



HAL
open science

Development and implementation of systematic model-development strategy using model-based experimental design

Zhengkun Jiang, Jean-François Portha, Jean-Marc Commenge, Laurent Falk

► To cite this version:

Zhengkun Jiang, Jean-François Portha, Jean-Marc Commenge, Laurent Falk. Development and implementation of systematic model-development strategy using model-based experimental design. Chemical Engineering Research and Design, 2019, 146, pp.290-310. 10.1016/j.cherd.2019.03.020 . hal-03058318

HAL Id: hal-03058318

<https://hal.univ-lorraine.fr/hal-03058318v1>

Submitted on 11 Dec 2020

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Development and implementation of systematic model-development strategy using model-based experimental design

Zhengkun JIANG*, Jean-François PORTHA, Jean-Marc COMMENGE, Laurent FALK

Laboratoire Réactions et Génie des Procédés, Université de Lorraine, CNRS, 1 rue Grandville, BP 20451, 54001 Nancy, France

*Corresponding author: zhengkun.jiang@univ-lorraine.fr

Abstract

In this paper, a systematic model-development strategy using model-based experimental design, consisting of initial data acquisition, model development, model identification, model validation and model refining, is developed. Model 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, model-based experimental design is performed by taking into consideration several reactors, which enlarges the explored experimental windows. Based on this strategy, a computer-aided software, facilitating the user's work, is developed for model development of liquid-phase reaction systems. The application of the strategy is demonstrated with an experimental case study involving NaOH-catalyzed ethanolysis of sunflower oil.

Highlights

- Systematic model-development strategy is developed based on model-based experimental design.
- Reaction networks are proposed within reaction supernetwork.

- Experimental design is performed by taking into account several reactors.
- Computer-aided software is developed for liquid-phase reaction systems.

Keywords

Model development, Reaction supernetwork, Model-based experimental design, Multi-equipment strategy, Ethanolysis

1 Introduction

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. Such a strategy is composed of the activities on model development, model evaluation as well as the ones following model evaluation.

The conventional model development is a sequential procedure consisting of three steps: (i) *model structure development*, i.e., constructing the mathematical representation (a specific set of equations) for describing the reaction system under investigation; (ii) *model structure analysis*, i.e., verifying whether the parameters involved in the model structure constructed above are identifiable; (iii) *model parameter estimation*.

Model structure development is based on the reaction network (otherwise called as reaction pathway, reaction mechanism or reaction scheme), that is the set of reactions supposed to occur in the reaction system. Traditionally, this reaction network is 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 network. Some advanced techniques have been proposed to guide the user.

Target Factor Analysis, proposed by Bonvin and Rippin (1990) and Amrhein et al. (1999), is able to test out the reactions within the chemical reaction network to verify whether they can be reasonably accepted or not. In order to possibly propose the right 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, is a subset extracted from this reaction supernetwork. Furthermore, reaction network can be also elucidated through evolutionary algorithms such as genetic programming (Koza et al., 2007) and genetic algorithm (Hii et al., 2014). Considering that the number of variables to be actually measured is commonly limited compared to that of variables needed to be measured (i.e. all time-dependent state variables), model structure analysis is then required to evaluate model structure in terms of identifiability (and distinguishability, if more than one candidate model structure), before any experimental activities followed by model parameter estimation (Asprey and Macchietto, 2000; Franceschini and Macchietto, 2008; Galvanin et al., 2013). For more details about the methods used for evaluating model identifiability and distinguishability, refer to the reviews given by Ollivier (1990) and by Miao et al. (2011). Once model structure has been verified structurally identifiable, model parameters can be estimated conventionally in a one-step manner or sequentially over several steps (Brendel et al., 2006; Michalik et al., 2009).

Following model development and model evaluation in terms of (adequacy of) structures and (accuracy of) parameters, there are potentially two main problems: (i) if several rival models are available, how to identify efficiently the model with the best structure; (ii) if model parameters are not accurate enough, how to improve efficiently parameter precision. Model-based experimental

design is an interesting tool to address such problems. Based on information already acquired on the models, it aims at designing an experiment for obtaining maximum experimental information used for different goals. A detail review about model-based experimental design used for model discrimination and parameter precision is given by Franceschini and Macchietto (2008). Furthermore, model-based experimental design can be also applied to verify model prediction quality (Zhang et al., 2015).

Table 1: Survey on different model-development strategies.

N°	1	2	3	4	5	6	7
Reference	Asprey and Macchietto, 2000	Brendel et al., 2006	Michalik et al., 2009	Mathieu et al., 2013	Hii et al., 2014	Zhang et al., 2015	Violet et al., 2016
Model structure development ^a	U	T	U	U	O	S	U
Model structure analysis	Yes	No	Yes	No	No	No	Yes
Model parameter estimation ^b	D	S	S	D	D	D	D
Model identification ^c	S	D	D	No	D	S	S
Model refining	Yes	No	No	Yes	No	No	No
Model validation	No	No	No	No	No	Yes	No
Multiple reactors	No	No	No	Yes	No	No	No
Software implementation	No	No	No	No	No	No	No
Application reaction system ^d	Homo	Homo	Hetero	Homo	Homo	Homo	Homo

^a There are four methods for model structure development: U, model structure is developed by user; T, target factor analysis 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: D, direct method; S, stepwise method.

^c The model with the best model structure can be identified through two methods: D, direct method; S, stepwise method.

^d Application reaction system can be classified into two categories: Homo, homogeneous reaction system; Hetero, heterogeneous reaction system.

A survey on different model-development strategies is given in Table 1. It can be seen that each strategy has both advantages and application limitations. For instance, (i) the application reaction system of all strategies is homogeneous apart from the strategy N°3, in which model refining and validation are not taken into consideration; (ii) most strategies only use one kind of reactor, except for the strategy N°4, in which the efficiency of model parameter precision improvement has been proven enhanced by using several types of reactors. It is also observed that none of the strategies has been implemented as software, therefore, user expertise on reaction kinetics, modeling and statistical analysis are required.

In this work, a systematic model-development strategy based on model-based experimental design for simultaneous model development and reactor performance optimization is proposed. The reaction network is developed within the reaction supernetwork containing all feasible chemical reactions and mass transfers. Based on this strategy, a computer-aided software, making model development easier and faster, is developed. The strategy application and the software usability are then demonstrated through a complex example involving NaOH-catalyzed ethanolysis of sunflower oil.

2 Systematic strategy for model development

2.1 Description of the strategy

The aim of the strategy (illustrated in Figure 1) is to identify a set of adequate and accurate models as well as the optimal reactor with its associated optimal operating conditions for a given synthesis, while minimizing the number of experiments to be performed. The steps involved in the strategy are grouped into five modules (shown in Figure 2), which will be now described. Note that, all the steps mentioned below refer to Figure 1.

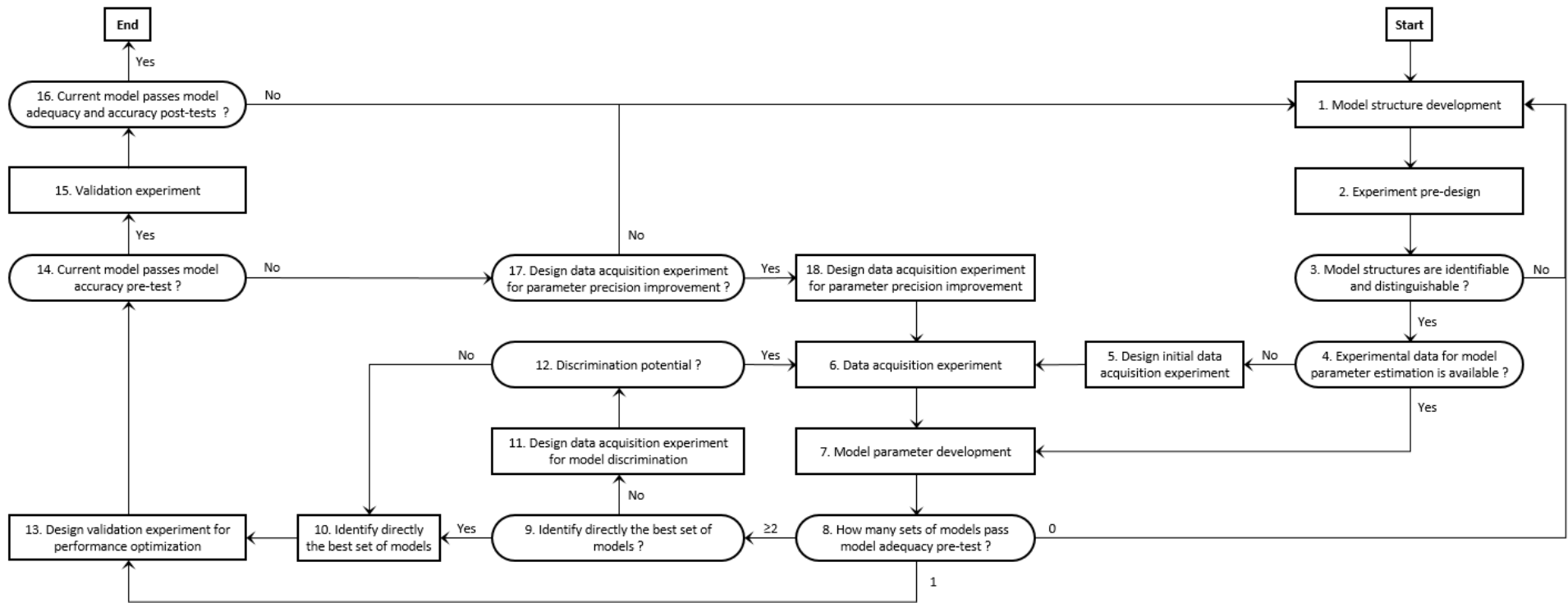


Figure 1: Systematic model-development strategy based on model-based experimental design.

