out respectively in Semi-Batch Stirred-Tank Reactor (SBSTR) under isoperibolic mode and in Continuous Tubular Reactor (CTR) under isothermal mode (see Table 4.2).

Before the experiments, two reactant solutions used for the synthesis have to be prepared. Multiple combinations (shown in Table 5.2) of solution constituents are possible, and can be divided into two modalities according to whether both Formic Acid (FA) and Hydrogen Peroxide (HP), i.e. the perhydrolysis reactants, are contained in one solution.

Table 5.2: Survey on different combination modalities of solution constituents (SA = Sulfuric Acid).

Combination modality	Solution 1	Solution 2	Reference
I	FA	Oil + HP	Zheng et al. (2016)
	HP	Oil + FA	Wu et al. (2016)
	HP	Oil + FA + SA	Kousaalya et al. (2018)
II	FA + HP	Oil + SA	Santacesaria et al. (2011); Moreno et al. (2017)
	FA + HP + SA	Oil	Santacesaria et al. (2012)

Considering the practical limit of the used pump (see Section 4.2.1.8.1), it can be expected that pumping the I-modality solution 2, i.e. the solution containing oil and one perhydrolysis reactant, increases significantly the difficulty of the practical operation (i.e. how to assure a reliable and reproducible flow rate), particularly when the ratio oil-to-HP or oil-to-FA is not constant. Therefore, in this work, the II-modality combination is adopted, specifically, solution 1 is defined as the mixture of FA solution and HP solution; solution 2 is defined as the mixture of SA solution (trace) and sunflower oil.

For the experiments in the SBSTR, solution 2 is charged into the reactor and heated to the desired initial reaction temperature. While heating solution 2, the mechanical stirring (550 rpm) is turned on. The synthesis is timed, when solution 1 at room temperature is added into the reactor with a constant flow rate by a pump. During the synthesis, the temperature of the fluid flowing in the jacket is maintained at the desired initial reaction temperature; samples (8 mL) are taken at the designed sampling times and quickly cooled.

For the experiments in the CTR, the preheated solution 2 and the solution 1 at room temperature are fed into the reactor through a T-mixer by two pumps. Samples (8 mL) are collected at the outlet and quickly cooled.

Samples withdrawn are centrifuged for 20 min at 7000 rpm to separate the aqueous and organic phases. The organic samples are firstly diluted in ethyl acetate, secondly washed by sodium bicarbonate solution (5 wt.%) to eliminate completely the acidity, and finally dried in TurboVap LV evaporator at 80 °C for 1 hour in order to remove ethyl acetate from the solution. Then, the aqueous and organic samples are stored in a fridge at 4 °C prior to the analyses (see Section 5.1.3).

5.1.3 Analytical methods

The concentrations of hydrogen ion and hydrogen peroxide in the aqueous phase are determined by respectively the conventional acid-base titration method (see Section 5.1.3.1) and the Greenspan and Mackellar method (see Section 5.1.3.2). The concentrations of double bond and epoxide in the organic phase are determined by respectively the Hanus method (see Section 5.1.3.3) and the Jay method (see Section 5.1.3.4). All the analyses are achieved with the aid of the automatic titrator (Mettler Toledo T70). The solutions used for analyses (shown in Table 5.3) have to be prepared in advance.

Table 5.3: Information about various solutions used for analyses.

Solution	Used as	Used for
Sodium hydroxide solution (0.2 N)	Titrant	Determination of the concentration of hydrogen ion in the aqueous phase
Bromothymol blue solution	Indicator	
Sulfuric acid solution (5 wt.%)	Solvent	Determination of the concentration of hydrogen peroxide in the aqueous phase
Ammonium cerium(IV) sulfate solution (0.1 N)	Titrant	
Ferroin solution	Indicator	
Potassium iodide solution (15 wt.%)	Reactant	Determination of the concentration of double bond in the organic phase
Sodium thiosulfate solution (0.1 N)	Titrant	
Starch solution	Indicator	
TEABR solution [a] (20 wt.%)	Reactant	Determination of the concentration of epoxide in the organic phase
Perchloric acid solution [a] (0.1 N)	Titrant	
Crystal violet solution [a]	Indicator	

[a] Acetic acid is used as solution solvent.

5.1.3.1 Concentration of hydrogen ion in the aqueous phase

The concentration of hydrogen ion in the aqueous phase is determined by the conventional acid-base titration method. The titration protocol is described as follow:

1. Take 0.4 mL sample into a 250 mL conical flask and add 50 mL distilled water to dilute sample.
2. Add about three drops of bromothymol blue indicator solution and titrate with sodium hydroxide solution (0.2 N) using the automatic titrator until the indicator changes (from yellow) to light green.

The concentration of hydrogen ion in the aqueous phase is calculated using Eq. (5.1).

$$c_{H^+,aq} = \frac{V \times N}{S} \quad (5.1)$$

where V is the volume of titrant solution used for titration of sample (mL); N is the normality of titrant solution (N); S is the volume of sample used (mL).

5.1.3.2 Concentration of hydrogen peroxide in the aqueous phase

The concentration of hydrogen peroxide in the aqueous phase is determined by the Greenspan and Mackellar method (Greenspan and Mackellar, 1948), which is based on the fact that the reaction between ammonium cerium(IV) sulfate (i.e. the titrant) and percarboxylic acid (e.g. performic acid) is not feasible, whereas the reaction between ammonium cerium(IV) sulfate and hydrogen peroxide (illustrated by Eq.(5.2)) is feasible.

$$H_2O_2 + 2(NH_4)_4Ce(SO_4)_4 \rightarrow Ce_2(SO_4)_3 + 4(NH_4)_2SO_4 + H_2SO_4 + O_2 \quad (5.2)$$

The measurement protocol is described as follow:

1. Take 0.15 mL sample into a 250 mL conical flask and add 50 mL sulfuric acid solution (5 wt.%) to dilute sample.
2. Add about three drops of ferroin indicator solution and titrate with ceric ammonium sulfate solution (0.1 N) using the automatic titrator until the disappearance of the color.

The concentration of hydrogen peroxide in the aqueous phase is calculated using Eq. (5.3).

$$c_{HP,aq} = \frac{V \times N}{2 \times S} \quad (5.3)$$

5.1.3.3 Concentration of double bond in the organic phase

The concentration of double bond in the organic phase is determined by the Hanus method (Paquot, 2013), which can be summarized in four steps:

1. Convert the unreacted double bond in the sample to the di-halogenated single bond according to Eq. (5.4) by adding excess iodine monobromide (IBr) solution;
2. Convert the remaining unreacted IBr to iodine (I_2) according to Eq. (5.5) by adding excess potassium iodide solution;
3. Determine the amount of the generated I_2 (i.e. the amount of the remaining unreacted IBr) by titrating with sodium thiosulfate solution, the corresponding titration reaction is represented by Eq. (5.6);
4. Determine the amount of the used excess IBr through a blank titration.

$$R\text{–}CH = CH\text{–}R' + IBr \rightarrow R\text{–}CHI\text{–}CHBr\text{–}R' \quad (5.4)$$

$$IBr + KI \rightarrow I_2 + KBr \quad (5.5)$$

$$I_2 + 2Na_2S_2O_3 \rightarrow 2NaI + Na_2S_4O_6 \quad (5.6)$$

The practical measurement protocol is designed as follow:

1. Take 0.25 mL sample into a 300 mL conical flask and add 10 mL chloroform to dilute sample.
2. Add 25 mL Hanus solution (0.1N IBr) into the conical flask, plug it with a stopper to avoid the sublimation of iodine when sample reacts with Hanus solution, shake it for 1 minute and leave it in a dark room for 30 minutes.
3. Add 10 mL potassium iodide solution (15 wt.%) and 100 mL distilled water (or water of equivalent purity), seal and shake for 30 seconds.

4. Titrate with sodium thiosulfate solution (0.1 N) using the automatic titrator until light yellow color is formed. Add about three drops of starch solution (and then the mixture solution turns to blue). Continue the titration until blue color has disappeared.
5. Measure the blank value by repeating steps (1)-(4) without sample.

The concentration of double bonds in the organic phase is calculated using Eq. (5.7).

$$c_{DB,org} = \frac{(B - V) \times N}{2 \times S} \quad (5.7)$$

where B is the volume of titrant solution used for titration of blank solution (mL).

5.1.3.4 Concentration of epoxide in the organic phase

The concentration of epoxide in the organic phase is determined by the Jay method (Maerker, 1965), which is based on the chemical reaction between epoxide and hydrogen bromide (HBr) generated in situ from perchloric acid and TEABr. The in-situ HBr generation reaction and the titration reaction are represented by respectively Eq. (5.8) and Eq. (5.9).

$$HClO_4 + TEABr \rightarrow HBr + TEAClO_4 \quad (5.8)$$

$$R\text{–}CH(O)CH\text{–}R' + HBr \rightarrow R\text{–}CHOH\text{–}CHBr\text{–}R' \quad (5.9)$$

The measurement protocol is described as follow:

1. Take 0.10 mL sample into a 300 mL conical flask and add 10 mL chloroform and 10 mL TEABr solution (20 wt.%).
2. Add about three drops of crystal violet indicator solution and titrate with perchloric acid solution (0.1 N) using the automatic titrator until the indicator changes (from violet) to light blue.

The concentration of epoxides in the organic phase is calculated using Eq. (5.10).

$$c_{Ep,org} = \frac{V \times N}{S} \quad (5.10)$$

5.2 Strategy application

This section presents the chronological application of the strategy to the in-situ epoxidation case study. Operational sketch of strategy application for this case study is the same as that for the ethanolysis case study (see Figure 4.4).

5.2.1 Stage I: Software application

5.2.1.1 Submodule A1: Name project

Case study 2 is named after epoxidation of sunflower oil by performic acid generated in situ. A file with this name is generated automatically, and all the simulation results will be stored in it.

5.2.1.2 Submodule A2: Characterize reactors

For this case study, the characteristics to be specified for the SBSTR, in which the experiments are carried out under isoperibolic mode, are not only its dimensions but also its mass, overall heat-transfer coefficient, heat-transfer surface and the heat capacity of its manufacturing material. Whereas those for the CTR, in which the experiments are carried out under isothermal mode, are only its dimensions. For more information about the reactor characteristics, please refer to Table 4.2.

5.2.1.3 Submodule A3: Define species

The actual constituents of the reaction mixture are very complex, which can be for instance illustrated by Table 5.4, listing all possible products (containing at least one double bond) generated from tri-olein (containing three double bonds). Therefore, simplification of reaction mixture constituents is necessary. For that, sunflower oil is assumed to be represented by a pseudo species containing only one double bond (hereinafter referred to as DB). Hence, a series of functional-group-containing pseudo species generated from the DB are adopted:

- epoxide-containing pseudo species (hereinafter referred to as Ep) generated from the DB due to the epoxidation;
- hydroxyl-containing and ketone-containing pseudo species (hereinafter referred to as respectively OH and Ket) generated from the Ep due to the epoxide ring-openings via respectively the nucleophilic substitution and the rearrangement. The nucleophilic substitution of Ep produces a hydroxyl and another functional group depending on the nucleophilic agent, therefore, the OH can be formiate-containing OH (OH_{FA}), perhydroxyl-containing OH (OH_{HP}), performiate-containing OH (OH_{PFA}), diol-containing OH (OH_{W}), pseudo dimer (Dim) and pseudo trimer (Trim).

Table 5.4: Diversity of the products generated from tri-olein.

N° Product	Functional group							
	Double bond −CH=CH−	Epoxide −CH(O)CH−	Hydroxyl −OH	Perhydroxyl −OOH	Formiate −OOCH	Performiate −OOOCH	Ether −O−	Ketone −CO−
1	2	1	0	0	0	0	0	0
2	2	0	2	0	0	0	0	0
3	2	0	1	1	0	0	0	0
4	2	0	1	0	1	0	0	0
5	2	0	1	0	0	1	0	0
6	2	0	0	0	0	0	0	1
7	1	2	0	0	0	0	0	0
8 [a]	1	1	2	0	0	0	0	0
9	1	1	1	1	0	0	0	0
10	1	1	1	0	1	0	0	0
11	1	1	1	0	0	1	0	0
12	1	1	0	0	0	0	0	1
13	1	0	4	0	0	0	0	0
14	1	0	2	2	0	0	0	0
15	1	0	2	0	2	0	0	0
16	1	0	2	0	0	2	0	0
17	1	0	0	0	0	0	0	2
18	1	0	3	1	0	0	0	0
19 [b]	1	0	3	0	1	0	0	0
20	1	0	3	0	0	1	0	0
21	1	0	2	0	0	0	0	1
22	1	0	2	1	1	0	0	0
23	1	0	2	1	0	1	0	0
24	1	0	1	1	0	0	0	1
25	1	0	2	0	1	1	0	0
26	1	0	1	0	1	0	0	1
27	1	0	1	0	0	1	0	1
28	1	0	2	0	0	0	1	0
29	1	0	1	1	0	0	1	0
30	1	0	1	0	1	0	1	0
31	1	0	1	0	0	1	1	0

[a] Its molecule structure is shown in Figure 5.1.

[b] Its molecule structure is shown in Figure 5.2.

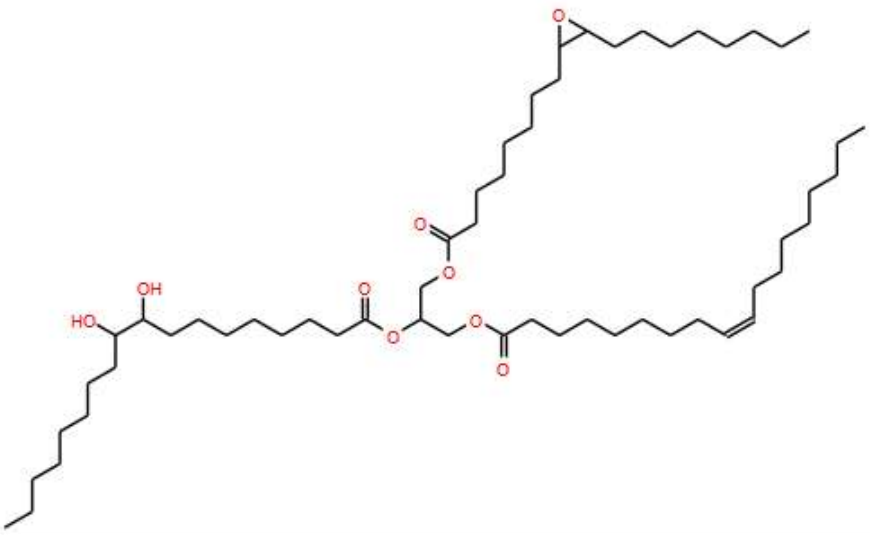


Figure 5.1: Molecule structure of the product N°8.

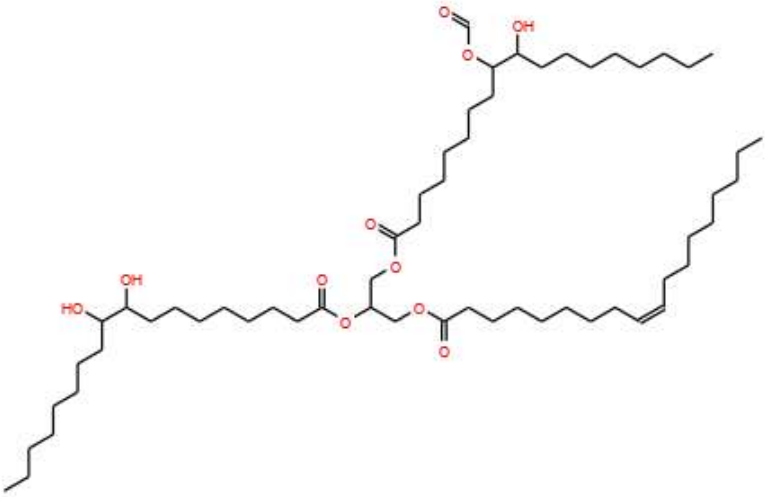


Figure 5.2: Molecule structure of the product N°19.

Thus, all possible species in the modeled reaction system are shown in Table 5.5. In the background, the operating conditions in general form (shown in Table 5.11) and the time-dependent state variables (see Figure 5.5) are automatically generated by the software.

Table 5.5: Information about all possible species in the reaction system for the case study 2.

Species	Molar mass [g·mol^{-1}]	Density [kg·m^{-3}]	Specific heat capacity [J·K^{-1}·kg^{-1}]	Species type	In phase I [a]?	In phase II [b]?
FA	46.03	1168.16	2213.84	Reactant	Yes	Yes
HP	34.01	1401.31	2745.48	Reactant	Yes	Yes
SA_{eq} [c]	49.04	1840.00	1468.84	Catalyst	Yes	No
PFA [d]	62.02	1168.16	2213.84	Reactant	Yes	Yes
W	18.02	985.03	4198.65	Solvent	Yes	Yes
DB	196.99	913.52	2093.40	Reactant	No	Yes
Ep [e]	212.99	913.52	2093.40	Product	No	Yes
OH_{FA} [e]	259.02	913.52	2093.40	Product	No	Yes
OH_{HP} [e]	247.01	913.52	2093.40	Product	No	Yes
OH_{PFA} [e]	275.01	913.52	2093.40	Product	No	Yes
OH_{W} [e]	231.01	913.52	2093.40	Product	No	Yes
Ket [e]	212.99	913.52	2093.40	Product	No	Yes
Dim [f]				Product	No	Yes
Trim [f]				Product	No	Yes
CO_2 [g]				Product	No	No
O_2 [g]				Product	No	No

[a] Phase I is the aqueous phase.