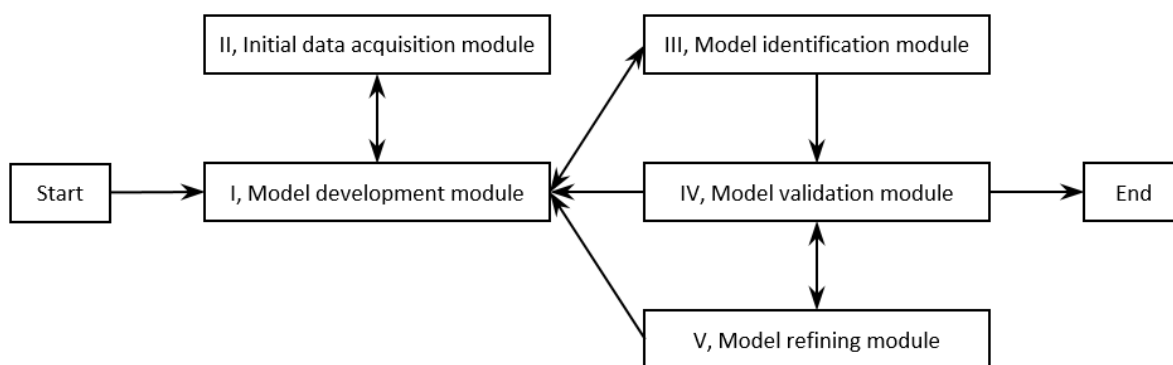


Figure 2: General procedure for model development.

2.1.1 Model development module and initial data acquisition module

The first steps consist of 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). 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): (i) initial model structure development; (ii) following model structure re-development due to the failure of model adequacy evaluation, the parameters involved in the new model structures are not identifiable using current experimental data. If one situation happens, initial data acquisition experiments, i.e. Preliminary Experiments (PEs), have to be designed and performed (steps 5-6).

Finally, one or several sets of Models Under Test (MUTs) are developed after model parameter estimation (step 7) (see Section 2.3). The number of Sets of MUTs (SMUTs) depends on the number of reaction networks under test; the number of MUTs contained in one SMUT depends on the number of reactors taken into consideration.

2.1.2 Model identification module

Following the model development module, the model identification module is executed. 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 2.4.1). According to the number of SMUTs passing model adequacy pre-test, different steps have to be performed: 0, step 1 (model development module); 1, step 13 (model validation module); ≥ 2 , step 9 (model identification module). For the last situation, i.e. if at least two SMUTs pass model adequacy pre-test, the most suitable SMUT can be identified through two methods:

- Direct identification method (steps 9-10): Identify directly the SMUT with the minimal parameter estimation criterion (see Section 2.3.3);
- Stepwise identification method (iterative steps 9-11-12-6-...-6-...-8/12): Design 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 (see Section 2.5.1). After the DAEMD, at least one SMUT can be removed, and the iterative steps continue until one only 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 same or similar parameter estimation criteria, the stepwise one seems better to be used.

2.1.3 Model validation module

Model validation module has two possible inputs: (i), output of model identification module, i.e. the SMUT passing model adequacy pre-test; (ii), output of model refining module, i.e. the SMUT passing both model adequacy and accuracy pre-tests.

After the Validation Experiment for Performance Optimization (VEPO) is designed (step 13) (see Section 2.5.2), 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 2.4.2). 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

Following the 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) is performed. At the beginning of each iteration, manual intervention is required (step 17), namely, the user has to evaluate on the basis of his expertise whether design of a series of Data Acquisition Experiments for Parameter Precision Improvement (DAEPPI) (step 18) (see Section 2.5.3) 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, design of DAEPPI is not necessary.

2.2 Model structure development

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

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \boldsymbol{\xi}, \boldsymbol{\theta}, t) \quad (1)$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \quad (2)$$

$$\hat{\mathbf{y}}(t) = \mathbf{x}(t)M_T \quad (3)$$

Where:

- $\mathbf{x}(t)$ is an n_{sv} -dimensional vector of time-dependent state variables. In this work, the time-dependent state variables taken into consideration are the species concentrations, the phase volumes and the system temperature,
- $\boldsymbol{\xi}$ is an $n_{sp} \times n_{oc} \times n_{exp}$ -dimensional array of operating conditions such as the sampling times, the initial conditions, the flow rate, the duration and the composition of each feed. Operating conditions depend on the reactor type, for more details about the ones characterizing the reactor taken into consideration in this work, see Appendix 1,
- $\boldsymbol{\theta}$ 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}_0 is the initial state for $t=t_0=0$,
- $\hat{\mathbf{y}}(t)$ 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 model structure for a given reaction system in a given reactor, one should (i) propose a reasonable reaction network; (ii) define the rate law for each of the involved chemical reactions and mass transfers; (iii) indicate the reactor type.

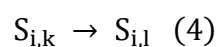
2.2.1 Reaction supernetwork and reaction network

Reaction network development is a sequential procedure composed of two steps: (i) development of reaction supernetwork; (ii) extraction of reaction network from the reaction supernetwork developed above.

The reaction supernetwork is defined to contain not only chemical reactions but also mass transfers, which allows it to be used for the heterogeneous reaction system. It can be developed through two methods:

- Theoretical method: The optimizer generates all possible chemical reactions and mass transfers, and then, eliminates the chemical reactions which do not satisfy the mass balances (Burnham et al., 2008) or the atomic balances (Zhang et al., 2015) as well as the mass transfers in which the involved species are (assumed to be) not soluble in the debit or credit phase;
- Empirical method: The user elucidates all possible chemical reactions and mass transfers on the basis of the literature and his expertise.

In this reaction supernetwork, mass transfers are described in the same manner as chemical reactions. For instance, mass transfer of species S_i between phases k and l is represented by Eq. (4), in which the reactant/product is the species in the debit/credit phase, and its stoichiometric coefficients is $1/-1$.



Furthermore, both chemical reactions and mass transfers are assumed to be irreversible. Therefore, a reversible chemical reaction should 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. The mass transfer proceeds in the opposite direction, if the calculated mass-transfer rate is negative.

The reaction network, a subset of chemical reactions and mass transfers extracted from the reaction supernetwork, can be proposed by the user or the optimizer. To illustrate this concept, the homogeneous alkali-catalyzed alcoholysis of oil (Richard et al., 2013; Likozar and Levec, 2014a; Likozar and Levec, 2014b; Eze et al., 2014; Reyro et al., 2015) is considered as a case study. All possible species in the reaction system 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 (see Table 2), developed on the basis of literature by user, comprises 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).

Table 2: Reaction supernetwork for the homogeneous alkali-catalyzed alcoholysis of 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	$\text{FAAE(I)} + \text{MG(I)} \xrightarrow{\text{Cat.}} \text{DG(I)} + \text{A(I)}$	
8	$\text{FAAE(I)} + \text{G(I)} \xrightarrow{\text{Cat.}} \text{MG(I)} + \text{A(I)}$	
9	$\text{TG(I)} + \text{AMH(I)} \rightarrow \text{AMSFA(I)} + \text{DG(I)}$	
10	$\text{DG(I)} + \text{AMH(I)} \rightarrow \text{AMSFA(I)} + \text{MG(I)}$	saponification of glycerides and FAAE
11	$\text{MG(I)} + \text{AMH(I)} \rightarrow \text{AMSFA(I)} + \text{G(I)}$	
12	$\text{FAAE(I)} + \text{AMH(I)} \rightarrow \text{AMSFA(I)} + \text{A(I)}$	
13	$\text{FFA(I)} + \text{AMH(I)} \rightarrow \text{AMSFA(I)} + \text{W(I)}$	neutralization of FFA
14	$\text{FFA(I)} + \text{AMA(I)} \rightarrow \text{AMSFA(I)} + \text{A(I)}$	
15	$\text{TG(I)} \rightarrow \text{TG(II)}$	
16	$\text{DG(I)} \rightarrow \text{DG(II)}$	
17	$\text{MG(I)} \rightarrow \text{MG(II)}$	
18	$\text{FAAE(I)} \rightarrow \text{FAAE(II)}$	mass transfer
19	$\text{A(I)} \rightarrow \text{A(II)}$	
20	$\text{G(I)} \rightarrow \text{G(II)}$	
21	$\text{FFA(I)} \rightarrow \text{FFA(II)}$	
22	$\text{W(I)} \rightarrow \text{W(II)}$	

The common reaction networks presented in the literature can be extracted from this reaction supernetwork:

- the pseudo-homogeneous reaction network (N°1-14) (Eze et al., 2014; Reyero et al., 2015) is the subset containing all chemical reactions;
- the reduced pseudo-homogeneous reaction network (N°3-8) (Likožar and Levec, 2014b) can be extracted from the above reaction network by retaining only the desired transesterification reactions;

- on the basis of the reduced pseudo-homogeneous reaction network, taking into consideration the mass transfers of TG, DG, MG, FAAE, A and G allows developing the biphasic reaction network (N°3-8, 15-20) (Richard et al., 2013; Likozar and Levec, 2014a).

2.2.2 Rate law

Once the reaction network is available, the rate law for each of the involved chemical reactions and mass transfers has to be defined. For example, in the liquid-phase reaction system, which is concerned in this work, the reaction rate of the chemical reaction j in the phase k can be represented by Eq. (5) according to a power law equation.

$$r_{j,k} = k_{j,k} \prod_i^{n_{s_k}} c_{i,k}^{\gamma_{i,j,k}} \quad (5)$$

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 .

The mass-transfer rate of the species i from phase k to phase l is by default written as:

$$J_{i,k \rightarrow l} = K_{i,k} (c_{i,k} - m_{i,k|l} c_{i,l}) \quad (6)$$

Where $K_{i,k}$ is the overall mass-transfer coefficient of the species i , defined for the debit phase k , $m_{i,k|l}$ is the distribution coefficient of the species i between the phases k and l .

2.2.3 Reactor type

In this work, 4 types of ideal reactors are taken into consideration: Batch Stirred-Tank Reactor (BSTR), Semi-Batch Stirred-Tank Reactor (SBSTR), Continuous Stirred-Tank

Reactor (CSTR) and Continuous Tubular Reactor (CTR), that belong to two reactor configurations: Stirred-Tank Reactor (STR) and Tubular Reactor (TR).

Once the reaction network and the rate law for each of the involved chemical reactions and mass transfers are available, for the three STR-type reactors, the ordinary differential equation describing the variation of the concentration of the species i in the phase k can be written as:

$$\frac{dc_{i,k}}{dt} = \frac{1}{V_k} \left(F_{i,k}^{\text{in}} - F_{i,k}^{\text{out}} - c_{i,k} \frac{dV_k}{dt} \right) + \sum_j^{n_{rk}} v_{i,j,k} r_{j,k} - \sum_l^{n_p-1} J_{i,k \rightarrow l \neq k} \quad (7)$$

Where:

- $F_{i,k}^{\text{in}}$ and $F_{i,k}^{\text{out}}$ are respectively the inflow and the outflow rates of the species i in the phase k , with the following constraints, for the BSTR, there is neither reactant addition ($\sum_k^{n_p} \sum_i^{n_{sk}} v_i F_{i,k}^{\text{in}}$) nor product removal ($\sum_k^{n_p} \sum_i^{n_{sk}} v_i F_{i,k}^{\text{out}}$); for the SBSTR, there is either reactant addition or product removal; for the CSTR, there are both reactant addition and product removal, and their volumetric flow rates are identical,
- n_{rk} represents the number of chemical reactions occurring in the phase k ,
- n_p is the number of phases,
- $v_{i,j,k}$ is the stoichiometric coefficient of the species i for the chemical reaction j in the phase k ,
- v_i is the molar volume of the species i ,
- V_k is the volume of the phase k , its variation with respect to time is written as:

$$\frac{dV_k}{dt} = \sum_i^{n_{sk}} v_i \left(F_{i,k}^{\text{in}} - F_{i,k}^{\text{out}} + V_k \sum_j^{n_{rk}} v_{i,j,k} r_{j,k} - V_k \sum_l^{n_p-1} J_{i,k \rightarrow l \neq k} \right) \quad (8)$$

The ordinary differential equation describing the variation of the temperature with respect to time is written as:

$$\frac{dT}{dt} = \frac{\sum_k^{n_p} \sum_i^{n_{s_k}} F_{i,k}^{in} M_i C_{p_i} (T^{in} - T) + US(T^{exch} - T) - \sum_k^{n_p} \sum_j^{n_{r_k}} V_k \Delta h_{j,k} \Gamma_{j,k}}{\sum_k^{n_p} \sum_i^{n_{s_k}} V_k C_{i,k} M_i C_{p_i}} \quad (9)$$

Where:

- T^{in} and T^{exch} are respectively the feed temperature of the process fluid and the temperature of the fluid flowing in the jacket,
- M_i and C_{p_i} are respectively the molar mass and the heat capacity of the species i ,
- U and S are respectively the global heat-transfer coefficient and the heat-transfer surface,
- $\Delta h_{j,k}$ is the molar enthalpy of the reaction j in the phase k .

For the CTR, similar ordinary differential equations are considered, where the time is substituted by the residence time.

2.3 Model parameter development

2.3.1 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 of the STR-type reactor constructed above, the involved parameters are the reaction rate constants, the kinetic orders (and the molar reaction enthalpies, if necessary) for the chemical reactions as well as the overall mass-transfer coefficients and the distribution coefficients for the mass transfers. Among them, the chemical reaction rate constant can be reformulated by Eq. (10) according to the Arrhenius law when it is considered as a temperature-dependent parameter.

$$k_j = A_j e^{-\frac{E_{a_j}}{RT}} \quad (10)$$

Where A_j and E_{a_j} are the independent kinetic parameters (pre-exponential factor and energy of activation respectively). Considering A_j and E_{a_j} may exhibit very different orders of magnitude, they can be further reformulated by Eqs. (11)-(12):

$$A_j = 10^{\frac{T^{\max} \log k_{j,T^{\max}} - T^{\min} \log k_{j,T^{\min}}}{T^{\max} - T^{\min}}} \quad (11)$$

$$E_{a_j} = \frac{RT^{\max}T^{\min}(\log k_{j,T^{\max}} - \log k_{j,T^{\min}}) \ln 10}{T^{\max} - T^{\min}} \quad (12)$$

Where $\log k_{j,T^{\max}}$ and $\log k_{j,T^{\min}}$ are the reaction rate constant logarithms at respectively the maximal and minimal possible reaction temperatures, namely, the parameters to be actually estimated instead of k_j .

2.3.2 Parameter pre-assignment

In order to decrease the computation time for parameter estimation, one effective way is to pre-assign reasonable values to parameters, for instance, the kinetic orders for the reactant, the solvent, the catalyst and the product can be by default taken equal to its stoichiometric coefficient, 0, 1 and 0 respectively.

2.3.3 Parameter estimation

After parameter reformulation and pre-assignment, the values of the model parameters are obtained by Eq. (13), in which the criterion is built based on Maximum Likelihood Estimator and represents the average likelihood of samples.

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \frac{1}{\sum_{e=1}^{n_{\text{exp}}} n_{\text{sp}_e}} \sum_{e=1}^{n_{\text{exp}}} \frac{n_{\text{sp}_e}}{2} \sum_{r=1}^{n_{\text{resp}}} \ln \left\{ \frac{2\pi}{n_{\text{sp}_e}} \sum_{f=1}^{n_{\text{sp}_e}} [y_r(\xi_{e,f}) - \hat{y}_r(\xi_{e,f}, \theta)]^2 \right\} \quad (13)$$

Where n_{exp} and n_{sp_e} represent respectively the number of experimental runs and the number of samples in the e^{th} experiment, $y_r(\xi_{e,f})$ is the r^{th} measured response obtained under the operating conditions $\xi_{e,f}$, which is the f^{th} set of the operating conditions of the e^{th} experiment, $\hat{y}_r(\xi_{e,f}, \theta)$ is the corresponding model prediction.

2.4 Model evaluation

2.4.1 Model adequacy test

Model adequacy tests can be classified into two categories: model adequacy pre-test and model adequacy post-test. At the step 8 of the strategy, the former one is implemented for each SMUT to evaluate whether it allows fitting all available experiments. At the step 16 of the strategy, the later one is implemented to verify whether the current model describes adequately the VEPO performed at the step 15.

Specifically, model 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 (Franceschini and Macchietto, 2008; Donckels et al., 2009 and Violet et al., 2016).

2.4.2 Model accuracy test

Model accuracy tests can also be classified into two categories: model accuracy pre-test and model accuracy post-test. At the steps 14/16 of the strategy, the former/later one is implemented to

evaluate/verify whether the current model has the potential to predict/has predicted accurately the VEPO.

Specifically, model accuracy is evaluated using the model prediction uncertainty (see Section 2.4.2.2) derived from the model parameter uncertainty (see Section 2.4.2.1).

2.4.2.1 Model parameter uncertainty

The $n_u \times n_u$ -dimensional parameter variance-covariance matrix, denoted as $\Sigma_{\hat{\theta}}$, quantifying the uncertainty on the parameter estimates, can be approximated using the Fisher Information Matrix M_F (Franceschini and Macchietto, 2008):

$$\Sigma_{\hat{\theta}}(\hat{\theta}, \xi_1, \dots, \xi_e, \dots, \xi_{n_{\text{exp}}}) = \left[\sum_{e=1}^{n_{\text{exp}}} M_F(\hat{\theta}, \xi_e) \right]^{-1} \quad (14)$$

With:

$$M_F(\hat{\theta}, \xi_e) = \sum_{r=1}^{n_{\text{resp}}} \sum_{s=1}^{n_{\text{resp}}} \frac{\mathbf{Q}_{r,e} \mathbf{Q}_{s,e}^T}{\Sigma_{y_{r,s,e}}} \quad (15)$$

Where $\mathbf{Q}_{r,e}$ represents the $n_{\text{sp}_e} \times n_u$ -dimensional dynamic sensitivity matrix of the r^{th} measurable variable evaluated at different sampling times of the e^{th} experiment and is written as:

$$\mathbf{Q}_{r,e} = \begin{bmatrix} \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_1} & \dots & \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_{n_u}} \leftarrow t_{e,1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_1} & \dots & \frac{\partial \hat{y}_{r,e}}{\partial \hat{\theta}_{n_u}} \leftarrow t_{e,n_{\text{sp}_e}} \end{bmatrix} \quad (16)$$

In Eq. (15), $\Sigma_{y_{r,s,e}}$ represents the $[r^{\text{th}}, s^{\text{th}}, e^{\text{th}}]$ element of the $n_{\text{resp}} \times n_{\text{resp}} \times n_{\text{exp}}$ -dimensional variance-covariance matrix of the residuals and is defined as:

$$\Sigma_{y_{r,s,e}} = \frac{1}{n_{sp_e}} \sum_{f=1}^{n_{sp_e}} [y_r(\xi_{e,f}) - \hat{y}_r(\xi_{e,f}, \hat{\theta})] [y_s(\xi_{e,f}) - \hat{y}_s(\xi_{e,f}, \hat{\theta})] \quad (17)$$

From $\Sigma_{\hat{\theta}}$ (Eq. (14)), the confidence interval of the u^{th} parameter is calculated by:

$$\Delta \hat{\theta}_u = t \sqrt{\Sigma_{\hat{\theta}_{uu}}} \quad (18)$$

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 u^{th} parameter, i.e., the u^{th} diagonal element of $\Sigma_{\hat{\theta}}$.

2.4.2.2 Model prediction uncertainty

The $n_{\text{resp}} \times n_{\text{resp}}$ -dimensional model prediction variance-covariance matrix, denoted as $\Sigma_{\hat{y}}(\hat{\theta}, \xi)$, quantifying the uncertainty on the model predictions of the experiment carried out under operating conditions ξ , can be calculated by propagating the uncertainty on the parameter estimates $\Sigma_{\hat{\theta}}$ (Eq. (14)), according to:

$$\Sigma_{\hat{y}}(\hat{\theta}, \xi) = \mathbf{G}^T \Sigma_{\hat{\theta}} \mathbf{G} \quad (19)$$

Where \mathbf{G} is the Jacobian matrix of $\hat{y}(\xi, \hat{\theta})$ and can be written as:

$$\mathbf{G} = \begin{bmatrix} \frac{\partial \hat{y}_1(\xi, \hat{\theta})}{\partial \hat{\theta}_1} & \dots & \frac{\partial \hat{y}_{n_{\text{resp}}}(\xi, \hat{\theta})}{\partial \hat{\theta}_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_1(\xi, \hat{\theta})}{\partial \hat{\theta}_{n_u}} & \dots & \frac{\partial \hat{y}_{n_{\text{resp}}}(\xi, \hat{\theta})}{\partial \hat{\theta}_{n_u}} \end{bmatrix} \quad (20)$$

The confidence intervals associated to the model predictions can then be calculated in a similar way as those associated to the parameter estimates.