[b] Phase II is the organic phase.

[c] Reaction rate of a chemical reaction in the homogeneous catalytic reaction system is a power function of the concentrations of the associated species (e.g. reactant molecules, catalyst anion or ion) and proportional to its reaction rate constant (see Eq. (2.6)). The present version of the software models only the concentrations of the species in molecule form. Therefore, the concentration of the catalyst in anionic or ionic form has to be represented as a function of the concentration(s) of the associated actually existing molecule(s) in the reaction system. The simplest function is that the concentration of the catalyst anion or ion is equal to that of an actually existing species molecule. Specifically, for this case study, the concentration of the hydrogen ion is equal to that of the equivalent SA, denoted as SA_{eq}, the molar mass of the SA_{eq} is equal to the half of that of the SA, the density and specific heat capacity of SA_{eq} are equal to those of the SA, since the concentration of the hydrogen ion is twice that of the SA by neglecting FA-autocatalyzed reaction.

[d] PerFormic Acid (PFA) is considered as reactant, since the feed solution 1, i.e. the mixture of FA solution and HP solution, is actually an oxidizing mixture containing not only FA and HP but also PFA and water due to FA-autocatalyzed reaction (Sun et al., 2011).

[e] The products generated from the DB, such as the Ep, the OHs and the Ket, are assumed to possess the same density and specific heat capacity as those of the DB.

[f] Properties of Dim and Trim are not given, since no initial reaction networks under test (see Table 5.8) include the reactions involving them.

[g] Properties of CO_2 and O_2 are not given, since while constructing the model structure for a liquid-phase reaction system, gases escaped from the liquid phase are not modeled.

5.2.1.4 Submodule A4: Define performance criteria

For this case study, the performances to be actually optimized are the yield of the epoxide group, the maximal temperature of the reaction system and the reaction duration. The initial version of the software does not allow the user to define the yield-type performance criterion, but the conversion-type and selectivity-type performance criteria (see Section 3.1.1.4). Therefore, practically, the yield

of the epoxide group is represented as the product of two alternative performances, i.e. the conversion of the double bond group and the selectivity of the epoxide group.

After the specification of the alternative performance criteria (shown in Table 5.6), the corresponding default criterion for the Validation Experiment for Performance Optimization (VEPO) represented by Eq. (5.11) has to be modified to Eq. (5.12). The definition of the new criterion for the VEPO is achieved by modifying the codes at the background of the software.

Table 5.6: Alternative performance criteria for the case study 2.

Performance	Symbol	Unit	Objective value	Veto value
Reaction duration	τ	s	300	28800
Conversion of double bond group	X_{DB}	%	100	0
Selectivity of epoxide group	Y_{Ep}	%	100	0
Maximal temperature of reaction system	T_{max}	K	298.15	368.15

$$\varphi(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) = \frac{1}{\sqrt{4}}\left(\tau^{n}(\boldsymbol{\xi}) + X_{DB}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) + Y_{Ep}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) + T_{max}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi})\right) \quad (5.11)$$

$$\varphi(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) = \frac{1}{\sqrt{3}}\left(\tau^{n}(\boldsymbol{\xi}) + X_{DB}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) \times Y_{Ep}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}) + T_{max}^{n}(\hat{\boldsymbol{\theta}},\boldsymbol{\xi})\right) \quad (5.12)$$

where the superscript ‘n’ stands for ‘normalized’, the first three normalized performances (i.e. τ^{n}, X_{DB}^{n} and Y_{Ep}^{n}) are calculated according to Eq. (1.21); whereas the last normalized performance (i.e. T_{max}^{n}) is calculated according to:

$$T_{max}^{n}\left(\hat{\boldsymbol{\theta}},\boldsymbol{\xi}_{R_q}\right) = \begin{cases} 1, \text{if } T_{max} \geq T_p \\ 0, \text{if } T_{max} < T_p \end{cases} \quad (5.13)$$

where T_p represents the veto value of the maximal temperature of reaction system, i.e. 368.15 K.

5.2.1.5 Submodule B1: Develop reaction supernetwork

On the basis of the reaction supernetwork for the epoxidation of vegetable oil by PerCarboxylic Acid (PCA) generated in situ (see Table 3.6), the reaction supernetwork for the epoxidation of sunflower oil by PFA generated in situ (see Table 5.7) can be obtained by:

- deleting all the reactions involving *the undecomposed epoxide group contained in the hydroxyl-containing decomposition compound* (i.e. the reactions with the superscripts ‘a’ and ‘c’ in Table 3.6), since *such groups* are represented by Ep;
- replacing the involved Carboxylic Acid (CA) and PCA by respectively FA and PFA.

Table 5.7: Reaction supernetwork for the case study 2.

N°	Chemical reaction or mass transfer	Essential degree	Description
1	$FA(I) + HP(I) \xrightarrow{H^+(I)} PFA(I) + W(I)$	I-level	Acid perhydrolysis of FA
2 [a]	$PFA(I) + W(I) \xrightarrow{H^+(I)} FA(I) + HP(I)$	I-level	Acid hydrolysis of PFA
3	$PFA(I) + HP(I) \xrightarrow{H^+(I)} FA(I) + W(I) + O_2(g)$	III-level	Acid perhydrolysis of PFA
4	$PFA(I) \xrightarrow{H^+(I)} FA(I) + 1/2O_2(g)$	III-level	Decomposition of PFA
5	$PFA(I) \rightarrow W(I) + CO_2(g)$	III-level	
6	$PFA(I) \rightarrow 2/3FA(I) + 1/3W(I) + 1/3O_2(g) + 1/3CO_2(g)$	II-level	
7	$HP(I) \xrightarrow{H^+(I)} W(I) + 1/2O_2(g)$	III-level	Decomposition of HP
8	$DB(II) + PFA(II) \rightarrow Ep(II) + FA(II)$	I-level	Epoxidation of DB
9	$Ep(II) + FA(I) \xrightarrow{H^+(I)} OH_{FA}(II)$	II-level	Conventional epoxide ring-opening at the aqueous-organic phase
10	$Ep(II) + HP(I) \xrightarrow{H^+(I)} OH_{HP}(II)$	II-level	
11	$Ep(II) + PFA(I) \xrightarrow{H^+(I)} OH_{PFA}(II)$	II-level	
12	$Ep(II) + W(I) \xrightarrow{H^+(I)} OH_W(II)$	II-level	
13	$Ep(II) + OH(II) \xrightarrow{H^+(I)} Dim(II)$	III-level	Oligomerization
14	$Ep(II) + Dim(II) \xrightarrow{H^+(I)} Trim(II)$	III-level	
15	$Ep(II) \xrightarrow{H^+(I)} Ket(II)$	III-level	Nucleophilic rearrangement of Ep
16	$Ep(II) + FA(II) \rightarrow OH_{FA}(II)$	III-level	Epoxide ring-opening in the organic phase
17	$Ep(II) + HP(II) \rightarrow OH_{HP}(II)$	III-level	
18	$Ep(II) + PFA(II) \rightarrow OH_{PFA}(II)$	III-level	
19	$Ep(II) + W(II) \rightarrow OH_W(II)$	III-level	
20	$OH_{FA}(II) + W(I) \xrightarrow{H^+(I)} OH_W(II) + FA(I)$	III-level	Acid hydrolysis of OH_{FA}
21	$FA(I) \rightarrow FA(II)$	I-level	Mass transfer
22	$HP(I) \rightarrow HP(II)$	III-level	
23	$PFA(I) \rightarrow PFA(II)$	I-level	
24	$W(I) \rightarrow W(II)$	III-level	

[a] Endothermic reaction. Without special declaration, chemical reactions are exothermic.

It can be seen that the essential degree of an item involved in this reaction supernetwork is ranked into three levels, according to whether it is essential in the initial Reaction Networks Under Test (RNUTs) proposed in Section 5.2.1.6.

1. I-level items, such as acid perhydrolysis of FA, epoxidation of DB by PFA, mass transfers of FA and PFA between the aqueous and organic phases, *are essential* in the initial RNUTs;
2. II-level items, such as generalized PFA decomposition (representing the other two PFA decompositions) and conventional epoxide ring-opening reactions (taking place at the aqueous-

organic interface), *seem essential* in the initial RNUTs, however, whether they are included in the final RNUT or not has to be tested;

3. III-level items *are not essential* in the initial RNUTs, however, following the failure of the evaluation of the model structures constructed based on the initial RNUTs, they have to be taken into consideration while proposing the new RNUTs.

5.2.1.6 Submodule B2: Propose candidate reaction networks

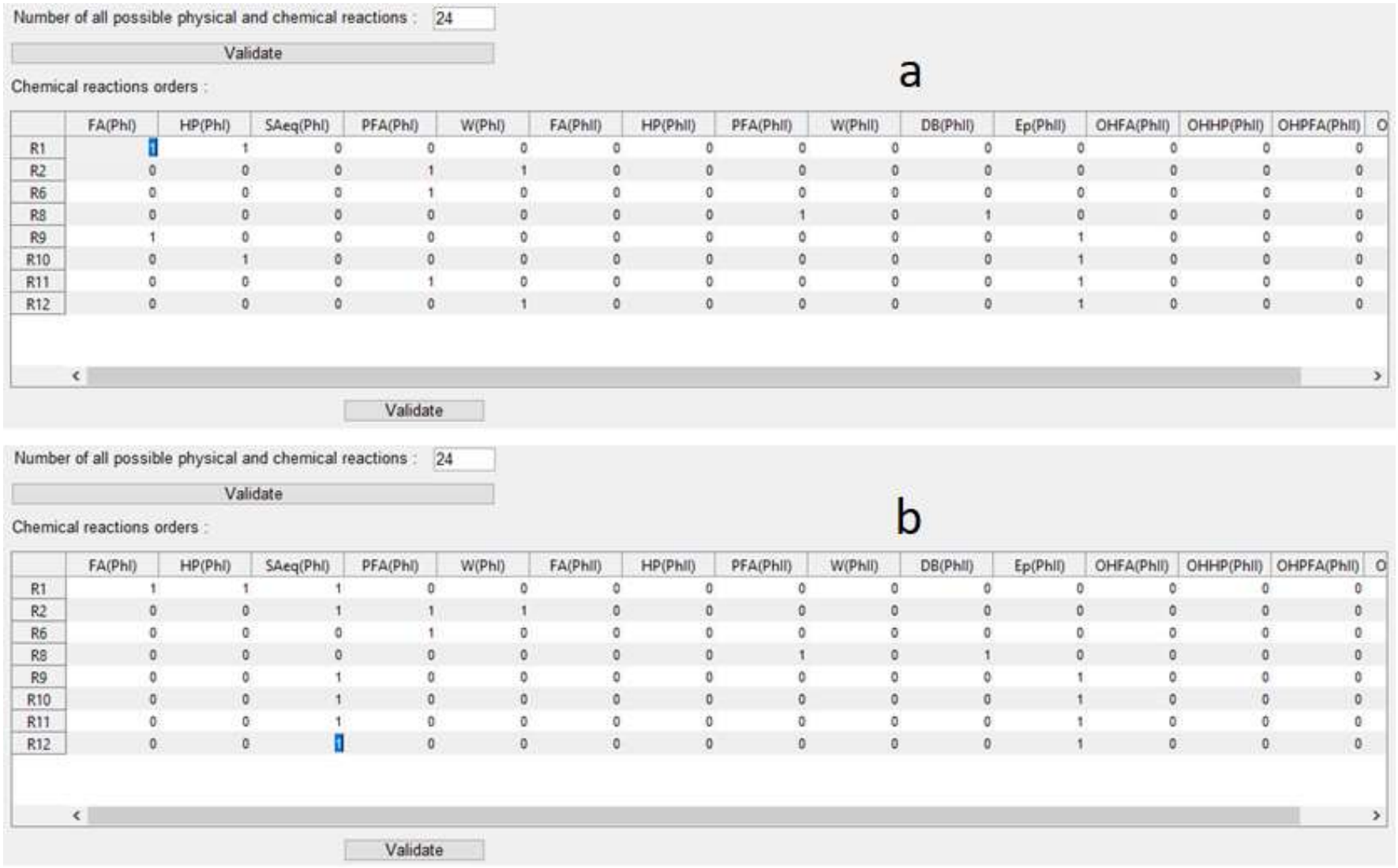
For this case study, six initial RNUTs (illustrated in Table 5.8) are proposed from the reaction supernetwork developed above.

Table 5.8: Reaction networks under test for the case study 2.

N° RNUT	Chemical reactions and mass transfers taken into consideration			
	Essential reaction	Decomposition	ring-opening	Mass transfer
1	1, 2, 8	–	9, 11	21, 23
2	1, 2, 8	6	9, 11	21, 23
3	1, 2, 8	–	12	21, 23
4	1, 2, 8	6	12	21, 23
5	1, 2, 8	–	9–12	21, 23
6	1, 2, 8	6	9–12	21, 23

Once the RNUTs are available, in the background of the software, the involved parameters (to be estimated) are automatically generated: (i) reaction rate constants, kinetic orders and molar reaction enthalpies for the chemical reactions, (ii) overall volumetric mass-transfer coefficients and distribution coefficients for the mass transfers (see Section 2.3.1).

The initial version of the software does not consider the kinetic orders as parameters to be estimated (see Section 3.1). Therefore, the user has to validate the kinetic orders generated automatically by the software or customize the kinetic orders on the basis of his expertise. Figure 5.3 illustrates how to customize the kinetic orders for the RNUT6. For example, the concentration of catalyst, equal to that of the SA_{eq}, has to be taken into consideration, while calculating the reaction rate of the chemical reaction N°1 (in Table 5.7), therefore, the kinetic order of the SA_{eq} for the chemical reaction N°1 is manually changed from 0 to 1.



	FA(PhI)	HP(PhI)	SAeq(PhI)	PFA(PhI)	W(PhI)	FA(PhII)	HP(PhII)	PFA(PhII)	W(PhII)	DB(PhII)	Ep(PhII)	OHFA(PhII)	OHHP(PhII)	OHPFA(PhII)	O
R1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
R2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	
R6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
R8	0	0	0	0	0	0	0	1	0	1	0	0	0	0	
R9	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
R10	0	1	0	0	0	0	0	0	0	0	1	0	0	0	
R11	0	0	0	1	0	0	0	0	0	0	1	0	0	0	
R12	0	0	0	0	1	0	0	0	0	0	1	0	0	0	

	FA(PhI)	HP(PhI)	SAeq(PhI)	PFA(PhI)	W(PhI)	FA(PhII)	HP(PhII)	PFA(PhII)	W(PhII)	DB(PhII)	Ep(PhII)	OHFA(PhII)	OHHP(PhII)	OHPFA(PhII)	O
R1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	
R2	0	0	1	1	1	0	0	0	0	0	0	0	0	0	
R6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
R8	0	0	0	0	0	0	0	1	0	1	0	0	0	0	
R9	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
R10	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
R11	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
R12	0	0	1	0	0	0	0	0	0	0	1	0	0	0	

Figure 5.3: Interface for the customization of the chemical reaction orders for the RNUT6: (a) interface before the definition, (b) interface after the definition.

5.2.1.7 Submodule B3: Select model parameters to be estimated

Given the involved parameters of any RNUT, the corresponding parameters to be actually estimated can be obtained by adopting the appropriate model parameter reformulations (shown in Table 5.9). Note that the minimal and maximal possible reaction temperatures correspond respectively to 298.15 K and 368.15 K (see Table 5.6).

Table 5.9: Model parameter reformulation for the case study 2.

Item	Involved parameter	Dependence type	Parameter logarithm [a] ?	Independent parameter
Chemical reaction j	k_j	Temperature-dependent	Yes	$\log(k_{j,Tmin})$, $\log(k_{j,Tmax})$
	Δh_j	Temperature-dependent	Yes	$\log(\Delta h_{j,Tref})$
Mass transfer of i	$m_{i,I/II}$	Independent	Yes	$\log(m_{i,I/II})$
	$K_{i,I,SBSTR}$, $K_{i,I,CTR}$	Independent	Yes	$\log(K_{i,I,SBSTR})$, $\log(K_{i,I,CTR})$

[a] If parameters exhibit very different orders of magnitude, it seems better to use the parameter logarithms, since in this situation, the parameters to be estimated exhibit the same or similar orders of magnitude.

5.2.1.8 Submodule C1: Specify constraints on operating conditions

The bounds on the practical operating conditions of each considered reactor are specified and shown in Table 5.10.

Table 5.10: Bounds on practical operating conditions of each considered reactor for the case study 2.

Practical operating condition	Symbol	Unit	SBSTR	CTR
Volume of formic acid solution in solution 1	$V_{FAS,1}$	mL	0 – 300	0 – 300
Volume of hydrogen peroxide solution in solution 1	$V_{HPS,1}$	mL	0 – 300	0 – 300
Volume of sulfuric acid solution in solution 2	$V_{SAS,2}$	mL	0.1 – 2.5	0.1 – 2.5
Volume of oil in solution 2	$V_{oil,2}$	mL	300 [a]	300
Initial reaction temperature [b]	T^{int}	°C	25 – 90	25 – 95
Temperature of feed 1 [c]	T_1	°C	25	25
Flow rate of feed 1	F_1	mL/min	2 – 20	2 – 20
Duration of feed 1	t_1	min	5 – 60	
Temperature of feed 2	T_2	°C		25 – 95
Flow rate of feed 2 [d]	F_2	mL/min		7.1
Duration of feed 2	t_2	min		
Experiment duration	τ	min	5 – 480	

[a] In order that any experiment in the SBSTR is identifiable, the volume of oil in solution 2 (i.e. the volume of oil initially charged into the reactor) is defined as constant. For more details, refer to Section 4.2.1.8.1.

[b] The initial reaction temperature is equal to the temperature of fluid flowing in the jacket (SBSTR) and to the water bath temperature (CTR).

[c] In order to avoid the decomposition of PFA generated from FA and HP through FA-autocatalyzed reaction in solution 1, the temperature of feed 1 is set to 25 °C, i.e. the room temperature.

[d] In order to assure the reliability and the reproducibility of the flow rate of feed 2, it is defined as constant. For more details, refer to Section 4.2.1.8.1.

Then, the constraints of the corresponding operating conditions in general form, actually input into the software, are specified. For consistency, the corresponding operating conditions in general form are subject not only to their bounds (shown in Table 5.11) derived from those of the practical operating conditions, but also to the equality and inequality constraints among them (described in Table 5.12).

Table 5.11: Bounds on operating conditions in general form of each considered reactor for the case study 2.

Operating condition in general form	Symbol	Unit	SBSTR	CTR
Experiment duration	τ	s	300 – 28800	
Initial reaction temperature	T^{int}	K	298.15 – 363.15	
Initial load of formic acid	n_{FA}^{int}	mol	0	
Initial load of hydrogen peroxide	n_{HP}^{int}	mol	0	
Initial load of equivalent sulfuric acid [a]	n_{SAeq}^{int}	mol	0.0038 – 0.0938	
Initial load of performic acid	n_{PFA}^{int}	mol	0	
Initial load of water	n_{W}^{int}	mol	0	
Initial load of double bond group	n_{DB}^{int}	mol	1.3912	
Temperature of feed I	T_{I}^{in}	K	298.15	298.15 – 368.15
Temperature of feed II	T_{II}^{in}	K	0 [b]	298.15 – 368.15
Flow rate of feed I	F_{I}^{in}	$m^3 \cdot s^{-1}$	3.3333×10^{-8} – 3.3333×10^{-7}	3.3373×10^{-8} – 3.3431×10^{-7}
Flow rate of feed II	F_{II}^{in}	$m^3 \cdot s^{-1}$	0 [b]	1.1736×10^{-7} – 1.1829×10^{-7}
Duration of feed I	t_{I}^{in}	s	300 – 3600	
Duration of feed II	t_{II}^{in}	s	0 [b]	
Concentration of formic acid in feed I	$c_{FA,I}^{in}$	$mol \cdot m^{-3}$	0 – 25378.2 [c]	0 – 25378.2 [c]
Concentration of hydrogen peroxide in feed I	$c_{HP,I}^{in}$	$mol \cdot m^{-3}$	0 – 41202.9 [c]	0 – 41202.9 [c]
Concentration of equivalent sulfuric acid in feed I	$c_{SAeq,I}^{in}$	$mol \cdot m^{-3}$	0	0 – 37520.4 [c]
Concentration of performic acid in feed I	$c_{PFA,I}^{in}$	$mol \cdot m^{-3}$	0 – 18835.2 [c]	0 – 18835.2 [c]
Concentration of water in feed I	$c_{W,I}^{in}$	$mol \cdot m^{-3}$	0 – 54663.2 [c]	0 – 54663.2 [c]
Concentration of formic acid in feed II	$c_{FA,II}^{in}$	$mol \cdot m^{-3}$	0 [b]	0
Concentration of hydrogen peroxide in feed II	$c_{HP,II}^{in}$	$mol \cdot m^{-3}$	0 [b]	0
Concentration of performic acid in feed II	$c_{PFA,II}^{in}$	$mol \cdot m^{-3}$	0 [b]	0
Concentration of water in feed II	$c_{W,II}^{in}$	$mol \cdot m^{-3}$	0 [b]	0
Concentration of double bond group in feed II	$c_{DB,II}^{in}$	$mol \cdot m^{-3}$	0 [b]	4637.4

[a] Considering that the amount of the concentrated SA (95-97 wt.%) is far less than those of other reagents, the amount of water (i.e. minor constituent) in SA solution is assumed to be neglected.

[b] For the experiments in the SBSTR, the (organic-phase) feed II containing only oil is defined to be non-existent, since oil is initially charged into the reactor.

[c] For the sake of simplicity, the bound on the concentration of each constituent in the (aqueous-phase) feed I is defined between 0 and the concentration of pure species (i.e. the reciprocal of its molar volume).

Table 5.12: Equality and inequality constraints among operating conditions in general form for the case study 2.

N°	Constraint	Description	Used for
1	Eq. (5.14)	The equilibrium of the FA-autocatalyzed reaction is assumed to be attained. The reaction quotient, represented as a function of the concentration of involved species, is equal to the equilibrium constant, which is set to 0.8 at 25°C on the basis of literature finding (Filippis et al., 2009).	SBSTR, CTR
2	Eq. (5.15)	Mass conservation of water before and after the FA-autocatalyzed reaction under the assumption that the volumes before and after mixing FA solution and HP solution to prepare the feed solution 1 are the same.	
3	Eq. (5.16)	Volume balance for the feed I.	
4	Eq. (5.17)	The volume of solution I charged by pump is not greater than that of oil initially charged.	SBSTR
5	Eq. (5.18)	The feed duration is not greater than the reaction duration.	
6	Eq. (5.19)	The temperatures of feed I and feed II are the same and equal to the reaction temperature.	CTR
7	Eq. (5.20)	Volumic ratio between the catalyst-type species (i.e. SA_{eq}, practically present in the feeding solution 2) and the non-catalyst-type species (i.e. FA, HP, PFA and W, practically present in the feeding solution 1) in the virtual feed I is equal to the flow-rate ratio of the corresponding practical feeds (illustrated by Figure 5.4).	

$$\frac{c_{PFA,I}^{in} \cdot c_{W,I}^{in}}{c_{FA,I}^{in} \cdot c_{HP,I}^{in}} = Q \quad (5.14)$$

$$\left(c_{FA,I}^{in} + c_{PFA,I}^{in}\right) \cdot MM_{FA} \cdot \frac{1-\omega_{FA}}{\omega_{FA}} + \left(c_{HP,I}^{in} + c_{PFA,I}^{in}\right) \cdot MM_{HP} \cdot \frac{1-\omega_{HP}}{\omega_{HP}} = \left(c_{W,I}^{in} - c_{PFA,I}^{in}\right) \cdot MM_{W} \quad (5.15)$$

$$c_{FA,I}^{in} \cdot v_{FA} + c_{HP,I}^{in} \cdot v_{HP} + c_{PFA,I}^{in} \cdot v_{PFA} + c_{W,I}^{in} \cdot v_{W} = 1 \quad (5.16)$$

$$F_{I}^{in} \cdot t_{I}^{in} \leq V_{oil,2} \quad (5.17)$$

$$t_{I}^{in} \leq \tau \quad (5.18)$$

$$T_{I}^{in} = T_{II}^{in} \quad (5.19)$$

$$\frac{v_{SAeq} \cdot c_{SAeq,I}^{in}}{v_{FA} \cdot c_{FA,I}^{in} + v_{HP} \cdot c_{HP,I}^{in} + v_{PFA} \cdot c_{PFA,I}^{in} + v_{W} \cdot c_{W,I}^{in}} = \frac{F_2 - F_{II}^{in}}{F_{I}^{in} - (F_2 - F_{II}^{in})} \quad (5.20)$$

where Q is the reaction quotient, ω_{FA} and ω_{HP} are the mass fractions of respectively the FA solution and the HP solution used for the preparation of the feed solution 1, F_2 is the practical flow rate of the feed solution 2 (i.e. the mixture of SA and oil), equal to 7.1 mL/min (see Table 5.10).

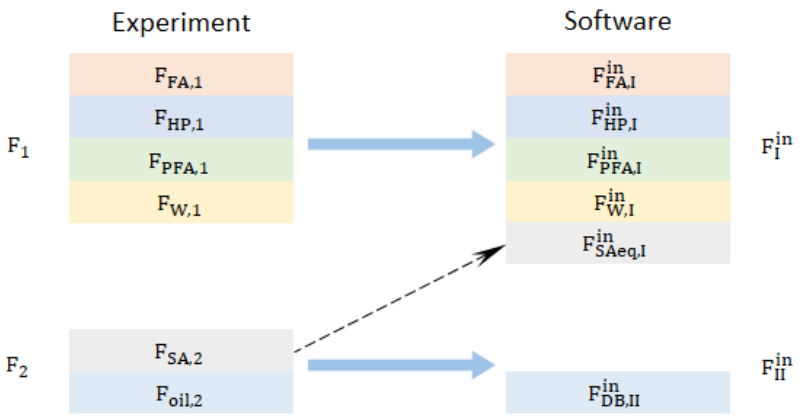


Figure 5.4: Illustration of the relationship between practical feeds and virtual feeds for the case study 2.

5.2.1.9 Submodule C2: Design measurable variables

Figure 5.5 illustrates how to define the measurable variables for this case study. The variables to be actually measured are divided into three types:

1. concentration-type conventional measurable variables: the concentration of HP in the phase I, the concentration of DB in the phase II and the concentration of Ep in the phase II,
2. temperature-type conventional measurable variable: reaction system temperature,
3. concentration-type unconventional measurable variable: the concentration of hydrogen ion in the phase I (i.e. the total concentration of FA, SA_{eq} and PFA in the phase I).

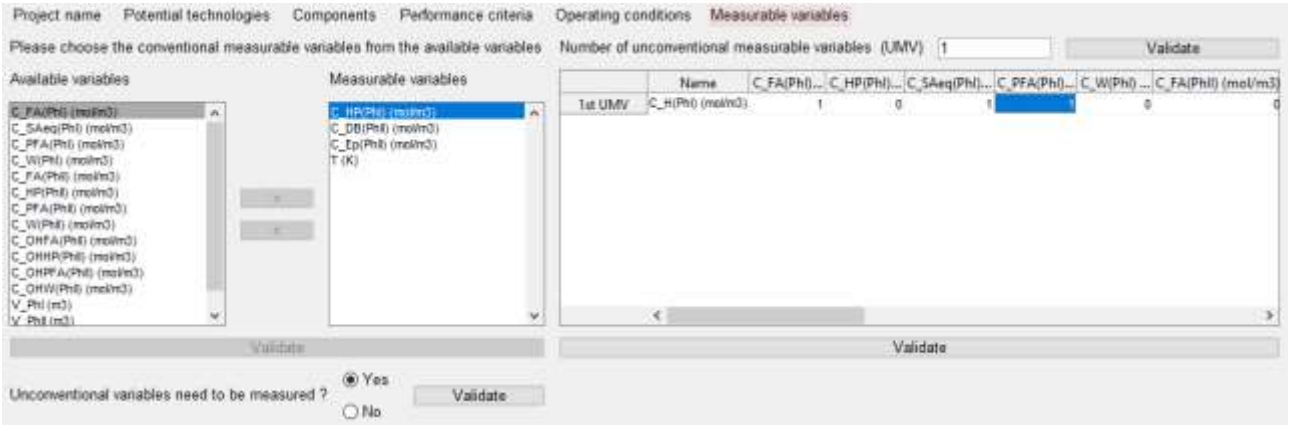


Figure 5.5: Interface for the selection of the measurable variables for the case study 2.

Given these measurable variables, the parameters involved in the model structures constructed are identifiable (Santacesaria et al., 2011).

5.2.1.10 Submodule D1: Design data acquisition experiments

The fractional factorial design of the initial data acquisition experiments, i.e. the Preliminary Experiments (PEs), is performed by the user and composed of three steps:

1. Identify the experimental factors under test, i.e. FA-to-DB molar ratio, HP-to-DB molar ratio, SA amount, initial reaction temperature;
2. Define the low and high levels for each considered experimental factor (see Table 5.13);

Table 5.13: Definition of the low and high levels for each considered experimental factor.

Factor	Low level (–)	High level (+)
FA-to-DB molar ratio [-]	0.23	0.34
HP-to-DB molar ratio [-]	0.75	1.12
SA amount [g per 100 g of oil]	0.64	1.28
Initial reaction temperature [°C]	55	75

3. Define the reference experiment for each considered reactor and the associated derivative experiments (see Table 5.14). Note that the reference experiment is carried out under the average-level initial reaction temperature, i.e. the average of the high-level and low-level initial reaction temperatures.

Table 5.14: Fractional factorial experimental design result for the case study 2 (+ = high-level; ± = average-level; - = low-level).

N°	Reactor	FA-to-DB molar ratio	HP-to-DB molar ratio	SA amount	Initial reaction temperature
1 [a]	SBSTR	-	-	-	±
2	SBSTR	-	-	+	±
3	SBSTR	-	-	-	+
4	SBSTR	-	-	-	-
5	SBSTR	-	+	-	±
6	SBSTR	+	-	-	±
7 [a]	CTR	-	-	-	±
8	CTR	-	-	-	+
9	CTR	-	-	-	-
10	CTR	-	+	-	±
11	CTR	+	-	-	±

[a] Reference experiment.

5.2.2 Stage II: Preliminary experiments

The PEs are performed according to the practical operating conditions (shown in Table 5.15) derived from the factorial experimental design result (see Table 5.14).

Table 5.15: Practical operating conditions of the PEs for the case study 2.

Run	1	2	3	4	5	6	7, 8, 9	10	11
Reactor	SBSTR	SBSTR	SBSTR	SBSTR	SBSTR	SBSTR	CTR	CTR	CTR
$V_{FAS,1}$ [mL]	16.5	16.5	16.5	16.5	16.5	24.75	13	13	19.5
$V_{HPS,1}$ [mL]	122	122	122	122	183	122	95	142.5	95
$V_{SAS,2}$ [mL]	1.3	2.6	1.3	1.3	1.3	1.3	1	1	1
$V_{oil,2}$ [mL]	410	410	410	410	410	410	320	320	320
T^{int} [°C]	65	65	75	55	65	65	55, 65, 75	65	65
T_1 [°C]	25	25	25	25	25	25	25	25	25
F_1 [mL·min^{-1}]	2.3	2.3	2.3	2.3	2.3	2.3	2.4	3.5	2.5
t_1 [min]	30	30	30	30	30	30			
T_2 [°C]							55, 65, 75	65	65
F_2 [mL·min^{-1}]							7.1	7.1	7.1
τ [min]	30, 60, 90, 120, 150, 180, 240, 300, 330, 360								

Figure 5.6 shows the evolution of experimental responses as a function of reaction time for the 4th PE. It is observed that the concentration of the hydrogen ion in the aqueous phase increases as the reaction progresses. This phenomenon may be explained by:

1. the decrease of the volume of the aqueous phase due to the consumption of the peroxidant, i.e. the HP,
2. the approximate constancy of the amount of hydrogen ion in the aqueous phase due to the fact that the couple FA/PFA is considered as the HP-carrier.

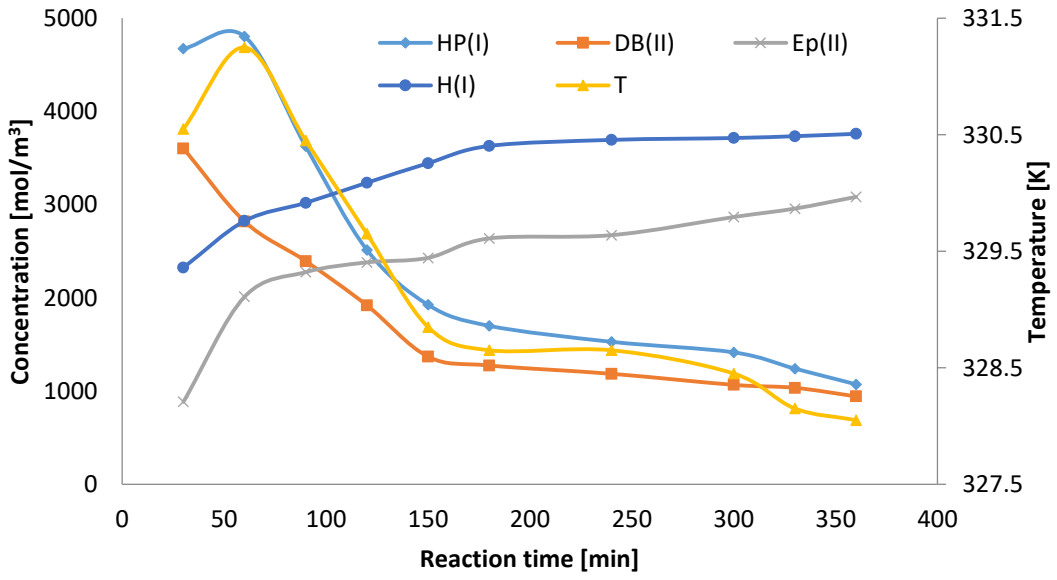


Figure 5.6: Evolution of the experimental responses for the 4th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 55°C, using the SBSTR.

5.2.3 Stage III: Software application

5.2.3.1 Submodule D2: Input experimental data

The operating conditions in general form (shown in Table A3.1 of Appendix 3), derived from the practical operating conditions of the PEs (see Table 5.15), and the corresponding experimental responses (see Table A3.2 of Appendix 3), are input into the software. The initial version of the software is designed to be capable to estimate the model parameters in the absence of some experimental data (e.g. if one or several samples are contaminated). To demonstrate this capacity, the aqueous samples of the experiments in the CTR are assumed to be contaminated, namely, the concentrations of hydrogen ion and hydrogen peroxide in the aqueous phase are missing.