2.5 Model-based experimental design

At the steps 11, 13 and 16 of the strategy, in order to identify the optimal reactor with its associated optimal operating conditions for respectively the DAEMD, the VEPO and the DAEPPi, model-based experimental design is performed. First, the optimal operating conditions of each considered reactor (e.g. q^{th} reactor) have to be identified by minimizing or maximizing the specific criterion according to Eq. (21). Then, the optimal reactor with the minimal or maximal criterion can be easily identified.

$$\xi_{R_q}^{\text{opt}} = \arg \begin{cases} \max_{\xi_{R_q} \in \Xi_{R_q}} C_{\text{DAEMD}}(\hat{\theta}, \xi_{R_q}) \\ \min_{\xi_{R_q} \in \Xi_{R_q}} C_{\text{VEPO}}(\hat{\theta}, \xi_{R_q}) \\ \min_{\xi_{R_q} \in \Xi_{R_q}} C_{\text{DAEPPI}}(\hat{\theta}, \xi_1, \dots, \xi_e, \dots, \xi_{n_{\text{exp}}}, \xi_{R_q}) \end{cases} \quad (21)$$

Where Ξ_{R_q} represents the field of research of operating conditions of the q^{th} reactor, C_{DAEMD} , C_{VEPO} and C_{DAEPPI} represent the criteria for respectively the DAEMD, the VEPO and the DAEPPi, more details about these criteria will be given below.

2.5.1 Criterion for DAEMD

The criterion for the DAEMD is defined as the highest pairwise criterion and can be represented by Eq. (22) if the simplest pairwise criterion proposed first by Hunter and Reiner (1965) is used. For more details about any other improved pairwise criteria, see, for example, Hunter and Reiner (1965), Buzzi-Ferraris and Forzati (1983), Donckel et al. (2009).

$$C_{\text{DAEMD}} = \max_{m,n} \left\{ \left[\hat{y}_m(\xi_{R_q}, \hat{\theta}) - \hat{y}_n(\xi_{R_q}, \hat{\theta}) \right] \left[\hat{y}_m(\xi_{R_q}, \hat{\theta}) - \hat{y}_n(\xi_{R_q}, \hat{\theta}) \right]^T \right\} \quad (22)$$

Where \hat{y}_m and \hat{y}_n are the responses predicted by the m^{th} and n^{th} SMUTs.

2.5.2 Criterion for VEPO

The criterion for the VEPO is defined by the user as:

$$C_{\text{VEPO}} = \frac{1}{\sqrt{n_{\text{perf}}}} \sum_p^{n_{\text{perf}}} J_{n_p}(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) \quad (23)$$

Where n_{perf} is the number of the performances to be optimized, J_{n_p} is the p^{th} normalized performance defined as:

$$J_{n_p}(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) = \begin{cases} 1 & \text{if } \frac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) - o_p}{v_p - o_p} > 1 \\ \frac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) - o_p}{v_p - o_p} & \text{if } 0 \leq \frac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) - o_p}{v_p - o_p} \leq 1 \\ 0 & \text{if } \frac{J_p(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_{R_q}) - o_p}{v_p - o_p} < 0 \end{cases} \quad (24)$$

Where \mathbf{J} is a vector function representing the performances, v_p and o_p are respectively the veto and objective values of the p^{th} performance defined by the user.

2.5.3 Criterion for DAEPPPI

The criterion for DAEPPPI is usually a statistical optimality criterion defined as:

$$C_{\text{DAEPPPI}} = h\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}\left(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_e, \dots, \boldsymbol{\xi}_{n_{\text{exp}}}, \boldsymbol{\xi}_{R_q}\right)\right) \quad (25)$$

Where h is an optimality function of the parameter variance-covariance matrix calculated based on the best currently available parameter estimates from all the existing experiments and the expected new experiment. The most common optimality functions are the determinant (D-optimality), the largest eigenvalue (E-optimality) and the trace (A-optimality). In this work, the criterion is the

average variance of model parameters, so h is the geometric mean of the diagonal elements defined as:

$$h\left(\Sigma_{\hat{\theta}}\left(\hat{\theta}, \xi_1, \dots, \xi_e, \dots, \xi_{n_{\text{exp}}}, \xi_{R_q}\right)\right) = \left[\prod_{u=1}^{n_u} \Sigma_{\hat{\theta}_{uu}}\left(\hat{\theta}, \xi_1, \dots, \xi_e, \dots, \xi_{n_{\text{exp}}}, \xi_{R_q}\right)\right]^{\frac{1}{n_u}} \quad (26)$$

In Eq. (25), while calculating $M_F\left(\hat{\theta}, \xi_{R_q}\right)$ referring to Eq. (15), the $n_{\text{resp}} \times n_{\text{resp}}$ -dimensional average variance-covariance matrix of the residuals, denoted as $\overline{\Sigma}_y$, is used, its $[r^{\text{th}}, s^{\text{th}}]$ element is defined as:

$$\overline{\Sigma}_{y_{r,s}} = \frac{1}{n_{\text{exp}}} \sum_{e=1}^{n_{\text{exp}}} \Sigma_{y_{r,s,e}} \quad (27)$$

2.6 Optimization algorithm

Both 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, concretely, genetic algorithm followed by interior-point algorithm.

2.7 Computer-aided software

The systematic model-development strategy for liquid-phase reaction systems is implemented as a computer-aided software using MATLAB R2014a. In order to reduce software development difficulty, following simplifications have been done without losing the fundamental functions:

- The reaction supernetwork is developed through empirical method;
- The reaction network under test is proposed by the user on the basis of his expertise;

- The kinetic orders cannot be considered as the parameters to be estimated;
- Model structure analysis is not integrated into the software;
- Model adequacy is qualitatively evaluated;
- If several SMUTs are available, the most suitable SMUT is directly identified.

The developed software provides user-friendly interfaces and guides the user for model development through the modules of the workflow (illustrated in Figure 3). Each module will be described in detail in Appendix 1.

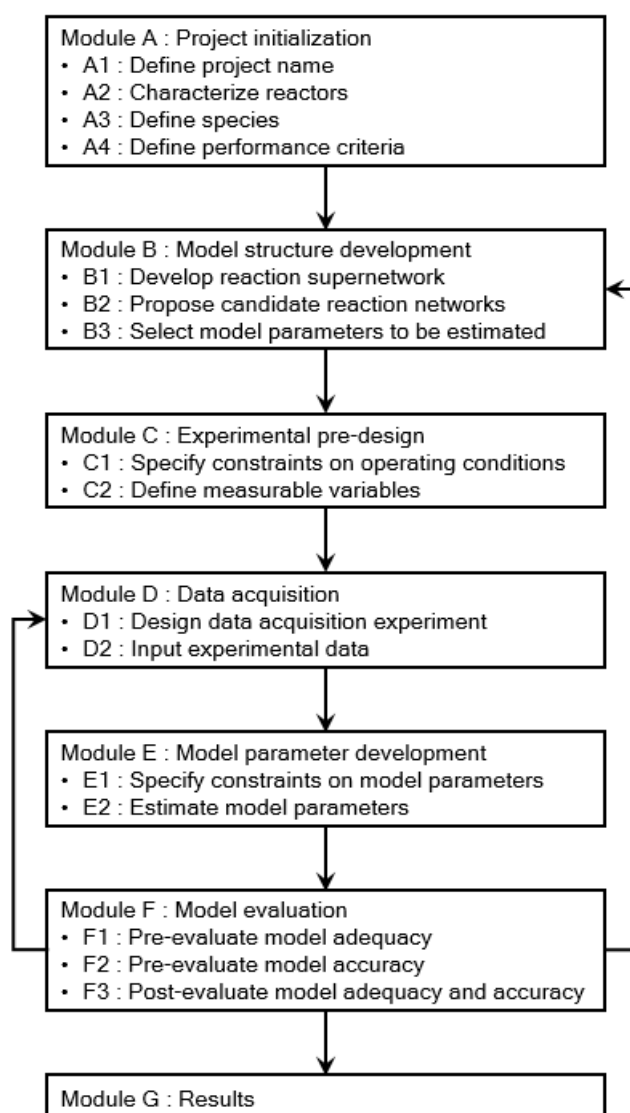


Figure 1: Software workflow.

3 Case study: NaOH-catalyzed ethanolysis of sunflower oil

To demonstrate the feasibility of the systematic model-development strategy and the practicability of the developed strategy-based software, the NaOH-catalyzed ethanolysis of sunflower oil has been selected as a case study for the low toxicity of all chemicals used and its easy handling.

3.1 Materials and methods

Commercial sunflower oil is purchased in a local supermarket. Its fatty acid composition (% by weight) is shown as follow: 11% saturated acids, 27% mono-unsaturated acids, 62% poly-unsaturated acids. Ethanol ($\geq 99.8\%$), hydrochloric acid (analytical quality, 37% solution in water) and sodium hydroxide ($\geq 98.0\%$, pellets) are supplied by Sigma-Aldrich.

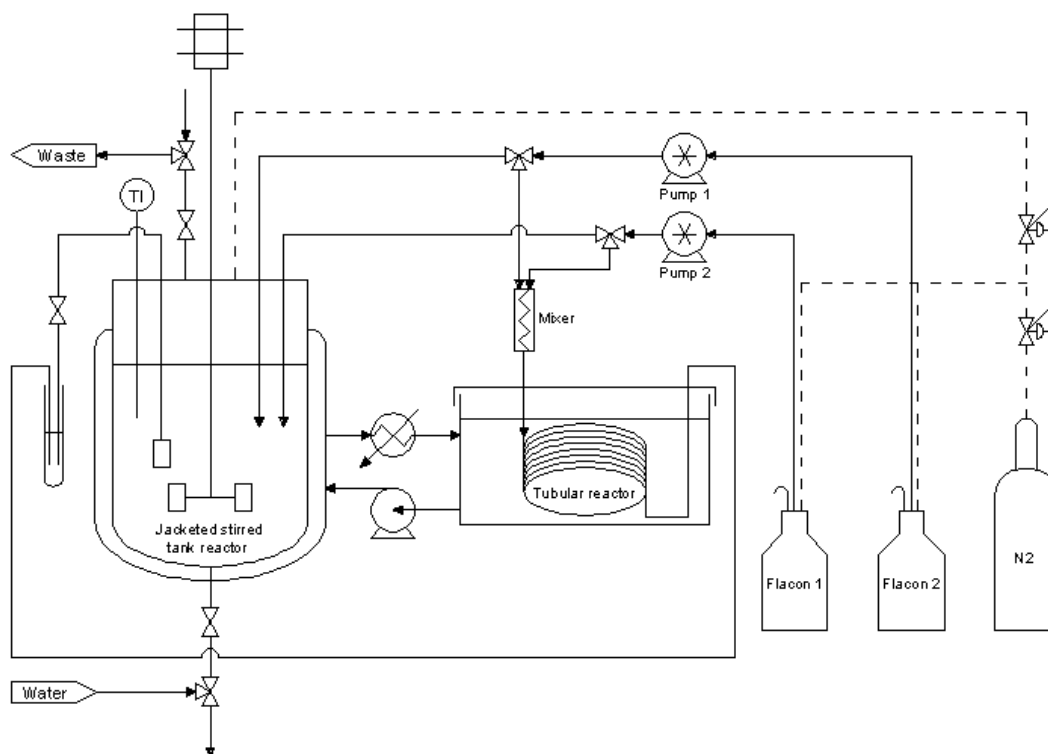


Figure 2: Experimental set-up.