The measurement error of each measurable variable, which will be used as the threshold to evaluate model prediction accuracy at the submodule F2 (see Sections 5.2.3.5, 5.2.5.4 and 5.2.7.4) and the submodule F3 (see Section 5.2.9.1), has to be defined here.

As described in Section 5.1.2, for the experiment, in which more than one sample is taken, the samples (except for the last one) are only quickly cooled (to reduce the reaction rates), but not immediately quenched (to stop the chemical reactions). Therefore, the measurement error of the *delay-time* concentration-type measurable variable, which is greater than that of the corresponding *real-time* concentration-type measurable variable, is set as 100 mol/m^3; the measurement error of the *real-time* temperature-type measurable variable is set as 2 K, which is greater than the usual value, since the model does not include the heat-loss term.

5.2.3.2 Submodule E1: Specify constraints on model parameters

The model parameters are subject not only to their bounds (shown in Table 5.16) but also to the equality and inequality constraints among them (described in Table 5.17). These constraints are established on the basis of the works of Santacesaria et al. (2011), Zheng et al. (2016) and De Haro et al. (2016).

Table 5.16: Bounds on model parameters for the case study 2.

N° involved item	Parameter	Lower bound	Upper bound
1	$\log(k_{1,Tmin})$, $\log(k_{1,Tmax})$	-9.5	-6.5
	$\log(\Delta h_{1,Tref})$	3.2	4.2
2	$\log(k_{2,Tmin})$, $\log(k_{2,Tmax})$	-9.5	-6.5
	$\log(\Delta h_{2,Tref})$	3.2	4.2
6	$\log(k_{6,Tmin})$, $\log(k_{6,Tmax})$	-4.1	-2.6
	$\log(\Delta h_{6,Tref})$	5.0	5.8
8	$\log(k_{8,Tmin})$, $\log(k_{8,Tmax})$	-6.0	-3.0
	$\log(\Delta h_{8,Tref})$	4.7	5.7
9	$\log(k_{9,Tmin})$, $\log(k_{9,Tmax})$	-10.0	-7.0
	$\log(\Delta h_{9,Tref})$	4.0	5.0
10	$\log(k_{10,Tmin})$, $\log(k_{10,Tmax})$	-10.0	-7.0
	$\log(\Delta h_{10,Tref})$	4.0	5.0
11	$\log(k_{11,Tmin})$, $\log(k_{11,Tmax})$	-10.0	-7.0
	$\log(\Delta h_{11,Tref})$	4.0	5.0
12	$\log(k_{12,Tmin})$, $\log(k_{12,Tmax})$	-10.0	-7.0
	$\log(\Delta h_{12,Tref})$	4.0	5.0
21	$\log(m_{FA,I/II})$	2.1	2.6
	$\log(K_{FA,I,SBSTR})$, $\log(K_{FA,I,CTR})$	-1.0	3.0
23	$\log(m_{PFA,I/II})$	1.1	1.4
	$\log(K_{PFA,I,SBSTR})$, $\log(K_{PFA,I,CTR})$	-1.0	3.0

Table 5.17: Equality and inequality constraints among model parameters for the case study 2.

N°	Constraint	Description
1	Eq. (5.21)	The constraints are constructed in order that the estimates of A_j and Ea_j, derived from $\log k_{j,T_{max}}$ and $\log k_{j,T_{min}}$ according to Eqs. (2.16)-(2.17), are reasonable, i.e. positive and not too big. Note that j is not equal to 1 or 2, since the activation energies of the reactions N°1 and N°2 can be negative (De Haro et al., 2016).
2	Eq. (5.22)	
3	Eq. (5.23)	The absolute values of the enthalpy of the exothermic FA perhydrolysis and that of the endothermic PFA hydrolysis (i.e. the reverse reaction of the former one) are identical. Note that the enthalpy sign, representing the reaction type, i.e. exothermic or endothermic, has been specified at the submodule B1 (see Section 5.2.1.5)
4	Eq. (5.24)	The determining step of the epoxide ring-opening reactions is the formation of the carbocation. Therefore, the reaction rate constants and the reaction enthalpies of the epoxide ring-opening reactions are assumed to be identical and equal to those of the formation of the carbocation.
5	Eq. (5.25)	
6	Eq. (5.26)	

$$\log k_{j,T_{max}} - \log k_{j,T_{min}} > 0.01 \quad (5.21)$$

$$\log k_{j,T_{max}} - \log k_{j,T_{min}} < 1.50 \quad (5.22)$$

$$\log \Delta h_{1,T_{ref}} = \log \Delta h_{2,T_{ref}} \quad (5.23)$$

$$\log k_{9,T_{min}} = \log k_{10,T_{min}} = \log k_{11,T_{min}} = \log k_{12,T_{min}} \quad (5.24)$$

$$\log k_{9,T_{max}} = \log k_{10,T_{max}} = \log k_{11,T_{max}} = \log k_{12,T_{max}} \quad (5.25)$$

$$\log \Delta h_{9,T_{ref}} = \log \Delta h_{10,T_{ref}} = \log \Delta h_{11,T_{ref}} = \log \Delta h_{12,T_{ref}} \quad (5.26)$$

5.2.3.3 Module E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs.

Table 5.18: Evolution of parameter estimations for the case study 2.

N° parameter estimation	1	2	3
Strategy application stage	Stage III	Stage V	Stage VII
Data used for parameter estimation	PEs	PEs + 1st DAEPPI	PEs + 1st DAEPPI + 2nd DAEPPI
Model adequacy index [a]	28.7852	28.5416	28.1609
Model accuracy index [b]	9.7083	0.4180	0.1276
$\log(k_{1,Tmin})$	-7.9060 ± 3.8310	-7.7737 ± 0.2303	-7.7505 ± 0.0464
$\log(k_{1,Tmax})$	-7.3575 ± 0.5072	-7.3421 ± 0.0700	-7.2483 ± 0.0674
$\log(\Delta h_{1,Tref})$	3.7094 ± 4.3069	3.4631 ± 1.8592	3.5175 ± 0.3899
$\log(k_{2,Tmin})$	-6.9635 ± 4.9772	-6.6767 ± 0.1654	-6.8222 ± 0.0454
$\log(k_{2,Tmax})$	-8.1050 ± 0.6150	-7.8126 ± 0.1181	-8.0717 ± 0.0755
$\log(\Delta h_{2,Tref})$	3.7094 ± 4.3069	3.4631 ± 1.8592	3.5175 ± 0.3899
$\log(k_{8,Tmin})$	-5.2430 ± 0.4341	-5.1258 ± 0.1268	-5.0970 ± 0.0124
$\log(k_{8,Tmax})$	-5.0899 ± 0.3137	-4.8985 ± 0.1490	-5.0424 ± 0.0360
$\log(\Delta h_{8,Tref})$	5.1024 ± 0.3524	4.8246 ± 0.1806	4.9886 ± 0.0493
$\log(k_{12,Tmin})$	-9.5195 ± 19.8284	-9.4148 ± 2.9591	-9.5819 ± 8.8917
$\log(k_{12,Tmax})$	-9.4054 ± 10.3529	-9.1395 ± 2.7154	-9.4095 ± 7.8162
$\log(\Delta h_{12,Tref})$	4.4859 ± 102.6963	4.4215 ± 142.5093	4.5491 ± 120.2308
$\log(m_{FA,I/II})$	2.1958 ± 0.1538	1.9386 ± 0.0416	2.0261 ± 0.0532
$\log(K_{FA,I,SBSTR})$	0.5793 ± 455.8369	0.4269 ± 64.6063	0.4982 ± 3.6920
$\log(K_{FA,I,CTR})$	-0.3758 ± 1.5042×10^3	-0.2003 ± 141.1698	-0.2788 ± 265.9856
$\log(m_{PFA,I/II})$	1.1993 ± 0.2739	1.1401 ± 0.0260	1.3404 ± 0.0185
$\log(K_{PFA,I,SBSTR})$	0.4412 ± 25.7433	0.5401 ± 1.9012	0.4994 ± 1.5243
$\log(K_{PFA,I,CTR})$	1.1534 ± 276.1010	1.0090 ± 53.9905	1.2526 ± 229.7026

[a] Model adequacy index, i.e. parameter estimation criterion calculated using Eq. (2.19), is used to *directly* characterize model adequacy.

[b] Model accuracy index, i.e. average variance of model parameters derived from Eq. (1.7), is used to *indirectly* characterize model accuracy. Practically, in this work, whether model accuracy is satisfactory or not has to be evaluated using model prediction errors.

The estimation result of the parameters involved in the Set of Models Under Test (SMUT), whose model structures are constructed based on the RNUT3, i.e. the best reaction network identified (see Section 5.2.3.4), is shown in the 1st column of Table 5.18, in which the results of parameter estimations performed using different experimental data updated after the 1st Data Acquisition Experiments for Parameter Precision Improvement (DAEPPI) and the 2nd DAEPPI are also included and will be discussed later (see Sections 5.2.5.2 and 5.2.7.2). For more information about the

estimation results of the parameters involved in the SMUTs, whose model structures are constructed based on the other RNUTs, please refer to Table A3.3 in Appendix 3.

Model parameters are obviously inaccurate, since the confidence intervals for some parameters (with gray shading) are very wide. Such parameters can be classified into three categories:

- parameters characterizing the reaction enthalpies,
- parameters characterizing the reaction rate of the ring-opening reaction, which is much less than those of other chemical reactions taken into consideration,
- parameters characterizing the mass-transfer rates, which are much greater than the reaction rates of the chemical reactions taken into consideration.

5.2.3.4 Submodule F1: Model adequacy pre-test

Model adequacy pre-test is firstly implemented for the SMUT with the largest (i.e. the worst) parameter estimation criterion, since if this SMUT passes model adequacy pre-test, the other SMUTs with smaller parameter estimation criteria (i.e. better model adequacy) pass evidently model adequacy pre-test. Table 5.19 shows the estimation criteria of the parameters involved in the SMUTs, whose model structures are constructed based on the initial RNUTs (see Section 5.2.1.6).

Table 5.19: Comparison of parameter estimation criteria.

Reaction network	RNUT1	RNUT2	RNUT3	RNUT4	RNUT5	RNUT6
Parameter estimation criterion	29.2186	29.4016	28.7852	29.3041	28.9732	28.9988

The worst SMUT (i.e. the SMUT, whose model structures are constructed based on the RNUT2 and parameters are the ones estimated from the experimental data of the PEs, denoted as SMUT[RNUT2,PEs]) is proven qualitatively feasible, since the experimental response points are reasonably scattered around or on the corresponding predicted curves for the SBSTR (see Figure 5.7) and the CTR (see Figure 5.8). It can be seen in the results of the SBSTR experiment (see Figure 5.7) at time 1800 s that the end of the feed creates a singular point in the curves predicted by the model. Therefore, the other SMUTs are also proven feasible; the most suitable SMUT, i.e. the SMUT[RNUT3,PEs], is easily identified, since it possesses the smallest (i.e. the best) parameter estimation criterion.

It can be seen that the figure is divided into a 3×3 grid: the 3 rows correspond to the number of measurable variable types, the 3 lines correspond to the maximum number of measurable variables with the same type. For more details, refer to Section 5.2.1.9.

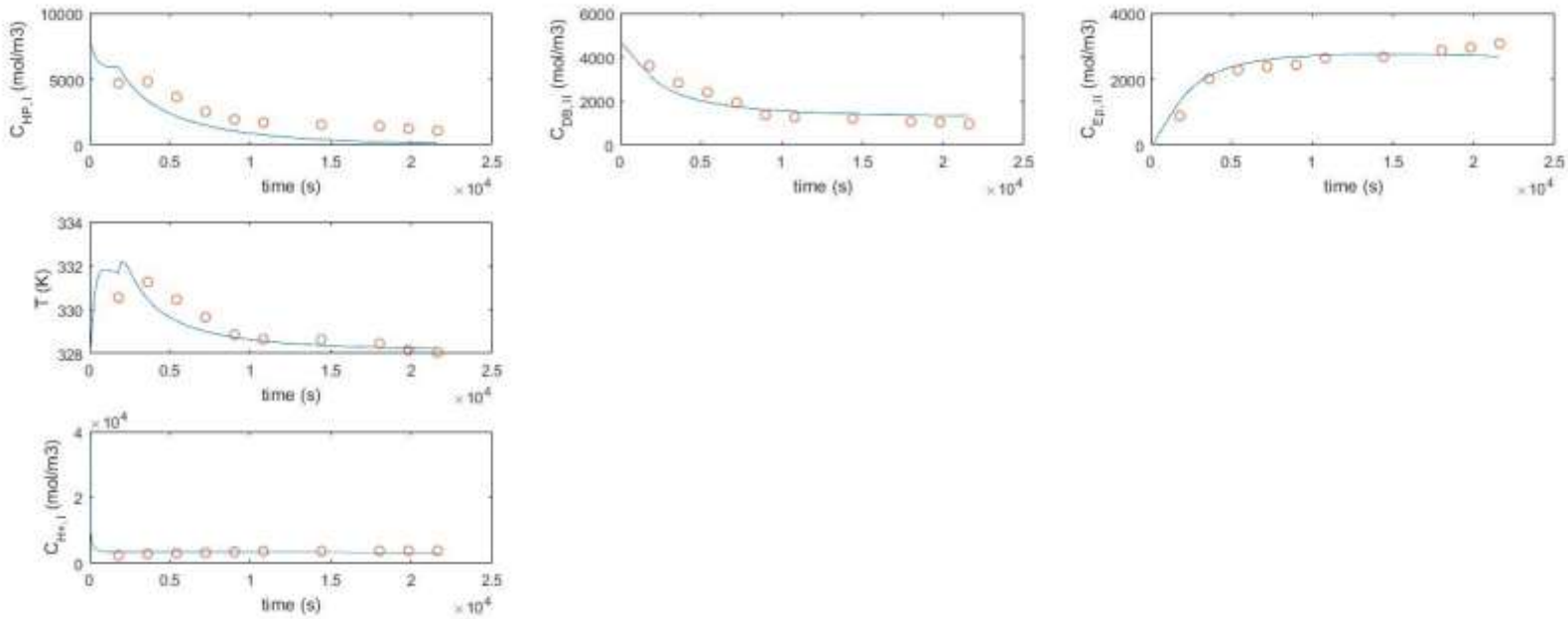


Figure 5.7: Evolution of the various measurable variables for the 4th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 55°C, using the SBSTR (continuous lines: responses predicted by the SMUT[RNUT2,PEs]; singular points: experimental responses).

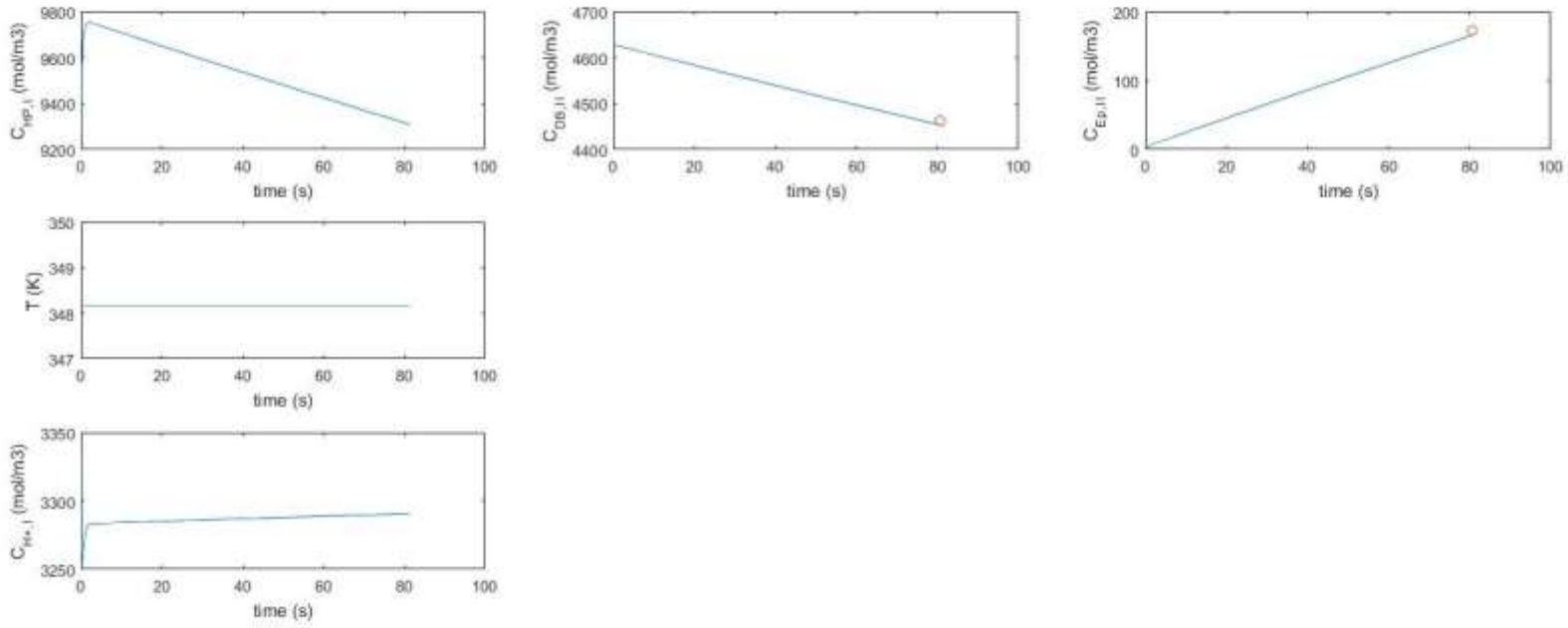


Figure 5.8: Evolution of the various measurable variables for the 3rd sample of the 7th PE carried out at FA-to-DB molar ratio of 0.23, HP-to-DB molar ratio of 0.75, SA concentration of 0.64 wt.% (per oil weight), initial reaction temperature of 75°, using the CTR (continuous lines: responses predicted by the SMUT[RNUT2,PEs]; singular points: experimental responses).

5.2.3.5 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is then performed on the basis of the SMUT[RNUT3,PEs], i.e. the most suitable SMUT identified at the submodule F1 (see Section 5.2.3.4). The result is shown in the 1st column of Table 5.20, in which the results of model-based experimental design for the VEPO performed based on the SMUTs updated after the 1st DAEPPI and the 2nd DAEPPI are also included and will be discussed later (see Sections 5.2.5.4 and 5.2.7.4).

The optimal reactor is the SBSTR, since its criterion for VEPO (0.0983) is much less than that of the CTR (0.4427). The corresponding optimal operating conditions are observed to have the following characteristics:

- the reaction duration does not approach its maximum value allowed (as in the case study 1), since it is expected to be reduced in order that the compromise between reaction time and epoxide yield can be reached;
- the amount of the feed solution 1 (i.e. the mixture of FA solution and HP solution), equal to the product F_1^{in} and t_1^{in}, reaches its maximum value allowed;
- the FA-to-HP molar ratio of the feed solution 1 is near to 1.

Furthermore, efforts have been made to interpret why the software proposed such operating conditions; however, actually, the reasons are not given.

Then, model accuracy pre-test is carried out to investigate whether the predicted error of each measurable variable for the VEPO (shown in the last part of Table 5.20) is less than its corresponding threshold, i.e. 100 mol/m^3 for the concentration-type measurable variable; 2 K for the temperature-type measurable variable.

The SMUT[RNUT3,PEs] does not pass model accuracy pre-test, since there is no measurable variable, whose predicted error is less than the corresponding threshold. Therefore, there are two possible choices for the next step: model structure development (submodule B1 or B2) or design of the DAEPPI (submodule D1). Considering that the SMUT[RNUT3,PEs] passes model adequacy pre-test, we decide to design and perform a DAEPPI and expect that the model parameter precision as well as the model prediction precision will increase after the 1st DAEPPI so that the SMUT[RNUT3,PEs+1st DAEPPI] may pass model accuracy pre-test.

Table 5.20: Results of model-based experimental design for the VEPO.

N° model-based experimental design		1		2		3
(Strategy application stage)		(Stage III)		(Stage V)		(Stage VII)
SMUT used		SMUT1 [a]		SMUT2 [b]		SMUT3 [c]
VEPO	Considered reactor	SBSTR	CTR	SBSTR	CTR	SBSTR
	τ [s]	4.3174×10^{3}		4.0285×10^{3}		4.6514×10^{3}
	T^{int} [K]	355.30		356.68		357.93
	n_{FA}^{int} [mol]	0		0		0
	n_{HP}^{int} [mol]	0		0		0
	n_{SAeq}^{int} [mol]	0.0832		0.0935		0.0601
	n_{PFA}^{int} [mol]	0		0		0
	n_{W}^{int} [mol]	0		0		0
	n_{DB}^{int} [mol]	1.3912		1.3912		1.3912
	T_{I}^{in} [K]	298.15	367.14	298.15	366.36	298.15
	T_{II}^{in} [K]	0	367.14	0	366.36	0
	F_{I}^{in} [$m^3\cdot s^{-1}$]	1.0053×10^{-7}	3.3373×10^{-8}	1.3237×10^{-7}	3.5752×10^{-8}	1.0221×10^{-7}
	F_{II}^{in} [$m^3\cdot s^{-1}$]	0	1.1736×10^{-7}	0	1.1739×10^{-7}	0
	t_{I}^{in} [s]	2.9841×10^{3}		2.2664×10^{3}		2.9349×10^{3}
	t_{II}^{in} [s]	0		0		0
	$c_{FA,I}^{in}$ [$mol\cdot m^{-3}$]	6.4226×10^{3}	9.4495×10^{3}	1.0898×10^{4}	1.1645×10^{4}	8.1729×10^{3}
	$c_{HP,I}^{in}$ [$mol\cdot m^{-3}$]	6.5665×10^{3}	4.3987×10^{3}	3.8927×10^{3}	3.2590×10^{3}	5.4714×10^{3}
	$c_{SAeq,I}^{in}$ [$mol\cdot m^{-3}$]	0	1.0943×10^{3}	0	9.9471×10^{2}	0
	$c_{PFA,I}^{in}$ [$mol\cdot m^{-3}$]	1.1753×10^{3}	1.4706×10^{3}	1.5837×10^{3}	1.5791×10^{3}	1.3821×10^{3}
	$c_{W,I}^{in}$ [$mol\cdot m^{-3}$]	2.8707×10^{4}	2.2612×10^{4}	2.1430×10^{4}	1.9226×10^{4}	2.5902×10^{4}
	$c_{FA,II}^{in}$ [$mol\cdot m^{-3}$]	0	0	0	0	0
	$c_{HP,II}^{in}$ [$mol\cdot m^{-3}$]	0	0	0	0	0
	$c_{PFA,II}^{in}$ [$mol\cdot m^{-3}$]	0	0	0	0	0
	$c_{W,II}^{in}$ [$mol\cdot m^{-3}$]	0	0	0	0	0
	$c_{DB,II}^{in}$ [$mol\cdot m^{-3}$]	0	4.6374×10^{3}	0	4.6374×10^{3}	0
Optimal performances	Criterion for VEPO (Eq. (5.12))	0.0983	0.4427	0.1032	0.4306	0.1102
	τ [s]	4.3174×10^{3}	84.7023	4.0285×10^{3}	83.3699	4.6514×10^{3}
	X_{DB} [%]	95.19	11.54	93.24	13.97	93.96
	Y_{Ep} [%]	99.95	100	99.92	100.00	99.96
	T_{max} [K]	367.16	367.14	367.27	366.36	366.92
Model prediction errors	$C_{HP,I}$ [mol/m^3]	1207.90		258.8225		9.78
	$C_{DB,II}$ [mol/m^3]	586.57		189.63		66.59
	$C_{Ep,II}$ [mol/m^3]	497.24		161.84		42.89
	T [K]	6.60		2.68		1.75
	$C_{H^+,I}$ [mol/m^3]	424.35		30.72		13.63

[a] SMUT1 = SMUT[RNUT3,PEs].

[b] SMUT2 = SMUT[RNUT3,PEs+1st DAEPPI].

[c] SMUT3 = SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI].

5.2.3.6 Submodule D1: Design data acquisition experiment

Firstly, the number of experiments (i.e. samples) contained in one DAEPPI has to be defined. For this case study, it is artificially set to 2 to make a compromise between information quantity and efficiency of data acquisition. Then, the model-based experimental design for the 1st DAEPPI is performed, the result is shown in the 1st column of Table 5.21, in which the result of the model-based experimental design for the 2nd DAEPPI is also included and will be discussed later (see Section 5.2.5.5).

Table 5.21: Results of model-based experimental design for the DAEPPI.

N° model-based experimental design (Strategy application stage)	1 (Stage III)		2 (Stage V)	
SMUT used	SMUT[RNUT3,PEs]		SMUT[RNUT3,PEs+1st DAEPPI]	
Criterion for DAEPPI (see Eq. (2.20))	0.5427		0.1266	
Optimal reactor	SBSTR		SBSTR	
N° experiment	1	2	1	2
τ [s]	914.2	1347.2	1265.1	15886
T^{int} [K]	300.8492		323.6448	
n_{FA}^{int} [mol]	0		0	
n_{HP}^{int} [mol]	0		0	
n_{SAeq}^{int} [mol]	0.0316		0.0716	
n_{PFA}^{int} [mol]	0		0	
n_{W}^{int} [mol]	0		0	
n_{DB}^{int} [mol]	1.3912		1.3912	
T_{I}^{in} [K]	298.1500		298.1500	
T_{II}^{in} [K]	0		0	
F_{I}^{in} [$m^3 \cdot s^{-1}$]	2.6181×10^{-7}		2.2102×10^{-7}	
F_{II}^{in} [$m^3 \cdot s^{-1}$]	0		0	
t_{I}^{in} [s]	914.1585		666.6934	
t_{II}^{in} [s]	0		0	
$c_{FA,I}^{in}$ [$mol \cdot m^{-3}$]	7.7807×10^{3}		2.8761×10^{3}	
$c_{HP,I}^{in}$ [$mol \cdot m^{-3}$]	5.6935×10^{3}		9.0712×10^{3}	
$c_{SAeq,I}^{in}$ [$mol \cdot m^{-3}$]	0		0	
$c_{PFA,I}^{in}$ [$mol \cdot m^{-3}$]	1.3391×10^{3}		601.7133	
$c_{W,I}^{in}$ [$mol \cdot m^{-3}$]	2.6464×10^{4}		3.4687×10^{4}	
$c_{FA,II}^{in}$ [$mol \cdot m^{-3}$]	0		0	
$c_{HP,II}^{in}$ [$mol \cdot m^{-3}$]	0		0	
$c_{PFA,II}^{in}$ [$mol \cdot m^{-3}$]	0		0	
$c_{W,II}^{in}$ [$mol \cdot m^{-3}$]	0		0	
$c_{DB,II}^{in}$ [$mol \cdot m^{-3}$]	0		0	

The optimal reactor is the SBSTR, and the corresponding optimal operating conditions are observed to have the following characteristics: (i) both sampling times are less than the 1st sampling time of PEs (i.e. 1800 s); (ii) the high concentration perhydrolysis reactant solution is quickly added into the reactor. At such operating conditions, the parameters characterizing the reaction rates of the reactions N°1, 2, 8 are more sensitive than the other parameters, since these reactions, particularly the reaction N°8, dominate at the initial reaction stage.

For the experiments carried out in the SBSTR, the sampling time is a very important operating condition, influencing so much the data acquisition efficiency. According to the model-based simulation result (illustrated for instance by Figure 5.7), there are actually at most four time periods during an experiment (shown in Table 5.22). However, in the previous experiments, there are no samples taken at the III-period. To maximize the acquired information:

1. the 1st sample appears to be designed to be taken at the end of feeding, corresponding to the most special moment within the whole process;
2. the 2nd sample is designed to be taken at the III-period, which can be confirmed by Figure 5.9. In order to ensure that the III-period will certainly be present, the solution I is charged into the reactor by pumping at a high flow rate (F_I^{in} = 2.6181×10^{-7} m^3·s^{-1}) within a short feed duration (t_I^{in} = 914.2 s).

Table 5.22: Various time periods of the experiment carried out in the SBSTR.

N° period [a]	With feed?	Period certainly present?	Reaction temperature trend [b]
I	Yes	Yes	Increase
II		No	Decrease
III	No	No	Increase
IV		Yes	Decrease

[a] Periods are arranged in order of chronological appearance.

[b] According to Eq. (A1.16), whether the reaction temperature increases or not, depends on three factors: (i) the rate of heat generation due to the reactions, denoted as q_1, which is always positive, (ii) heat-inflow rate, denoted as q_2, which is negative during the feeding period, and equal to 0 after feeding, since the temperature of feed 1, which is fixed at the room temperature, i.e. 25 °C, is always less than the reaction temperature, (iii) heat-production rate exchanged, denoted as q_3, which is always negative, since the temperature of fluid flowing in the jacket, which is fixed at the initial reaction temperature, is always lower than the reaction temperature. When $q_1 > q_2 + q_3$, the reaction temperature increases, otherwise, the reaction temperature decreases.

5.2.4 Stage IV: First DAEPPI

The 1st DAEPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at the submodule D1 (see Section 5.2.3.6). In order to minimize the measurement error of the concentration-type measurable variable (see Section 5.2.3.1),

for this experiment as well as the following experiments, including the 2nd DAEPPI and the VEPO, each withdrawn sample is immediately quenched.

5.2.5 Stage V: Software application

5.2.5.1 Submodule D2: Input experimental data

The experimental responses of the 1st DAEPPI (shown in Table A3.2 of Appendix 3) are input into the software.

5.2.5.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs and the 1st DAEPPI. The result is shown in 2nd column of Table 5.18. The current SMUT is updated to SMUT[RNUT3,PEs+1st DAEPPI]. It can be noticed that:

1. Model parameter precision can be considered to significantly increase, since the average variance of model parameters (denoted as 'model accuracy index' in Table 5.18) of the SMUT[RNUT3,PEs+1st DAEPPI] decreases by 95.69% in comparison to that of the SMUT[RNUT3,PEs].
2. The unique parameter, whose confidence interval increases, is $\log(\Delta h_{12,Tref})$, i.e. the parameter characterizing the enthalpy of the ring-opening reaction. This can be explained by the fact that the information for improving the confidence interval of this parameter is not acquired through the 1st DAEPPI. In the in-situ epoxidation reaction system, the epoxidation dominates for the whole process, particularly at the initial reaction stage. Therefore, the influence of this reaction on the reaction system (temperature) can be neglected with respect to the exothermic epoxidation to the reaction system (temperature).
3. Even if the 1st DAEPPI does not provide any CTR-related experimental information, the confidence intervals of $\log(K_{FA,I,CTR})$ and $\log(K_{FA,I,CTR})$, i.e. the parameters characterizing the mass-transfer rates for the CTR process, also decrease, which can be due to the interactions among the considered items, including the chemical reactions and the mass-transfers, as well as the significant decrease of the confidence intervals of the parameters characterizing the chemical reactions.

5.2.5.3 Submodule F1: Model adequacy pre-test

The SMUT[RNUT3,PEs] passes model adequacy pre-test, but that does not mean that SMUT[RNUT3,PEs+1st DAEPPI], whose reaction network is the same as the SMUT[RNUT3,PEs], passes also the test. Therefore, model adequacy pre-test should be carried out for the SMUT

[RNUT3,PEs+1st DAEPPI]. Here, the SMUT[RNUT3,PEs+1st DAEPPI] passes model adequacy pre-test, since its parameter estimation criterion (28.5416) is smaller than that of SMUT[RNUT3,PEs] (28.7852) (see Table 5.18), which can be also confirmed by Figure 5.9.

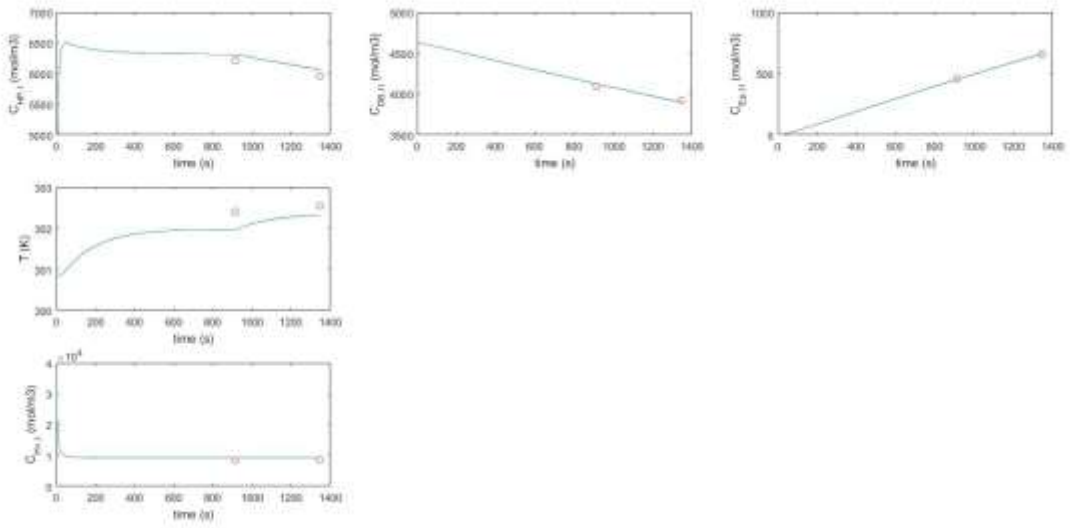


Figure 5.9: Evolution of the various measurable variables for the 1st DAEPPI (continuous lines: responses predicted by the SMUT[RNUT3,PEs+1st DAEPPI]; singular points: experimental responses).

5.2.5.4 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is performed on the basis of the SMUT[RNUT3,PEs+1st DAEPPI]. The result is shown in the 2nd column of Table 5.20. In comparison to the 1st predicted VEPO, the 2nd one is also carried out in the SBSTR and possesses similar operating conditions (e.g. the reaction duration does not approach its maximum allowed; the amount of the feed solution 1 reaches its maximum allowed), but significantly differs in the FA-to-HP molar ratio of the feed. Such a difference may be due to the difference between (the parameters of) the SMUTs used, even if the SMUTs are constructed based on the same RNUT.

Model accuracy pre-test fails again, since the model prediction errors of $c_{HP,I}$ (258.8225 mol/m^3), $c_{DB,II}$ (189.6333 mol/m^3) and $c_{Ep,II}$ (161.8377 mol/m^3) are still larger than the threshold (100 mol/m^3). In view of the fact that the predicted error of each measurable variable decreases after the 1st DAEPPI, we decide to do one more DAEPPI and expect that the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI] will pass model accuracy pre-test.