The experimental set-up used (Mathieu et al., 2013), enabling to perform experiments in four different types of ideal reactors: BSTR, SBSTR, CSTR and CTR, is presented in Figure 4. The STR

is equipped with a jacket, whereas the TR is immersed in a bath. In this case study, only BSTR and CTR are considered.

For the experiments in the BSTR, oil is charged into the BSTR at the initial time and heated to the desired temperature, which is then maintained at a constant value. The mechanical stirring 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. The samples (2 mL) are removed from the reaction mixture at the designed sampling times.

For the experiments in the CTR, the CTR is dipped in a temperature-controlled water bath to maintain the desired temperature. The preheated solution 1 (i.e. mixture of ethanol and sodium hydroxide) and solution 2 (i.e. oil) are fed into the CTR through a T-mixer by two HPLC calibrated pumps. Samples (2 ml) are collected at the outlet of the CTR.

After being immediately quenched by adding the hydrochloric acid solution (3.7%, 2 ml) and vigorously shaken for 1 min, 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 mass compositions of glycerides and esters in the treated light phase are analyzed by gas chromatography analysis referring respectively to the standardized protocols EN 14103:2011 and EN 14105:2011.

3.2 Strategy application

This section presents the chronological application of the strategy to the ethanolysis case study (illustrated in Figure 5).

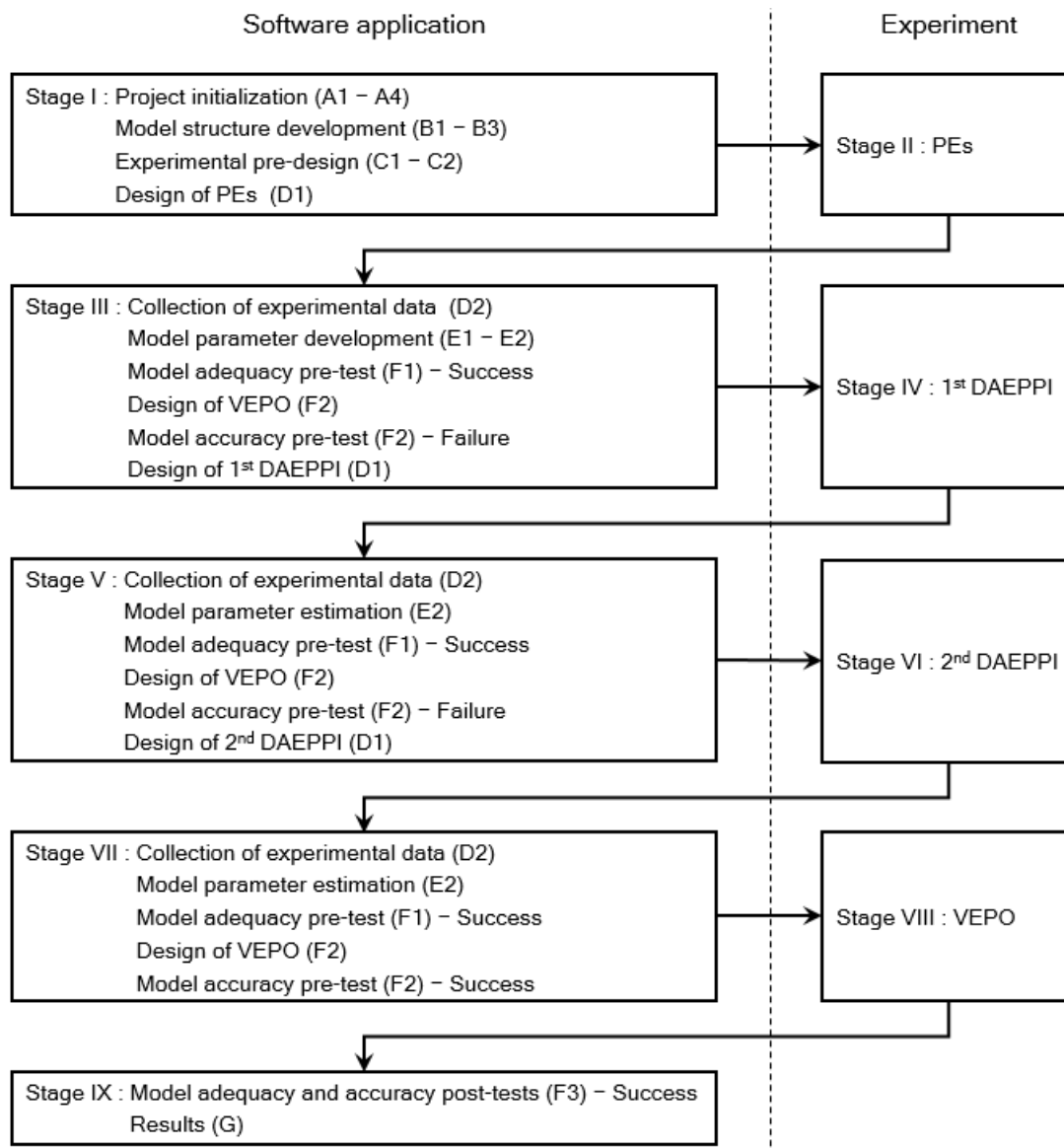


Figure 5: Operational sketch of strategy application for the ethanolysis case study.

3.2.1 Stage I: Software application

3.2.1.1 Submodule A1: Name project

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

3.2.1.2 Submodule A2: Characterize reactors

For this case study, all experiments are carried out under isothermal mode, therefore, the characteristics needed to be specified for the considered reactors are only their dimensions (shown in Table 3).

Table 3: Dimensions for each considered reactor.

Reactor	BSTR	CTR
Internal diameter (cm)	10	0.16
Height for BSTR / Length for CTR (cm)	10	635

3.2.1.3 Submodule A3: Define species

To simplify the reaction system of this case study, the following assumptions are made:

- Sunflower oil consists only of the triglycerides where the three fatty acids are identical;
- All mono-unsaturated triglycerides are represented by tri-olein, since, according to the chromatogram of esters (see Figure 10, Appendix 2), compared with the amount of ethyl oleate generated from tri-olein, the amounts of other mono-unsaturated fatty acid ethyl esters such as ethyl palmitoleate, ethyl margaroleate and ethyl eicosonate can be neglected;
- By analogy, all poly-unsaturated triglycerides are represented by tri-linolein;
- By analogy, all saturated triglycerides are represented by tri-palmitin.

Thus, all possible species in the reaction system are: Tri-Olein (TO), Tri-Linolein (TL), Tri-Palmitin (TP), Di-Olein (DO), Di-Linolein (DL), Di-Palmitin (DP), Mono-Olein (MO), Mono-Linolein (ML), Mono-Palmitin (MP), Ethyl Oleate (EO), Ethyl Linoleate (EL), Ethyl Palmitate (EP), G (Glycerol), Ethanol (E) and Sodium hydroxide (NaOH).

3.2.1.4 Submodule A4: Define performance criteria

Table 4 specifies the performance criteria used for this case study. The objective values of the conversion of the considered triglycerides as well as the total selectivity of Fatty Acid Ethyl Esters (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.

Table 4: Performance criteria.

Performance	Variable	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%

3.2.1.5 Submodule B1: Develop reaction supernetwork

On the basis of the biphasic reaction network of the homogeneous alkali-catalyzed alcoholysis of oil (see Section 2.2.1), in which only one triglyceride is considered, the reaction supernetwork for this case study (illustrated in Table 5), comprising 18 chemical reactions and 12 mass transfers, is developed by:

- taking into consideration three triglycerides (i.e. TO, TL and TP);
- assuming that ethanol and glycerol are present only in the polar phase (I).

Table 5: Reaction supernetwork for the case study.

N°	Reaction	N°	Reaction	N°	Reaction
1	$TO(I) + E(I) \xrightarrow{Cat} EO(I) + DO(I)$	11	$TL(I) + E(I) \xrightarrow{Cat} EL(I) + DL(I)$	21	$TP(I) + E(I) \xrightarrow{Cat} EP(I) + DP(I)$
2	$DO(I) + E(I) \xrightarrow{Cat} EO(I) + MO(I)$	12	$DL(I) + E(I) \xrightarrow{Cat} EL(I) + ML(I)$	22	$DP(I) + E(I) \xrightarrow{Cat} EP(I) + MP(I)$