5.2.5.5 Submodule D1: Design data acquisition experiment

Model-based experimental design for the 2nd DAEPPI is performed for only the SBSTR due to the following two reasons: (i) the experiment in CTR is carried out under isothermal mode, therefore,

no information for improving the parameters characterizing the reaction enthalpies, such as $\log(\Delta h_{12,Tref})$, can be provided; (ii) according to the first two model-based experimental design results for the VEPO, the SBSTR is always the optimal reactor. The result (shown in 2nd column of Table 5.21) indicates that the samples of the 2nd DAEPPI have to be taken at relatively long experiment durations. At such operating conditions, the chemical reaction rates decrease due to the decrease of the concentration of the peroxidant, i.e. the HP (note that the couple FA/PFA is considered as the HP-carrier). Therefore, the parameters characterizing the mass-transfer rates become more sensitive, and their confidence intervals are expected to be reduced after the 2nd DAEPPI. This will be confirmed later by the fact that the confidence interval of $\log(K_{FA,I,SBSTR})$ will decrease from 64.6063 to 3.6920 (i.e. the greatest drop), and that of $\log(K_{PFA,I,SBSTR})$ will decrease from 1.9012 to 1.5243.

5.2.6 Stage VI: Second DAEPPI

The 2nd DAEPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at D1 (see Section 5.2.5.5).

5.2.7 Stage VII: Software application

5.2.7.1 Submodule D2: Input experimental data

The experimental responses of the 2nd DAEPPI (shown in Table A3.2) are input into the software.

5.2.7.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs, the 1st DAEPPI and the 2nd DAEPPI. The current SMUT is updated to SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI]. It has been noticed that the average variance of model parameters decreases by 98.69% in comparison to the SMUT[RNUT3,PEs] ('model accuracy index' in Table 5.18). However, there are also some inaccurate parameters (see the 3rd column of Table 5.18). But this is normal, since the DAEPPI is designed to reduce the average variance of model parameters, not the variance of each parameter. Integration of more criteria for experimental design for parameter precision into the software, which will allow the users to select a reasonable one according to the parameter estimation result, seems necessary and will be discussed later (see Section 5.4).

5.2.7.3 Submodule F1: Model adequacy pre-test

The SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI] passes model adequacy pre-test, since its parameter estimation criterion (28.1609) is smaller than the previous one (28.5416) (see Table 5.18).

5.2.7.4 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is performed on the basis of the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI]. The result is shown in the 5th column of Table 5.20. The predicted error of each measurable variable is less than its corresponding threshold, therefore, the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI] passes model accuracy pre-test. The next step is to perform the VEPO experimentally.

5.2.8 Stage VIII: VEPO

The VEPO is performed according to the practical operating conditions derived from the operating conditions in general form obtained at the submodule F2 (see Section 5.2.7.4).

5.2.9 Stage IX: Software application

5.2.9.1 Submodule F3: Model adequacy and accuracy post-tests

Figure 5.10 presents the comparison between the experimental responses of the VEPO (shown in Table A3.2 of Appendix 3) and those predicted by the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI]. The SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI] seems to be able to predict well all the measurable variables. However, the predicted $c_{Ep,II}$ is out of the corresponding measurement interval. Therefore, the SMUT[RNUT3,PE+1st DAEPPI+2nd DAEPPI] is not validated.

It can be noticed that the predicted $c_{Ep,II}$ is greater than the corresponding measured value. This can be probably due to the fact that the estimated values of the parameters characterizing the reaction rate of the ring-opening reaction are too small, leading to the fact that the calculated reaction rate of the ring-opening reaction can be neglected with respect to those of other chemical reactions, which can be illustrated by Table 5.23.

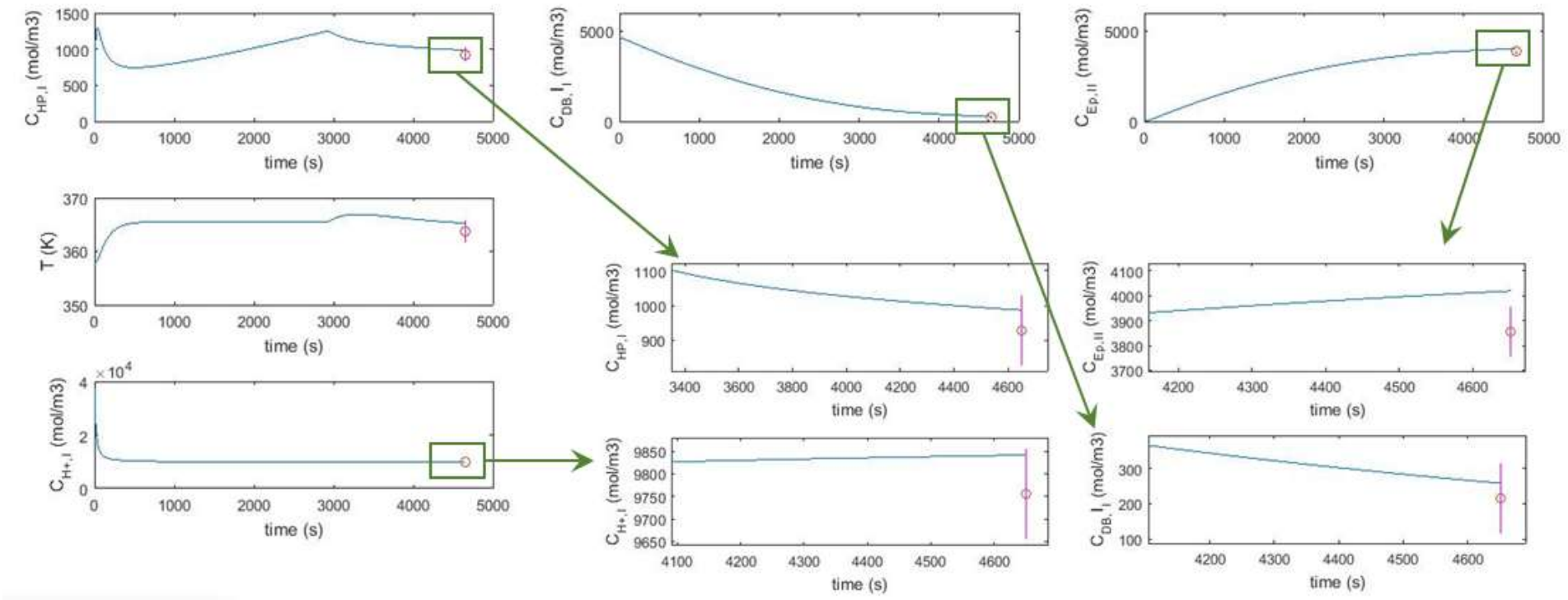


Figure 5.10: Experimental responses for the VEPO using the SBSTR (continuous lines: responses predicted by the SMUT[RNUT3,PEs+1st DAEPPI+2nd DAEPPI]; singular points: experimental responses; vertical bars around the singular points: experimental uncertainties).

In order to obtain reasonable values of such parameters, the user can follow two suggestions:

1. Define the lower and upper bounds of the parameters characterizing the reaction rate of the ring-opening reaction, i.e. $\log(k_{12,Tmin})$ and $\log(k_{12,Tmax})$, as the same and equal to the reasonable values issue from other researches (submodule E1), namely, pre-assign the reasonable values to $\log(k_{12,Tmin})$ and $\log(k_{12,Tmax})$, then re-estimate the model parameters (submodule E2) and follow the instructions proposed by the software.
2. Study independently the ring-opening reaction system by referring to the researches of Campanella and Baltanas (2006), Zheng et al. (2016), Cai et al. (2018).

5.3 Strategy feasibility analysis

Table 5.23 shows the comparison of kinetic constants obtained in this work and in the literature. The reaction enthalpies in the research of Santacesaria et al. (2011) are pre-assigned, not estimated, therefore, the estimated reaction enthalpies in the research of Zheng et al. (2016) are used for the comparison.

Table 5.23: Comparison of kinetic constants obtained from this work and from literatures.

Kinetic constant	This work	Santacesaria et al. (2011)	Zheng et al. (2016)
$k_{1,70°C}$ [$m^6 \cdot mol^{-2} \cdot s^{-1}$]	0.78×10^{-8}	1.09×10^{-8}	
$k_{2,70°C}$ [$m^6 \cdot mol^{-2} \cdot s^{-1}$]	1.22×10^{-11}	2.11×10^{-9}	
$k_{8,70°C}$ [$m^3 \cdot mol^{-1} \cdot s^{-1}$]	1.21×10^{-5}	3.33×10^{-5}	
$k_{12,70°C}$ [$m^3 \cdot mol^{-1} \cdot s^{-1}$]	0.96×10^{-9}	1.33×10^{-8}	
$\Delta h_{1,66.85°C}$ [$J \cdot mol^{-1}$]	-3.29×10^{3}		-5.58×10^{3}
$\Delta h_{2,,66.85°C}$ [$J \cdot mol^{-1}$]	3.29×10^{3}		5.58×10^{3}
$\Delta h_{8,66.85°C}$ [$J \cdot mol^{-1}$]	-0.97×10^{5}		-2.30×10^{5}
$\Delta h_{12,66.85°C}$ [$J \cdot mol^{-1}$]	-3.54×10^{4}		-9.00×10^{4}
$m_{FA,I/II}$ [-]	106.19	177.78	
$m_{PFA,I/II}$ [-]	21.90	17.02	
$K_{FA,I,SBSTR}$ [min^{-1}]	188.95	250.00	
$K_{PFA,I,SBSTR}$ [min^{-1}]	189.47	250.00	

The kinetic constants we determined (derived from the parameter actually estimated) are of the same orders of magnitude as those obtained from the literatures, except:

- k_2, the reaction rate constants of the acid perhydrolysis of PFA is much less than the corresponding research value. According to the research of De Haro et al. (2016), the acid-perhydrolysis activation energy can be negative, therefore, the fact that the reaction rate constant of this reaction (significantly) decreases with increasing temperature, is also normal.

- k_{12}, the reaction rate constants of the ring-opening reaction is also much less than the corresponding research value. This can be probably due to: (i) the complex interactions among the considered items, including the chemical reactions and the mass transfers, (ii) the fact that even if the estimates of the parameters characterizing the ring-opening reaction with the relatively low reaction rate are much less than the theoretical values (i.e. the other research values), a satisfactory parameter estimation criterion can be obtained. Therefore, the 2nd suggestion proposed in Section 5.2.9.1, i.e. the independent study of the ring-opening reaction system, seems to be necessary.

Even if the unneglectable deviation between the estimate and the corresponding reference value is observed for two parameters, in the view of the fact that all the other parameters are well estimated, the feasibility of the strategy is also highlighted.

5.4 Software practicability analysis

The initial version of the Strategy-Based Software used for Model Development (hereinafter referred to as the SBS_MD V1.0) is developed and generalized on the basis of the case study 1. Some ill-considered aspects are found while applying it to the case study 2, a more complex case study in comparison to the case study 1. The application of the SBS_MD V1.0 to the case study 2 is feasible. Nevertheless, necessary modifications and additions of the codes at the back-end are required. Therefore, the updates of the SBS_MD V1.0 appear necessary; improvements can be brought to the following submodules:

- Submodule A4: introduce a function allowing the user to customize the performance, which is out of the suggestion list;
- Submodule B2: introduce a function allowing the user to customize the function associating the concentration of catalyst anion or ion and the concentration(s) of species in molecular form;
- Submodule C1: introduce a function allowing the user to customize the constraints on operating conditions;
- Submodule D1: integrate more criteria for parameter precision, such as A-optimality, D-optimality and E-optimality, which allows the user to select a reasonable criterion according to his need. Let us imagine that if this function had been implemented in the SBS_MD V1.0, after the 1st DAEPPI, the confidence interval of $\log(\Delta h_{12,Tref})$ would increase and become the greatest (see Section 5.2.5.2), A-optimality aiming at reducing the greatest confidence interval, could have been used for designing the 2nd DAEPPI;

- Submodule E1: introduce a function allowing the user to customize the constraints on model parameters.

5.5 Conclusion

In this chapter, in order to develop an adequate and accurate SMUT for the case study 2, i.e. epoxidation of sunflower oil by performic acid generated in situ, a series of experiments for different purposes have been performed:

- 11 PEs designed by the user for providing the basic information for the initial model development, as well as allowing to identify the best reaction network,
- 2 DAEPPIs designed by the strategy-based software for increasing model accuracy,
- 1 VEPO designed also by the strategy-based software for verifying the prediction quality.

Even if an expected adequate and accurate SMUT is not finally developed, the concrete application process demonstrates also the feasibility of the strategy and the practicability of the software. Furthermore, in order that the software can be applied to more projects, some improvement suggestions allowing the users to customize their needs, are proposed.

References

De Haro, J.C., Izarra, I., Rodríguez, J.F., Pérez, Á., Carmona, M., 2016. Modelling the epoxidation reaction of grape seed oil by peracetic acid. Journal of cleaner production, 138, 70-76.

Cai, X., Zheng, J.L., Aguilera, A.F., Vernières-Hassimi, L., Tolvanen, P., Salmi, T., Leveneur, S., 2018. Influence of ring-opening reactions on the kinetics of cottonseed oil epoxidation. International Journal of Chemical Kinetics, 50(10), 726-741.

Campanella, A., Baltanás, M.A., 2006. Degradation of the oxirane ring of epoxidized vegetable oils in liquid–liquid heterogeneous reaction systems. Chemical Engineering Journal, 118(3), 141-152.

Filippis, P.D., Scarsella, M., Verdone, N., 2009. Peroxyformic acid formation: a kinetic study. Industrial & Engineering Chemistry Research, 48(3), 1372-1375.

Greenspan, F.P., Mackellar, D.G., 1948. Analysis of aliphatic per acids. Analytical Chemistry, 20(11), 1061-1063.

Kousaalya, A.B., Beyene, S.D., Gopal, V., Ayalew, B., Pilla, S., 2018. Green epoxy synthesized from Perilla frutescens: A study on epoxidation and oxirane cleavage kinetics of high-linolenic oil. Industrial Crops and Products, 123, 25-34.

Maerker, G., 1965. Determination of oxirane content of derivatives of fats. Journal of the American Oil Chemists' Society, 42(4), 329-332.

Moreno, V.C., Russo, V., Tesser, R., Di Serio, M., Salzano, E., 2017. Thermal risk in semi-batch reactors: The epoxidation of soybean oil. Process Safety and Environmental Protection, 109, 529-537.

Paquot, C., 2013. Standard methods for the analysis of oils, fats and derivatives. Elsevier.

Santacesaria, E., Tesser, R., Di Serio, M., Turco, R., Russo, V., Verde, D., 2011. A biphasic model describing soybean oil epoxidation with H_2O_2 in a fed-batch reactor. Chemical Engineering Journal, 173(1), 198-209.

Santacesaria, E., Renken, A., Russo, V., Turco, R., Tesser, R., Di Serio, M., 2012. Biphasic model describing soybean oil epoxidation with H_2O_2 in continuous reactors. Industrial & Engineering Chemistry Research, 51(26), 8760-8767.

Sun, X., Zhao, X., Du, W., Liu, D., 2011. Kinetics of formic acid-autocatalyzed preparation of performic acid in aqueous phase. Chinese Journal of Chemical Engineering, 19(6), 964-971.

Wu, Z., Nie, Y., Chen, W., Wu, L., Chen, P., Lu, M., Ji, J., 2016. Mass transfer and reaction kinetics of soybean oil epoxidation in a formic acid-autocatalyzed reaction system. The Canadian Journal of Chemical Engineering, 94(8), 1576-1582.

Zheng, J.L., Wärnå, J., Salmi, T., Burel, F., Taouk, B., Leveneur, S., 2016. Kinetic modeling strategy for an exothermic multiphase reactor system: Application to vegetable oils epoxidation using Prileschajew method. AIChE Journal, 62(3), 726-741.

Conclusion and Perspectives

Conclusion

Given a chemical synthesis and several available reactors, how to identify efficiently the optimal reactor with its associated optimal operating conditions, is always a problem, since there is no unified approach as a reference; exploratory approaches certainly will be time consuming and experimentally expensive. Therefore, a unified approach to tackle this issue is necessary.

To reach that goal, in this work, a model-development strategy based on model-based experimental design is developed. It can be used to guide the user to develop the adequate and accurate models for all considered reactors for a given chemical synthesis, and further to identify the optimal reactor with its associated optimal operating conditions.

The proposed strategy is methodologically systematic, since it provides not only the unified procedures for model development and evaluation, but also the solutions to solve all possible problems met following model evaluation in terms of model adequacy and accuracy. Furthermore, it is also practically systematic due to the following three reasons:

1. Various reactor models can be simultaneously developed. In addition, for model identification, validation and refining, model-based experimental design is performed by taking into consideration all the available reactors.
2. Model structure development is based on the reaction proposed within a reaction supernetwork containing all feasible chemical reactions and mass transfers: that allows the strategy to be applied not only for homogeneous reaction systems but also for heterogeneous reaction systems.
3. During the practical application of such a strategy, the user has not only to perform the experiments, but also to write the codes to solve the optimization problems, such as model parameter estimation and model-based experimental design. Therefore, in order to reduce the user works (i.e. facilitate the application of the strategy), a software, providing user-friendly interfaces and integrating model development, model evaluation, model-based experimental design for model refining and performance optimization, is developed using MATLAB R2014a. It should be noted that the application range of the initial version of the strategy-based software is limited in comparison to that of the strategy. Specifically, the most common reaction systems for syntheses of fine and pharmaceutical chemicals, namely, liquid-phase reaction systems, including liquid, liquid-liquid and liquid-liquid-liquid reaction systems, are taken into consideration; reactor configuration should be either stirred-tank reactor or tubular reactor.