3	$\text{MO(I)} + \text{E(I)} \xrightarrow{\text{Cat.}} \text{EO(I)} + \text{G(I)}$	13	$\text{ML(I)} + \text{E(I)} \xrightarrow{\text{Cat.}} \text{EL(I)} + \text{G(I)}$	23	$\text{MP(I)} + \text{E(I)} \xrightarrow{\text{Cat.}} \text{EP(I)} + \text{G(I)}$
4	$\text{EO(I)} + \text{DO(I)} \xrightarrow{\text{Cat.}} \text{TO(I)} + \text{E(I)}$	14	$\text{EL(I)} + \text{DL(I)} \xrightarrow{\text{Cat.}} \text{TL(I)} + \text{E(I)}$	24	$\text{EP(I)} + \text{DP(I)} \xrightarrow{\text{Cat.}} \text{TP(I)} + \text{E(I)}$
5	$\text{EO(I)} + \text{MO(I)} \xrightarrow{\text{Cat.}} \text{DO(I)} + \text{E(I)}$	15	$\text{EL(I)} + \text{ML(I)} \xrightarrow{\text{Cat.}} \text{DL(I)} + \text{E(I)}$	25	$\text{EP(I)} + \text{MP(I)} \xrightarrow{\text{Cat.}} \text{DP(I)} + \text{E(I)}$
6	$\text{EO(I)} + \text{G(I)} \xrightarrow{\text{Cat.}} \text{MO(I)} + \text{E(I)}$	16	$\text{EL(I)} + \text{G(I)} \xrightarrow{\text{Cat.}} \text{ML(I)} + \text{E(I)}$	26	$\text{EP(I)} + \text{G(I)} \xrightarrow{\text{Cat.}} \text{MP(I)} + \text{E(I)}$
7	$\text{TO(I)} \rightarrow \text{TO(II)}$	17	$\text{TL(I)} \rightarrow \text{TL(II)}$	27	$\text{TP(I)} \rightarrow \text{TP(II)}$
8	$\text{DO(I)} \rightarrow \text{DO(II)}$	18	$\text{DL(I)} \rightarrow \text{DL(II)}$	28	$\text{DP(I)} \rightarrow \text{DP(II)}$
9	$\text{MO(I)} \rightarrow \text{MO(II)}$	19	$\text{ML(I)} \rightarrow \text{ML(II)}$	29	$\text{MP(I)} \rightarrow \text{MP(II)}$
10	$\text{EO(I)} \rightarrow \text{EO(II)}$	20	$\text{EL(I)} \rightarrow \text{EL(II)}$	30	$\text{EP(I)} \rightarrow \text{EP(II)}$

3.2.1.6 Submodule B2: Propose candidate reaction networks

For this case study, two reaction networks under test (illustrated in Table 6) are initially proposed from the reaction supernetwork developed above. Then, the same subsequent steps such as model parameter development and model adequacy pre-test for both the RN1 and the RN2 will be carried out below. For concision of the present paper, this procedure will be solely illustrated for the RN1.

Table 6: Reaction networks under test.

N°	Reaction network	Reactions taken into consideration
1	Reduced pseudo-homogeneous reaction network (RN1)	1-6, 11-16, 21-26
2	Biphasic reaction network (RN2)	1-30

3.2.1.7 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 for 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 8).

3.2.1.8 Submodule C1: Specify constraints on operating conditions

The bounds on the practical operating conditions and the corresponding operating conditions in general form are shown respectively in Table 7 and Table 8.

Table 7: 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 ^a	200
Mass of NaOH in solution 1	$m_{NaOH,1}$	g	0.2 – 3 ^b	0.0380 – 0.7605 ^c
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 – 78 ^d
Flow rate of feed 1	F_1	mL/min		2 – 20
Temperature of feed 2	T_2	°C		30 – 78 ^d
Flow rate of feed 2	F_2	mL/min		4
Experiment duration	τ	s	30 – 8000	

^a Ethanol-to-oil molar ratio corresponds to 3 – 24.

^b NaOH amount corresponds to 0.1 – 1.5 g per 100 g of oil.

^c Concentration of NaOH in solution 1 corresponds to 0.15 – 3 g/L.

^d Boiling point of ethanol.

Table 8: 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 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	
Initial load of ethanol	n_E^{int}	mol	0.6837 – 5.4696	
Initial load of NaOH	n_{NaOH}^{int}	mol	0.0050 – 0.0750	
Temperature of feed	T^{in}	K		303.15 – 351.15
Flow rate of feed	F^{in}	m ³ /s		10^{-7} – 4×10^{-7}
Concentration of TO in feed	c_{TO}^{in}	mol/m ³		46.4446 – 185.7786
Concentration of TL in feed	c_{TL}^{in}	mol/m ³		107.3841 – 429.5365
Concentration of TP in feed	c_{TP}^{in}	mol/m ³		19.7246 – 78.8983
Concentration of ethanol in feed	c_E^{in}	mol/m ³		5708.9272 – 14272.3180
Concentration of NaOH in feed	c_{NaOH}^{in}	mol/m ³		1.25 – 62.5

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, which are constructed based on the oil composition constraints (Eqs. (28)-(30)) and the volume balance (Eq. (31)) by neglecting the contribution of NaOH to liquid volume.

$$\frac{c_{TO}^{in}}{c_{TO}^{in} + c_{TL}^{in} + c_{TP}^{in}} = \chi_{TO} \quad (28)$$

$$\frac{c_{TL}^{in}}{c_{TO}^{in} + c_{TL}^{in} + c_{TP}^{in}} = \chi_{TL} \quad (29)$$

$$\frac{c_{TP}^{in}}{c_{TO}^{in} + c_{TL}^{in} + c_{TP}^{in}} = \chi_{TP} \quad (30)$$

$$\frac{v_{TO}c_{TO}^{in} + v_{TL}c_{TL}^{in} + v_{TP}c_{TP}^{in}}{v_Ec_E^{in}} = \frac{F_2}{F^{in} - F_2} \quad (31)$$

Where χ_{TO} , χ_{TL} and χ_{TP} are the molar fractions of TO, TL and TP in the used oil, F_2 is the practical flow rate of solution 2 (i.e. oil), equal to 4 ml/min (see Table 7).

3.2.1.9 Submodule C2: Define measurable variables

Figure 6 illustrates how to define the measurable variables for this case study. 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, since according to the chromatogram of glycerides (see Figure 11, Appendix 2), glyceride amounts could not be distinctively determined due to the superposition of their peaks. Given these variables, the parameters involved in the model structure constructed based on the RN1 are identifiable (Likozar and Levec, 2014b).

The interface shows the 'Measurable variables' tab. It includes a list of available variables on the left and a list of selected measurable variables on the right. A table on the right shows the selection of 3 unconventional measurable variables (UMV) for MGs, DGs, and TGs.

	Name	C_TO...	C_TL...	C_TP...	C_DO...	C_DL...	C_DP...	C_MO...	C_ML...	C_MP...	C_EO...	C_EL...	C_EP...	C...
1st UMV	MGs	0	0	0	0	0	0	1	1	1	0	0	0	
2nd UMV	DGs	0	0	0	1	1	1	0	0	0	0	0	0	
3rd UMV	TGs	1	1	1	0	0	0	0	0	0	0	0	0	

Figure 6: Interface for the selection of the measurable variables.

3.2.1.10 Submodule D1: Design data acquisition experiments

The initial data acquisition experiments (i.e. the PEs) are designed by user, and the experiment design results are shown in Table 9 in terms of practical operating variables, from which the operating conditions in general form are derived and input into the software.

Experiments under different operating variables can be carried out during one single run of the set-up: for instance, 8 experiments under the same initial operating variables with different experiment durations (i.e. sampling times) using the BSTR are carried out in the 1st run. For the PEs using 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 9: Practical operating variables of the PEs for the case study.

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

* The residence time in the CTR is equal to the sampling time.

3.2.2 Stage II: PEs

The PEs are performed according to the practical operating conditions derived from the practical operating variables obtained at D1 (see Section 3.2.1.10).

3.2.3 Stage III: Software application

3.2.3.1 Submodule D2: Input experimental data

The measured responses of the PEs 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 F2 (see Sections 3.2.3.5, 3.2.5.4 and 3.2.7.4) and F3 (see Section 3.2.9.1), has to be defined here. For this case study, that for the measurable variable of type concentration is set as 5 mol/m³.

3.2.3.2 Submodule E1: Specify constraints on model parameters

In order that the estimates of A_j and E_{a_j} , derived from $\log k_{j,T_{\max}}$ and $\log k_{j,T_{\min}}$ according to Eqs. (11)-(12), 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 initial bounds specified between -10 and -6 as well as the linear inequality constraints (Eqs. (32)-(33)).

$$\log k_{j,T^{\max}} - \log k_{j,T^{\min}} > 0.4 \quad (32)$$

$$\log k_{j,T^{\max}} - \log k_{j,T^{\min}} < 1.4 \quad (33)$$

3.2.3.3 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 10, in which the results of parameter estimations performed using different experimental data updated after the 1st DAEPPi and the 2nd DAEPPi are also included and will be discussed later (see Sections 3.2.5.2 and 3.2.7.2). Model parameters are obviously inaccurate, since the confidence intervals for certain parameters (with gray shading) are very wide. Among these parameters, the confidence intervals for $\log(k_{1,T^{\min}})$, $\log(k_{1,T^{\max}})$, $\log(k_{21,T^{\min}})$, $\log(k_{21,T^{\max}})$, $\log(k_{4,T^{\min}})$, $\log(k_{4,T^{\max}})$, $\log(k_{24,T^{\min}})$, $\log(k_{24,T^{\max}})$, namely, the parameters to characterize respectively the reactions N°1, 21, 4, 24, 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 (Eq. (13)), which will be used for pre-evaluating model adequacy at F1 (see Sections 3.2.3.4, 3.2.5.3 and 3.2.7.3);
- the average variance of model parameters derived from the parameter variance-covariance matrix (Eq. (14)), which will be used to calculate the model prediction variance-covariance matrix (Eq. (19)) and further the confidence intervals associated to the model predictions, used for pre-evaluating model accuracy at F2 (see Sections 3.2.3.5, 3.2.5.4 and 3.2.7.4).

Table 10: Evolution of parameter estimations.

N° parameter estimation (Strategy application stage)	1 (Stage III)	2 (Stage V)	3 (Stage VII)
Data used for parameter estimation	PEs	PEs + 1 st DAEPPi	PEs + 1 st DAEPPi + 2 nd DAEPPi
Parameter estimation criterion	24.3357	24.0603	23.5021
Average variance of model parameters	1.6669	0.8850	0.0262
$\log(k_{1,Tmin})$	-5.9969 ± 15.8021	-6.0267 ± 35.8940	-6.0237 ± 0.0522
$\log(k_{1,Tmax})$	-5.3552 ± 65.5235	-5.3603 ± 60.9098	-5.3612 ± 0.1397
$\log(k_{2,Tmin})$	-7.3507 ± 1.7846	-7.3408 ± 3.2309	-7.3349 ± 0.6286
$\log(k_{2,Tmax})$	-6.2215 ± 3.3151	-6.2314 ± 3.6027	-6.2310 ± 0.7899
$\log(k_{3,Tmin})$	-8.1889 ± 0.1435	-8.1840 ± 0.0735	-8.1841 ± 0.0724
$\log(k_{3,Tmax})$	-7.5314 ± 0.0939	-7.5315 ± 0.0200	-7.5312 ± 0.0436
$\log(k_{11,Tmin})$	-6.9258 ± 3.5347	-6.9216 ± 0.5593	-6.9323 ± 0.0262
$\log(k_{11,Tmax})$	-6.1192 ± 5.7219	-6.0949 ± 0.3254	-6.1095 ± 0.2277
$\log(k_{12,Tmin})$	-7.1161 ± 1.4514	-7.1472 ± 0.4771	-7.1545 ± 0.0667
$\log(k_{12,Tmax})$	-6.6423 ± 2.8467	-6.6458 ± 0.2083	-6.6416 ± 0.1352
$\log(k_{13,Tmin})$	-7.1940 ± 5.2812	-7.1879 ± 0.2459	-7.1871 ± 0.3508
$\log(k_{13,Tmax})$	-6.0631 ± 9.8274	-6.0570 ± 0.2125	-6.0643 ± 0.3777
$\log(k_{21,Tmin})$	-6.2790 ± 284.1310	-6.3043 ± 64.0053	-6.3030 ± 0.0253
$\log(k_{21,Tmax})$	-5.4899 ± 515.8986	-5.5350 ± 70.2819	-5.5289 ± 0.7176
$\log(k_{22,Tmin})$	-7.2657 ± 26.4442	-7.2684 ± 8.3569	-7.2763 ± 0.6396
$\log(k_{22,Tmax})$	-6.3470 ± 46.5832	-6.3314 ± 8.3906	-6.3381 ± 0.9003
$\log(k_{23,Tmin})$	-8.0997 ± 1.4453	-8.1023 ± 0.5528	-8.1066 ± 0.4064
$\log(k_{23,Tmax})$	-6.7961 ± 2.0084	-6.8208 ± 0.8792	-6.8249 ± 0.3120
$\log(k_{4,Tmin})$	-6.5558 ± 13.0749	-6.5906 ± 47.3796	-6.6000 ± 0.1319
$\log(k_{4,Tmax})$	-6.0975 ± 63.4358	-6.1184 ± 72.6217	-6.1196 ± 0.3524
$\log(k_{5,Tmin})$	-7.1936 ± 2.2469	-7.1792 ± 2.9003	-7.1705 ± 0.6594
$\log(k_{5,Tmax})$	-5.8815 ± 2.4924	-5.8888 ± 3.1192	-5.8928 ± 0.9617
$\log(k_{6,Tmin})$	-9.9672 ± 7.3913	-9.9474 ± 0.5112	-9.9499 ± 1.8693
$\log(k_{6,Tmax})$	-8.9398 ± 3.8995	-8.9327 ± 3.1633	-8.9371 ± 2.9321

$\log(k_{14,Tmin})$	-5.8518 ± 5.1718	-5.8823 ± 1.2472	-5.8829 ± 0.1716
$\log(k_{14,Tmax})$	-5.3665 ± 8.9354	-5.3812 ± 0.5149	-5.3867 ± 0.5290
$\log(k_{15,Tmin})$	-8.4806 ± 1.4640	-8.4375 ± 0.1691	-8.4286 ± 0.0791
$\log(k_{15,Tmax})$	-7.6121 ± 2.6692	-7.5846 ± 0.2432	-7.5766 ± 0.2474
$\log(k_{16,Tmin})$	-7.1416 ± 5.3193	-7.1331 ± 0.3532	-7.1359 ± 0.3453
$\log(k_{16,Tmax})$	-6.5939 ± 9.9045	-6.5865 ± 0.3624	-6.5912 ± 0.3685
$\log(k_{24,Tmin})$	-5.9541 ± 255.6997	-5.9635 ± 64.0622	-5.9632 ± 1.0698
$\log(k_{24,Tmax})$	-5.4827 ± 466.5295	-5.4797 ± 76.2883	-5.4758 ± 0.5316
$\log(k_{25,Tmin})$	-6.1097 ± 19.9058	-6.1088 ± 9.7819	-6.1121 ± 0.7501
$\log(k_{25,Tmax})$	-5.6166 ± 37.2714	-5.6134 ± 8.0466	-5.6068 ± 0.3030
$\log(k_{26,Tmin})$	-9.1929 ± 4.8020	-9.1772 ± 1.9314	-9.1845 ± 2.1089
$\log(k_{26,Tmax})$	-8.7842 ± 6.2776	-8.7568 ± 3.8840	-8.7350 ± 2.8705

New 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.

3.2.3.4 Submodule F1: Model adequacy pre-test

The SMUT[RN1,PEs] has been proven qualitatively feasible, since the measured response points are reasonably scattered around or on the corresponding predicted curves for the BSTR (see Figure 7) and the CTR (see Figure 8). The esters (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. 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.

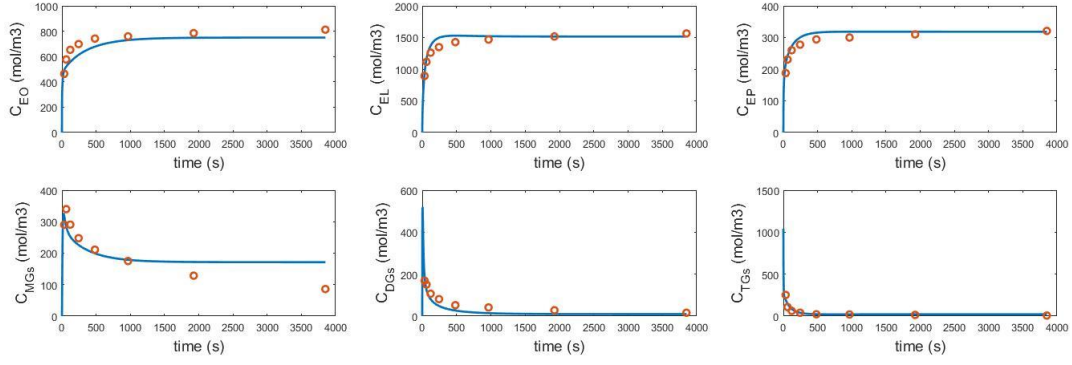


Figure 7: Evolution of the various measurable variables for the 1st series of PEs carried out at 60°C, NaOH concentration of 0.50 wt.% (per oil weight), ethanol-to-oil molar ratio of 6:1, using BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: measured responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).

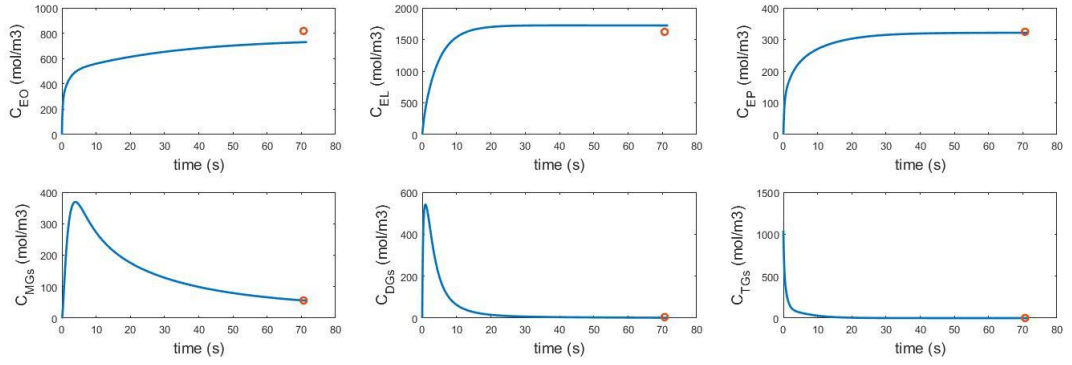


Figure 8: Evolution of the various measurable variables for the 1st experiment in the 5th series of PEs carried out at 60°C, NaOH concentration of 2.34 wt.% (per oil weight), ethanol-to-oil molar ratio of 28:1, using CTR (continuous lines: responses predicted by the SMUT[RN1,PEs]; singular points: measured responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{MGs}=C_{MO}+C_{ML}+C_{MP}$).

After model adequacy pre-test is implemented for the SMUT[RN1,PEs], the same test for the SMUT[RN2,PEs] is carried out. Both the SMUT[RN1,PEs] and the SMUT[RN2,PEs] pass 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. (13)) for the SMUT[RN2,PEs] (27.0638) is larger than that of the SMUT[RN1,PEs] (24.3357).

3.2.3.5 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is then performed on the basis of the SMUT[RN1,PEs], i.e. the most suitable SMUT identified at F1 (see Section 3.2.3.4). The result is shown in the 1st column of Table 11, in which the results of model-based experimental design for the VEPO performed based on the SMUTs updated after the 1st DAEPPPI and the 2nd DAEPPPI are also included and will be discussed later (see Sections 3.2.5.4 and 3.2.7.4). The optimal reactor (i.e. the BSTR) and the optimal operating conditions (i.e. long experiment duration, near to maximal ethanol-to-oil molar ratio, near to maximal reaction temperature and high NaOH amount), which are quite different from those of the PEs, are identified for the VEPO enabling 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 D2 (see Section 3.2.3.1).

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 and 5.9789, are larger than the threshold. Therefore, there are two possible choices for the next step: model structure development (B2) or design of the DAEPPPI (D1). Considering that the SMUT[RN1,PEs] passes model adequacy pre-test, we decide to design and perform a DAEPPPI and expect that the model parameter precision as well as the model prediction precision increases after the 1st DAEPPPI so that the SMUT[RN1,PEs+1st DAEPPPI] passes model accuracy pre-test.

Table 11: Results of the optimal experimental design for the VEPO.

N° optimal experimental design	1	2	3
SMUT used	SMUT[RN1,PEs]	SMUT[RN1,PEs+1 st DAEPPPI]	SMUT[RN1,PEs+1 st DAEPPPI+2 nd DAEPPPI]

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. (23))	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

3.2.3.6 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 et al., 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 12) indicates that the 1st DAEPPi will be carried out using the BSTR at long experiment durations and near to maximal ethanol-to-oil molar ratio (inhibiting inverse reactions). It is not easy to interpret quantitatively why the software proposes such operating conditions, since the interaction among the reactions defined in the RN1 is complex.

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 reactions with relatively long characteristic times. It is noticed that the inverse reactions contain more such reactions in comparison to the forward reactions (see Table 10). Data acquisition efficiency is related to the sensitivity coefficient, which is defined as the time-varying derivatives of model responses with respect to the model parameters (see Eq. (16)). The parameters involved in the inverse 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%).