In order to validate experimentally the feasibilities of the developed strategy and strategy-based software, considering the available reactors and analysis instruments in the laboratory, two chemical syntheses, related to the valorization of sunflower oil, namely, NaOH-catalyzed ethanolysis of sunflower oil and epoxidation of sunflower oil by performic acid generated in situ, have been selected as validation case studies.

The feasibility of the strategy and the practicability of the software have been firstly demonstrated by the ethanolysis case study. An adequate and accurate Set of Models Under Test (SMUT), whose model prediction quality has been verified experimentally, has been developed after a series of experiments for different purposes, including: (i) five Preliminary Experiments (PEs) designed by the user on the basis of his expertise, providing the basic information for the initial model development, as well as allowing to identify the best reaction network; (ii) two Data Acquisition Experiments for Parameter Precision Improvement (DAEPPIs) designed by the strategy-based software; (iii) one Validation Experiment for Performance Optimization (VEPO) designed also by the strategy-based software. The kinetic parameters (i.e. the activation energies and pre-exponential factors) we determined (derived from the parameter actually estimated) are of the same orders of magnitude as those obtained from literature, except the pre-exponential factors for two reactions. Furthermore, the parameter actually estimated, used to calculate the pre-exponential factors mentioned above, are reasonable, since they are comprised in their corresponding confidence intervals derived from literature.

The similar chronological application of the strategy to the in-situ epoxidation case study has then been performed. The corresponding designed experimental strategy was composed of eleven PEs, two DAEPPIs and one VEPO. An expected SMUT is not finally developed, since in the VEPO, the measured value of one measurable variable (i.e. the concentration of epoxide in the organic phase) is out of its corresponding predicted confidence interval. This can be due to the fact that the estimates of the parameters characterizing the ring-opening reaction with the relatively low reaction rate are much less than the theoretical values (i.e. the other research values). Even so, in the view of the fact that all the other parameters are well estimated, the feasibility of the strategy is also highlighted.

Perspectives

The feasibility and the generality of the strategy have been experimentally validated by two liquid-phase synthesis case studies with the help of the strategy-based software. The future works should focus on further validating them by more reaction systems with different characteristics, e.g. the gas–solid catalytic reaction system in fixed-bed reactor. For that, it is necessary to expand the software

database by integrating the general model structures for more reactors. Hence, once the reaction system description is proposed by the users, the specific model structure for each considered reactor can be automatically generated. Furthermore, in order to be more practical for the users, the software can be also updated by implementing the following functions:

- The promising functions, which have not been implemented during the development of the initial version of the software, including (i) development of the reaction supernetwork through theoretical method by optimizer, (ii) identification of the most suitable reaction network through theoretical method by optimizer, (iii) model structure analysis, (iv) quantitative evaluation of model adequacy, (v) model-based experimental design for model discrimination;
- The functions allowing the users to customize their needs, for example, to customize the performance, which is out of the suggestion list, to personalize the linear or nonlinear equality or inequality constraints between the model parameters or operating conditions in general form.
- The functions making the software more ergonomic, such as arbitrary switch between interfaces, data input by loading a data-containing excel file.

Appendix 1: General mass, volume and heat balances

The deduction process of Eqs. (2.12)-(2.14), i.e. the general model structure for any homogeneous catalytic reaction system in an ideal stirred-tank reactor, from the corresponding general mass, volume and heat balances, is given in this paragraph.

General mass balance

The general mass balance for the species i in the phase k can be represented as:

$$F_{i,k}^{in} + F_{i,k}^{p} = F_{i,k}^{out} + F_{i,k}^{acc} \quad (A1.1)$$

where:

- $F_{i,k}^{in}$ and $F_{i,k}^{out}$ are respectively the molar inflow and outflow rates of the species i in the phase k,
- $F_{i,k}^{p}$ is the molar production rate of the species i in the phase k, and can be represented by Eq. (A1.2), in which the terms on the right represent the molar production rates due to respectively the chemical reactions and the mass transfers,

$$F_{i,k}^{p} = V_k \sum_{j}^{n_{r_k}} \nu_{i,j,k} r_{j,k} - V_k \sum_{l}^{n_p - 1} J_{i,k \to l \neq k} \quad (A1.2)$$

  where V_k is the volume of the phase k, n_{r_k} is the number of chemical reactions occurring in the phase k, $\nu_{i,j,k}$ is the stoichiometric coefficient of the species i for the chemical reaction j in the phase k, $r_{j,k}$ is the reaction rate of the chemical reaction j in the phase k, n_p is the number of phases, $J_{i,k \to l \neq k}$ is mass-transfer rate of the species i from the phase k to the phase l,

- $F_{i,k}^{acc}$ is the molar accumulation rate of the species i in the phase k, and can be represented as:

$$F_{i,k}^{acc} = \frac{dn_{i,k}}{dt} = V_k \frac{dc_{i,k}}{dt} + c_{i,k} \frac{dV_k}{dt} \quad (A1.3)$$

  where $n_{i,k}$ and $c_{i,k}$ are respectively the mole and the concentration of the species i in the phase k.

Thus Eq. (A1.1) becomes:

$$\frac{dc_{i,k}}{dt} = \frac{1}{V_k}\left(F_{i,k}^{in} - F_{i,k}^{out} - c_{i,k}\frac{dV_k}{dt}\right) + \sum_{j}^{n_{r_k}} \nu_{i,j,k} r_{j,k} - \sum_{l}^{n_p - 1} J_{i,k \to l \neq k} \quad (A1.4)$$

General volume balance

The general volume balance for the phase k can be represented as Eq. (A1.5) under the assumption that the density is constant

$$q_k^{in} + q_k^{p} = q_k^{out} + q_k^{acc} \quad (A1.5)$$

where:

- q_k^{in}, q_k^{p}, q_k^{out} are respectively the volumetric inflow, production and outflow rates of the phase k, i.e. the volumetric non-accumulation rates of the phase k, denoted as $q_k^{t\neq acc}$, and can be represented by Eq. (A1.6) under the assumption that the mixture is ideal, its volume is equal to the sum of the volume of each constituent.

$$q_k^{t\neq acc} = \sum_{i}^{n_{sk}} v_i F_{i,k}^{t\neq acc} \quad (A1.6)$$

  where v_i is the molar volume of the species i,

- q_k^{acc} is the volumetric accumulation rate of the phase k, and can by represented as:

$$q_k^{acc} = \frac{dV_k}{dt} \quad (A1.7)$$

Thus Eq. (A1.5) becomes:

$$\frac{dV_k}{dt} = \sum_{i}^{n_{sk}} v_i \left(F_{i,k}^{in} - F_{i,k}^{out} + V_k \sum_{j}^{n_{rk}} \nu_{i,j,k} r_{j,k} - V_k \sum_{l}^{n_p - 1} J_{i,k\rightarrow l\neq k} \right) \quad (A1.8)$$

General heat balance

The general heat balance for the reaction system can be represented as:

$$Q^{in} + Q^{p} = Q^{out} + Q^{out} \quad (A1.9)$$

where:

- Q^{in} is the heat-inflow rate, and can be represented as:

$$Q^{in} = \sum_{k}^{n_p} \sum_{i}^{n_{sk}} F_{i,k}^{in} M_i Cp_i \left(T^{in} - T \right) \quad (A1.10)$$

where M_i is the molar mass of the species i, Cp_i is the heat capacity of the species i, T^{in} is the feed temperature, T is the reaction system temperature,

- Q^p is the heat-production rate exchanged, and can be represented by Eq. (A1.11) if the heat losses due to liquid evaporation and vapor condensation and with the environment are neglected.

$$Q^p = US\left(T^{exch} - T\right) \quad (A1.11)$$

where U is the global heat-transfer coefficient, S is the heat-transfer area, T^{exch} is the temperature of the fluid flowing in the jacket,

- Q^{out} is the heat-outflow rate, and can be considered to be zero under the assumption that the reaction mixture temperature is uniform/homogeneous and equal to that of output,
- Q^{acc} is the heat-accumulation rate of the reaction system, and can be represented by Eq. (A1.12), since (i) the reaction system is composed of the reactor and the reaction mixture; (ii) the reaction mixture enthalpy is a function of the reaction system temperature and the mole quantity of each constituent of reaction mixture. In Eq. (A1.12), the terms on the right can be further represented by Eqs. (A1.13)-(A1.15).

$$Q^{acc} = \frac{dH_r}{dt} + \frac{dH_{rm}}{dt} = \frac{dH_r}{dt} + \left(\frac{\partial H_{rm}}{\partial T}\right)_{n_i}\frac{dT}{dt} + \sum_i \left(\frac{\partial H_{sys}}{\partial n_i}\right)_{T,n_{h\neq i}}\frac{dn_i}{dt} \quad (A1.12)$$

$$\frac{dH_r}{dt} = m_r Cp_r \frac{dT}{dt} \quad (A1.13)$$

$$\left(\frac{\partial H_{rm}}{\partial T}\right)_{n_i}\frac{dT}{dt} = \sum_k^{n_p}\sum_i^{n_{sk}} V_k c_{i,k} M_i Cp_i \frac{dT}{dt} \quad (A1.14)$$

$$\sum_i \left(\frac{\partial H_{sys}}{\partial n_i}\right)_{T,n_{h\neq i}}\frac{dn_i}{dt} = \sum_k^{n_p}\sum_j^{n_{rk}} V_k \Delta h_{j,k} r_{j,k} \quad (A1.15)$$

- where H_r and H_{rm} are the enthalpies of respectively the reactor and the reaction mixture, m_r is the empty reactor mass, Cp_r is the heat capacity of the reactor manufacturing material, $\Delta h_{j,k}$ is the molar enthalpy of the reaction j in the phase k.

Thus Eq. (A1.9) becomes:

$$\frac{dT}{dt} = \frac{\sum_k^{n_p}\sum_i^{n_{sk}} F_{i,k}^{in} M_i Cp_i\left(T^{in} - T\right) + US\left(T^{exch} - T\right) - \sum_k^{n_p}\sum_j^{n_{rk}} V_k \Delta h_{j,k} r_{j,k}}{\sum_k^{n_p}\sum_i^{n_{sk}} V_k c_{i,k} M_i Cp_i + m_r Cp_r} \quad (A1.16)$$

Appendix 2: Experimental responses for the case study 1

The experimental responses of the PEs, the 1st DAEPPI, the 2nd DAEPPI and the VEPO are shown in Table A2.1.

Table A2.1: Experimental responses of the case study 1.

Experiment	N° sample	C_{EO} [mol/m^3]	C_{EL} [mol/m^3]	C_{EP} [mol/m^3]	C_{MGs} [mol/m^3]	C_{DGs} [mol/m^3]	C_{TGs} [mol/m^3]
1st PE	1	463.44	893.59	187.75	290.32	168.82	252.12
	2	576.78	1115.72	230.54	340.03	150.69	106.81
	3	652.84	1261.19	259.93	290.70	106.74	62.31
	4	698.35	1348.82	277.50	247.68	80.97	40.10
	5	741.81	1427.25	294.10	211.07	52.04	22.14
	6	759.11	1467.98	300.02	174.99	41.17	19.25
	7	784.83	1516.16	310.07	128.53	28.01	14.39
	8	811.80	1564.32	320.79	85.97	16.11	6.52
2nd PE	1	422.70	846.90	176.01	332.53	198.96	252.78
	2	499.24	1001.30	205.97	358.47	167.31	168.53
	3	565.60	1134.23	231.46	375.82	132.84	102.65
	4	617.54	1237.67	251.71	320.87	114.37	70.77
	5	663.38	1329.17	270.22	284.55	85.76	46.12
	6	705.01	1409.91	286.97	229.16	63.44	30.41
	7	725.12	1446.56	294.46	204.77	50.77	24.31
	8	764.07	1525.89	309.52	138.04	30.64	13.45
	9	794.81	1583.89	321.84	78.74	15.94	8.06
3rd PE	1	627.10	1251.39	256.19	250.75	105.13	92.66
	2	675.79	1348.73	275.23	226.01	80.56	57.62
	3	707.44	1413.14	287.09	203.94	61.22	39.09
	4	749.90	1497.92	304.58	153.72	36.87	20.79
	5	770.64	1538.77	312.09	120.61	23.80	16.39
	6	787.50	1569.90	319.00	91.52	16.21	12.05
	7	806.69	1609.09	325.65	67.86	8.50	2.05
	8	805.87	1608.74	325.59	72.18	7.94	1.29
	9	801.45	1600.45	323.75	88.43	8.22	0.45
4th PE	1	606.33	1211.48	246.34	294.86	112.17	97.58
	2	685.72	1370.20	277.75	268.71	66.20	39.04
	3	751.28	1500.92	302.95	190.54	28.04	12.13
	4	787.79	1573.26	317.52	116.38	12.26	4.73
	5	804.59	1605.36	324.73	82.50	6.08	0.76
	6	806.52	1608.91	325.49	78.98	5.26	0.31
	7	803.81	1603.25	325.30	87.47	5.31	0.30
	8	802.46	1600.78	324.61	91.39	5.59	0.32
	9	804.35	1604.19	325.23	86.21	5.27	0.27
5th PE	1	817.41	1621.41	324.31	56.00	4.67	0.99
	2	785.15	1557.64	312.11	111.17	17.76	11.63
1st DAEPPI	1	762.18	1730.56	338.34	31.23	5.21	3.12
	2	777.20	1739.78	340.44	22.11	4.14	1.76
2nd DAEPPI	1	202.22	92.31	52.56	8.03	288.44	734.44
	2	232.47	128.11	70.87	15.21	345.02	629.74
VEPO	1	746.23	1702.51	317.42	48.12	4.22	1.78

Appendix 3: Supplementary information for the case study 2

The supplementary information for the case study 2, including the operating conditions in general form of the PEs, the experimental responses and the parameter estimation results using the experimental data of the PEs, are given.

Table A3.1: Operating conditions in general form of the PEs for the case study 2.

Run	1	2	3	4	5	6	7, 8, 9	10	11
Reactor	SBSTR	SBSTR	SBSTR	SBSTR	SBSTR	SBSTR	CTR	CTR	CTR
τ [s]	1800, 3600, 5400, 7200, 9000, 10800, 14400, 18000, 19800, 21600								
T^{int} [K]	338.15	338.15	348.15	328.15	338.15	338.15			
n_{FA}^{int} [mol]	0	0	0	0	0	0			
n_{HP}^{int} [mol]	0	0	0	0	0	0			
n_{SA}^{int} [mol]	0.0488	0.0976	0.0488	0.0488	0.0488	0.0488			
n_{PFA}^{int} [mol]	0	0	0	0	0	0			
n_{W}^{int} [mol]	0	0	0	0	0	0			
n_{DB}^{int} [mol]	1.9013	1.9013	1.9013	1.9013	1.9013	1.9013			
T_1^{in} [K]	298.15	298.15	298.15	298.15	298.15	298.15	328.15, 338.15, 348.15	338.15	338.15
T_2^{in} [K]	0	0	0	0	0	0	328.15, 338.15, 348.15	338.15	338.15
F_1^{in} [$\times 10^{-8}$ $m^3 \cdot s^{-1}$]	3.8333	3.8333	3.8333	3.8333	3.8333	3.8333	4.0369	5.8702	4.2035
F_2^{in} [$\times 10^{-8}$ $m^3 \cdot s^{-1}$]	0	0	0	0	0	0	11.7965	11.7965	11.7965
t_1^{in} [s]	1800	1800	1800	1800	1800	1800			
t_2^{in} [s]	0	0	0	0	0	0			
$c_{FA,1}^{in}$ [$mol \cdot m^{-3}$]	2467.9	2467.9	2467.9	2467.9	1709.3	3506.1	2471.6	1717.4	3509.9
$c_{HP,1}^{in}$ [$mol \cdot m^{-3}$]	9378.2	9378.2	9378.2	9378.2	9959.5	8603.6	9273.566	9882.9	8503.1
$c_{SA,1}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	342.6	235.6	329.0
$c_{PFA,1}^{in}$ [$mol \cdot m^{-3}$]	523.5	523.5	523.5	523.5	371. 4	718.2	523.2	372.6	717.8
$c_{W,1}^{in}$ [$mol \cdot m^{-3}$]	35386.3	35386.3	35386.3	35386.3	36690.5	33612.7	35018.9	36428.0	33259.4
$c_{FA,2}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	0	0	0
$c_{HP,2}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	0	0	0
$c_{PFA,2}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	0	0	0
$c_{W,2}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	0	0	0
$c_{DB,2}^{in}$ [$mol \cdot m^{-3}$]	0	0	0	0	0	0	4637.4	4637.4	4637.4

Table A3.2: Experimental responses of the case study 2.

Experiment	N° sample	$C_{HP.aq}$ [mol/m^3]	$C_{DB.org}$ [mol/m^3]	$C_{Ep.org}$ [mol/m^3]	T [K]	C_{H^+aq} [mol/m^3]
1st PE	1	2741.33	3476.75	890.88	345.15	265
	2	2313.00	3229.59	1914.00	341.15	1375
	3	2604.27	2481.75	2623.92	340.35	1750
	4	1634.52	2314.86	3069.36	339.75	1910
	5	1168.49	2024.39	3312.96	339.15	1875
	6	849.81	1587.40	3410.40	339.15	1910
	7	476.31	1410.32	3507.84	338.65	2015
	8	315.25	1331.33	3431.28	338.45	1945
	9	229.59	1228.14	3424.32	338.35	1935
	10	181.61	1184.82	3486.96	338.15	2025
2nd PE	1	2302.72	2697.06	2255.04	347.15	1785
	2	1425.49	2452.45	2449.92	349.15	1855
	3	1117.09	2228.23	2930.16	340.15	1900
	4	870.37	2075.35	3006.72	339.75	1920
	5	702.47	2009.10	3048.48	339.05	1930
	6	524.28	1953.04	3090.24	338.85	1910
	7	414.63	1912.27	3145.92	338.65	1870
	8	270.71	1875.33	3006.72	338.45	1865
	9	181.61	1839.66	3104.16	338.35	1855
	10	147.35	1675.31	3132.00	338.15	1835
3rd PE	1	3313.59	3636.00	863.04	348.25	1755
	2	3227.92	3099.64	1893.12	350.55	2075
	3	2319.85	1440.89	2721.36	351.55	2165
	4	1428.92	1360.63	2999.76	348.75	2230
	5	846.39	1296.93	3222.48	348.55	2430
	6	603.09	1210.30	3264.24	347.65	2450
	7	339.24	1168.26	3312.96	346.55	2460
	8	167.91	1082.90	3173.76	346.45	2485
	9	126.79	1102.01	3194.64	345.95	2495
	10	113.08	1031.94	3250.32	345.95	2505
4th PE	1	4673.97	3604.15	883.92	330.55	2325
	2	4804.19	2819.36	2011.44	331.25	2825
	3	3625.41	2393.85	2275.92	330.45	3020
	4	2515.17	1919.92	2380.32	329.65	3235
	5	1925.79	1369.55	2429.04	328.85	3445
	6	1699.63	1274.00	2637.84	328.65	3630
	7	1528.29	1184.82	2672.64	328.65	3695
	8	1415.21	1066.34	2867.52	328.45	3715
	9	1240.45	1034.49	2958.00	328.15	3735

	10	1072.55	942.76	3083.28	328.05	3760
5th PE	1	4149.69	3551.91	1197.12	343.45	1550
	2	4064.03	3037.22	2206.32	340.25	1485
	3	3971.51	1624.35	2484.72	339.75	1440
	4	2487.76	1351.71	2630.88	339.05	1425
	5	2282.16	1179.72	2651.76	338.65	1420
	6	1843.55	1068.89	3034.56	338.65	1400
	7	1634.52	996.27	3466.08	338.55	1360
	8	1411.79	856.13	3862.80	338.05	1315
	9	1319.27	840.84	3911.52	337.75	1270
	10	1202.76	810.26	3967.20	337.75	1215
6th PE	1	1987.47	3604.15	1044.00	344.35	3655
	2	1936.07	3495.86	1788.72	343.75	3645
	3	1761.31	2355.63	2533.44	342.85	4000
	4	877.23	1922.47	2999.76	341.95	3780
	5	651.07	1801.44	3097.20	340.45	3640
	6	411.20	1735.19	3194.64	338.65	3575
	7	212.45	1551.73	3138.96	338.15	3550
	8	99.37	1533.90	3187.68	338.05	3050
	9	78.81	1484.21	3410.40	337.75	2795
	10	61.68	1391.21	3535.68	337.05	2660
7th, 8th, 9th PEs	1	-	4502.16	179.57	-	-
	2	-	4476.99	314.94	-	-
	3	-	4462.37	172.61	-	-
10th PE	1	-	4577.48	220.63	-	-
11th PE	1	-	4308.42	349.39	-	-
1st DAEPPI	1	6148.47	4089.34	456.37	302.90	8250
	2	5960.02	3923.22	658.81	303.05	8350
2nd DAEPPI	1	1732.86	714.18	3642.95	324.60	4070
	2	1647.73	651.56	3625.48	324.50	3990
VEPO	1	928.94	217.05	3857.67	363.80	9755

Table A3.3: Parameter estimation results using the experimental data of the PEs.

Reaction network	RNUT1	RNUT2	RNUT3	RNUT4	RNUT5	RNUT6
Parameter estimation criterion (Eq. (2.19))	29.2186	29.4016	28.7852	29.3041	28.9732	28.9988
$\log(k_{1,Tmin})$	-7.9716	-7.7554	-7.9060	-7.8641	-7.7642	-9.1430
$\log(k_{1,Tmax})$	-7.9164	-7.6330	-7.3575	-7.3882	-8.3496	-8.0837
$\log(\Delta h_{1,Tref})$	3.3149	3.7107	3.7094	3.6551	3.7518	3.2828
$\log(k_{2,Tmin})$	-6.6904	-6.6822	-6.9635	-6.1380	-6.5155	-6.7959
$\log(k_{2,Tmax})$	-6.6811	-6.6009	-8.1050	-6.0366	-6.9061	-6.5309
$\log(\Delta h_{2,Tref})$	3.3149	3.7107	3.7094	3.6551	3.7518	3.2828
$\log(k_{6,Tmin})$		-3.4645		-3.4775		-2.7504
$\log(k_{6,Tmax})$		-2.6395		-2.3827		-2.6998
$\log(\Delta h_{6,Tref})$		5.0233		5.3032		5.0832
$\log(k_{8,Tmin})$	-4.1652	-4.2056	-5.2430	-3.9515	-4.0893	-3.4123
$\log(k_{8,Tmax})$	-3.1353	-3.9503	-5.0899	-2.9327	-3.4562	-3.1048
$\log(\Delta h_{8,Tref})$	5.0705	4.8331	5.1024	4.7077	5.3165	5.2657
$\log(k_{9,Tmin})$	-8.3405	-7.9505			-8.8828	-9.7505
$\log(k_{9,Tmax})$	-8.2202	-7.7097			-8.5958	-8.4581
$\log(\Delta h_{9,Tref})$	4.0588	4.9506			4.0647	4.0963
$\log(k_{10,Tmin})$					-8.8828	-9.7505
$\log(k_{10,Tmax})$					-8.5958	-8.4581
$\log(\Delta h_{10,Tref})$					4.0647	4.0963
$\log(k_{11,Tmin})$	-8.3405	-7.9505			-8.8828	-9.7505
$\log(k_{11,Tmax})$	-8.2202	-7.7097			-8.5958	-8.4581
$\log(\Delta h_{11,Tref})$	4.0588	4.9506			4.0647	4.0963
$\log(k_{12,Tmin})$			-9.5195	-8.5512	-8.8828	-9.7505
$\log(k_{12,Tmax})$			-9.4054	-7.3010	-8.5958	-8.4581
$\log(\Delta h_{12,Tref})$			4.4859	4.2945	4.0647	4.0963
$\log(m_{FA,I/II})$	2.2445	2.1297	2.1958	2.1347	2.1261	2.1317
$\log(K_{FA,I,SBSTR})$	0.4167	0.8869	0.5793	0.0017	0.7809	-0.0980
$\log(K_{FA,I,CTR})$	0.2916	1.0805	-0.3758	-0.1419	0.6627	0.6247
$\log(m_{PFA,I/II})$	1.2046	1.1291	1.1993	1.3970	1.2387	1.1501
$\log(K_{PFA,I,SBSTR})$	-0.7846	0.0150	0.4412	-0.0820	1.1384	0.4052
$\log(K_{PFA,I,CTR})$	-0.3674	-0.0946	1.1534	-0.1302	0.5480	0.2309

Résumé élargi

Des modèles adéquats et précis décrivant quantitativement les synthèses de produits chimiques fins et pharmaceutiques sont essentiels pour designer et optimiser les procédés chimiques. Cependant, il est difficile, expérimentalement long et coûteux de développer de tels modèles. Des stratégies appropriées, efficaces et systématiques servant au développement du modèle sont donc nécessaires. Dans ce contexte, les objectifs de ce travail consistent au développement méthodologique, à l'implémentation numérique et à la validation expérimentale d'une stratégie systématique de développement du modèle.

La stratégie développée servant au développement du modèle se compose de cinq modules :

1. Module d'acquisition initiale de données : Des expériences préliminaires sont planifiées et effectuées dans le but d'avoir une première impression sur le système réactionnel. Les données expérimentales servant à l'estimation des paramètres, comprenant les conditions opératoires et les réponses mesurées, sont initialisées et seront actualisées après les expériences dans les modules suivants.
2. Module de développement du modèle : Ce module est composé de trois étapes: le développement de la structure du modèle, l'analyse de la structure du modèle et le développement des paramètres du modèle. La structure du modèle est développée en se basant sur le réseau réactionnel, extrait du super-réseau réactionnel comprenant toutes les réactions chimiques et les transferts de matière faisables. Evidemment, le réseau réactionnel théoriquement optimal et celui finalement identifié représentant mieux le système réactionnel sont compris dans le super-réseau réactionnel. L'application du super-réseau réactionnel permet la stratégie d'être appliquée pour non seulement les systèmes réactionnels homogènes, mais aussi ceux hétérogènes.
3. Module d'identification du modèle : Pour décrire les systèmes réactionnels dans les réacteurs pris en compte, une ou plusieurs séries de Modèles à Tester (MTs) sont développées. Le nombre des Séries de MTs (SMTs) égale celui des réseaux réactionnels candidats; le nombre des MTs contenus dans une SMT est égal à celui des réacteurs considérés. Pour identifier le réseau réactionnel le plus adéquat parmi ceux proposés, une expérience maximisant la différence entre les prédictions des modèles est planifiée et effectuée. Puis, suite à l'estimation paramétrique en utilisant les données expérimentales actualisées, au moins une SMT (correspondant à un réseau réactionnel candidat) est enlevée comme prévu. Après quelques expériences, la SMT restante est considérée comme la meilleure.

4. Module d'affinage du modèle : Après avoir identifié le réseau réactionnel le plus adéquat, la précision du modèle est évaluée. Si la précision du modèle ne satisfait pas aux exigences prédéfinies, les expériences d'acquisition de données maximisant la fonction d'optimalité de la matrice de variance-covariance des paramètres sont planifiées et conduites jusqu'à ce que les exigences de précision sont atteintes.
5. Module de validation du modèle : L'objective de ce module est de vérifier expérimentalement la capacité prédictive du modèle adéquat et précis. Le critère de planification de l'expérience de validation est la somme des performances normalisées, qui doit être maximisée.

La stratégie développée est méthodologiquement systématique, car elle fournit non seulement une procédure unifiée pour le développement du modèle, mais aussi les solutions pour résoudre tous les problèmes possibles rencontrés suite à l'évaluation du modèle en termes d'adéquation et de précision, incluant :

- si plusieurs SMTs concurrentes basées sur différents réseaux réactionnels candidats, sont disponibles, comment identifier efficacement la SMT basée sur le réseau réactionnel le plus adéquat?
- si les paramètres du modèle ne sont pas assez précis, comment raffiner efficacement leur précisions?

De plus, cette stratégie permet de modéliser simultanément plusieurs réacteurs. Les expériences servant à l'identification, au raffinage et à la validation du modèle, sont donc planifiées en compte tous les réacteurs disponibles, ce qui élargit les fenêtres expérimentales.

La stratégie développée guide les utilisateurs à la modélisation des réacteurs considérés et à l'identification du réacteur optimal et des conditions opératoires correspondantes. Lors de l'application pratique de la stratégie, les utilisateurs doivent non seulement effectuer les expériences, mais également faire la programmation pour résoudre les problèmes d'optimisation, tels que l'estimation des paramètres du modèle et la planification des expériences d'acquisition de données. Dans ce cas, afin de réduire les travaux des utilisateurs, un logiciel, fournissant des interfaces conviviales et intégrant les fonctions essentielles de la stratégie, telles que l'estimation des paramètres, l'évaluation du modèle et la planification expérimentale pour l'affinage du modèle et l'optimisation des performances, est développée en utilisant Matlab R2014a. En prenant compte de la difficulté de l'implémentation numérique de la stratégie, le domaine d'application et les fonctionnalités de la version initiale du logiciel sont beaucoup réduits par rapport à ceux de la stratégie, concrètement :

- La version initiale du logiciel ne convient qu'aux systèmes réactionnels les plus communs pour les synthèses de produits chimiques fins et pharmaceutiques, soit les systèmes réactionnels liquides en catalyse homogène, quatre réacteurs idéaux sont pris en compte: réacteur fermé en cuve agitée, réacteur semi-fermé en cuve agitée, réacteur continu en cuve agitée et réacteur tubulaire continu.
- Les simplifications fonctionnelles suivantes sont appliquées sans perdre les fonctions fondamentales:
  - Le super-réseau réactionnel et les réseaux réactionnels candidats sont proposés par l'utilisateur sur la base de son expertise;
  - Les ordres cinétiques ne sont pas considérés comme les paramètres à estimer;
  - L'analyse de la structure du modèle n'est pas intégrée au logiciel;
  - L'adéquation du modèle est qualitativement évaluée;
  - Si plusieurs SMTs sont disponibles, la SMT la plus appropriée, dont la valeur de la fonction objective de l'estimation paramétrique est minimum, est directement identifié.

Afin de valider les faisabilités et généralités de la stratégie et du logiciel basé sur la stratégie, tenant compte des réacteurs et des instruments d'analyse disponibles en laboratoire, deux cas d'étude, relatifs à la valorisation de l'huile de tournesol, sont sélectionnés. Il s'agit de l'éthanolyse de l'huile de tournesol catalysée par NaOH et de l'époxydation de l'huile de tournesol par acide performique généré in situ.

La faisabilité de la stratégie et la praticabilité du logiciel ont été tout d'abord démontrées par le cas d'étude de l'éthanolyse. Le réacteur fermé en cuve agitée et le réacteur tubulaire sont considérés. Une série des modèles adéquats et précis, dont la capacité prédictive du modèle du réacteur fermé en cuve agitée (le réacteur optimal) a été vérifiée expérimentalement, a été développé après une série d'expériences avec différents objectives, incluant:

- 5 expériences préliminaires désignées par l'utilisateur sur la base de son expertise,
- 2 expériences d'acquisition de données pour l'affinage de la précision des paramètres désignée par le logiciel,
- 1 expérience de validation pour l'optimisation des performances désignée également par le logiciel.

Les paramètres cinétiques, soit les énergies d'activation et les facteurs pré-exponentiels, que nous avons déterminés (dérivés des paramètres réellement estimés, soit les logarithmes des paramètres cinétiques) ont des même ordres de grandeur que ceux de la littérature, à l'exception des facteurs pré-exponentiels de deux réactions. De plus, les paramètres réellement estimés, utilisés pour

calculer les deux facteurs pré-exponentiels susmentionnés, sont raisonnables, puisqu'ils sont dans les intervalles de confiance correspondants issus de la littérature.

L'application de la stratégie au cas d'étude d'époxydation a été ensuite exécutée. La stratégie expérimentale est composée de 11 expériences préliminaires, 2 expériences d'acquisition de données pour l'affinage de la précision des paramètres et 1 expérience de validation pour l'optimisation des performances. La SMT désirable n'est pas finalement développée, car dans l'expérience de validation, la concentration d'époxyde dans la phase organique, n'est pas dans son intervalle de confiance prédit. Cela peut être dû au fait que les paramètres estimés, caractérisant la réaction d'ouverture de cycle, donc la vitesse de réaction est relativement faible, sont beaucoup moins que les valeurs théoriques référentielles. Néanmoins, étant donné que tous les autres paramètres sont bien estimés, la faisabilité de la stratégie est considéré (quasi-)faisable.

Les faisabilités et généralités de la stratégie et du logiciel basé sur la stratégie ont été validées avec succès par deux cas d'étude expérimentaux en catalyse homogène. Les futurs travaux devront se focaliser sur leur validation par plus de cas d'étude avec des caractéristiques différentes, par exemple: ceux en catalyse hétérogène, telles que le système réactionnel de catalyse gaz-liquide-solide dans un réacteur à lit fixe, le système réactionnel de catalyse gaz-solide dans un réacteur à lit fluidisé, etc. Pour cela, il est nécessaire d'enrichir la base de données du logiciel en intégrant les structures générales du modèle pour plus de réacteurs. Par conséquent, une fois que la description du système réactionnel est proposée, la structure spécifique du modèle pour chaque réacteur considéré peut être automatiquement générée. De plus, afin que le logiciel soit plus pratique pour les utilisateurs, les fonctions suivantes seront ajoutées :

- Les fonctions prometteuses, qui n'ont pas été implémentées lors du développement de la version initiale du logiciel, incluant :
    1. le développement du super-réseau réactionnel par l'optimiseur,
    2. l'identification du réseau réactionnel le plus approprié par l'optimiseur,
    3. analyse de la structure du modèle,
    4. évaluation quantitative du modèle en termes d'adéquation,
    5. planification expérimentale pour la discrimination des modèles;
- Les fonctions permettant aux utilisateurs de personnaliser leurs besoins, par exemple, de personnaliser les performances, qui ne sont pas dans la liste de suggestions, de personnaliser les contraintes d'égalité ou d'inégalité linéaires ou non linéaires des paramètres du modèle ou des conditions opératoires.

- Les fonctions rendant le logiciel plus ergonomique, telles que la commutation arbitraire entre les interfaces, la saisie de données en chargeant un fichier Excel contenant des données.