Table 12: Results of model-based experimental design for the 1st DAEPPi.

Criterion for DAEPPi (see Eq. (25))	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
$n_{\text{TO}}^{\text{int}}$ [mol]	0.0610	0.0610
$n_{\text{TL}}^{\text{int}}$ [mol]	0.1410	0.1410
$n_{\text{TP}}^{\text{int}}$ [mol]	0.0259	0.0259
$n_{\text{E}}^{\text{int}}$ [mol]	5.4324	5.4324
$n_{\text{NaOH}}^{\text{int}}$ [mol]	0.0197	0.0197

3.2.4 Stage IV: First DAEPPPI

The 1st DAEPPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at D1 (see Section 3.2.3.6).

3.2.5 Stage V: Software application

3.2.5.1 Submodule D2: Input experimental data

The measured responses of the 1st DAEPPPI (shown in Table 18, Appendix 3) are input into the software.

3.2.5.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs and the 1st DAEPPPI. The current SMUT is updated to SMUT[RN1,PEs+1st DAEPPPI]. It has been noticed that model parameter precision, quantified by the average variance of model parameters, increases by 46.91%. Meanwhile, not every parameter confidence interval of the SMUT[RN1,PEs+1st DAEPPPI] (see the 2nd column of Table 10) decreases in comparison to that of the SMUT[RN1,PEs] (see the 1st column of Table 10), since the DAEPPPI is designed to reduce the average variance of model parameters, not the variance of each parameter. For example, the parameters associated to reaction (1) (see Table 10) 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.

3.2.5.3 Submodule F1: Model adequacy pre-test

The SMUT[RN1,PEs] passes model adequacy pre-test, but this does not mean that SMUT[RN1,PEs+1st DAEPPPI], 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 DAEPPPI]. Here, the SMUT[RN1,PEs+1st DAEPPPI] passes model adequacy pre-test, since its parameter estimation criterion (24.0603) is smaller than that of SMUT[RN1,PEs] (24.3357) (see Table 10).

3.2.5.4 Submodule F2: Model accuracy pre-test

Model-based experimental design for the VEPO is performed on the basis of the SMUT[RN1,PEs+1st DAEPPPI]. The result is shown in the 2nd column of Table 11. 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} is still larger than the threshold. In view of the fact that the predicted error of each measurable variable decreases after the 1st DAEPPPI, we decide to do one more DAEPPPI and expect that the SMUT[RN1,PEs+1st DAEPPPI+2nd DAEPPPI] passes model accuracy pre-test.

3.2.5.5 Submodule D1: Design data acquisition experiment

Model-based experimental design for the DAEPPPI is performed. The result (shown in Table 13) indicates that the 2nd DAEPPPI will be carried out using the CTR (different from the reactor used in the 1st DAEPPPI) under the minimal reaction temperature at the maximal and minimal ethanol-to-oil molar ratios. In this DAEPPPI, the experiments are carried out at two extreme ethanol-to-oil molar ratios. We consider 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, this DAEPPPI (using the CTR) will provide more information at the beginning of the synthesis, enabling to complete the available information for parameter estimation.

Table 13: Result of model-based experimental design for the 2nd DAEPPPI.

Criterion for DAEPPPI (see Eq. (25))	0.6752	
Optimal reactor	CTR	
N ^o experiment	1	2
T ⁱⁿ [K]	303.1500	303.1500
F ⁱⁿ [m ³ /s]	4.0000×10 ⁻⁷	1.0000×10 ⁻⁷
c _{TO} ⁱⁿ [mol/m ³]	46.4457	185.6920
c _{TL} ⁱⁿ [mol/m ³]	107.3841	429.3261
c _{TP} ⁱⁿ [mol/m ³]	19.7342	78.8983
c _E ⁱⁿ [mol/m ³]	1.4265×10 ⁴	5.7138×10 ³
c _{NaOH} ⁱⁿ [mol/m ³]	2.6151	2.2506

3.2.6 Stage VI: Second DAEPPPI

The 2nd DAEPPPI is performed according to the practical operating conditions derived from the operating conditions in general form obtained at D1 (see Section 3.2.5.5).

3.2.7 Stage VII: Software application

3.2.7.1 Submodule D2: Input experimental data

The measured responses of the 2nd DAEPPPI (shown in Table 18, Appendix 3) are input into the software.

3.2.7.2 Submodule E2: Estimate model parameters

Parameter estimation is performed using the experimental data of the PEs, the 1st DAEPPPI and the 2nd DAEPPPI. The current SMUT is updated to SMUT[RN1,PEs+1st DAEPPPI+2nd DAEPPPI]. 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 10).

3.2.7.3 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 former one (24.0603) (see Table 10).

3.2.7.4 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 11. 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.

3.2.8 Stage VIII: VEPO

The VEPO is performed according to the practical operating conditions derived from the operating conditions in general form obtained at F2 (see Section 3.2.7.4).

3.2.9 Stage IX: Software application

3.2.9.1 Submodule F3: Model adequacy and accuracy post-tests

Figure 9 presents the comparison between the measured responses of the VEPO (shown in Table 18, Appendix 3) 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), therefore, the SMUT[RN1,PE+1st DAEPPPI+2nd DAEPPPI] is validated.

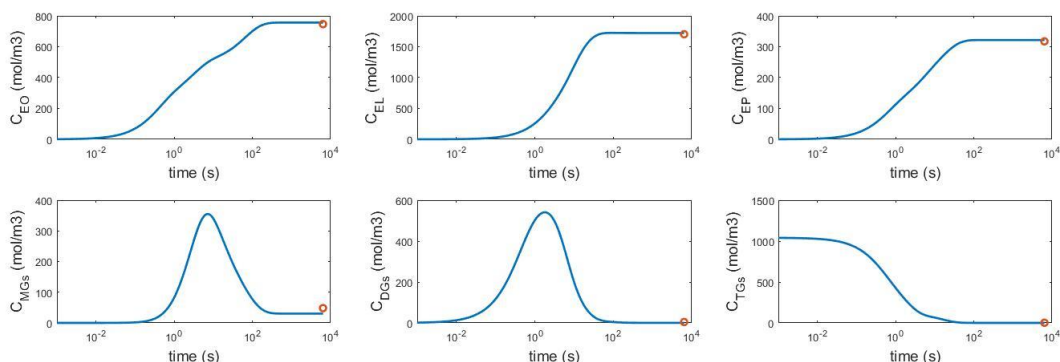


Figure 9: Concentration profiles for the VEPO using the BSTR (continuous lines: responses predicted by the SMUT[RN1,PEs+1st DAEPPPI+2nd DAEPPPI]; singular points: measured responses; $C_{MGs}=C_{TO}+C_{TL}+C_{TP}$; $C_{DGs}=C_{DO}+C_{DL}+C_{DP}$; $C_{TGs}=C_{MO}+C_{ML}+C_{MP}$).

3.2.9.2 Module G: Results

The final results about the validated SMUT (i.e. the SMUT[RN1,PEs+1st DAEPPPI+2nd DAEPPPI]) and the optimal reactor (i.e. the BSTR) with its associated optimal operating conditions (see the 3rd column of Table 12) are displayed and stored in the file created in A1.

3.3 Results

Table 14 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 parameters 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° 23 and 26. Further assessment about the rationality of those kinetic parameters characterizing the reactions N° 23 and 26 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 15. It is observed that each parameter estimate is comprised in its corresponding

confidence interval derived from literature, even if $\log(k_{26,T_{\max}})$ 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 14: Comparison of pre-exponential factors and activation energies obtained from this work and from literature.

N° Reaction	This work ^a		Likozar and Levec, 2014b ^b	
	Pre-exponential factor, A [m ⁶ /(mol ² ·s)]	Activation energy, Ea [J/mol]	Pre-exponential factor, A [m ⁶ /(mol ² ·s)]	Activation energy, Ea [J/mol]
21	2.29×10 ⁻¹	3.29×10 ⁴	4.50×10 ⁻¹	5.50×10 ⁴
22	3.87×10 ⁻¹	3.98×10 ⁴	7.50×10 ⁻¹	4.92×10 ⁴
23	1.86×10 ¹	5.44×10 ⁴	2.42×10 ⁰	4.80×10 ⁴
24	4.00×10 ⁻³	2.07×10 ⁴	2.00×10 ⁻³	3.80×10 ⁴
25	3.84×10 ⁻³	2.15×10 ⁴	2.17×10 ⁻³	3.10×10 ⁴
26	1.27×10 ⁻⁶	1.91×10 ⁴	1.17×10 ⁻³	3.42×10 ⁴

^a The oil used is sunflower oil; the catalyst used is NaOH; the practical operating variables are: (i) reaction temperature of 45, 60 °C, (ii) ethanol-to-oil molar ratio of 6:1, 12:1, (iii) NaOH amount of 0.5, 1.0 g per 100 g of oil.

^b The oils used are canola, palm, peanut, soybean and sunflower oils; the catalysts used are KOH and NaOH, the practical operating variables are: (i) reaction temperature of 40, 50, 60 °C, (ii) ethanol-to-oil molar ratio of 6:1, (iii) NaOH amount of 0.8 g per 100 g of oil.

Table 15: Check of parameter estimates in this work and parameter confidence intervals derived from literature.

N° Reaction	Estimate in this work		Confidence interval derived from Likozar and Levec, 2014b	
	log(k _{Tmin}) [m ⁶ /(mol ² ·s)]	log(k _{Tmax}) [m ⁶ /(mol ² ·s)]	log(k _{Tmin}) [m ⁶ /(mol ² ·s)]	log(k _{Tmax}) [m ⁶ /(mol ² ·s)]
23	-8.1066	-6.8249	[-8.6084 , -7.5659]	[-7.3733 , -6.4750]
26	-9.1845	-8.7350	[-9.6551 , -8.9242]	[-8.7352 , -8.1039]

4 Conclusion

A systematic strategy for model development has been successfully developed and implemented as a computer-aided software, which guides the user through a workflow to identify simultaneously a

set of adequate and accurate models as well as the optimal reactor with its associated optimal operating conditions for a given liquid-phase synthesis. For model identification, validation and refining, model-based experimental design is performed by taking into consideration several reactors. Model structure development is based on the reaction proposed within the reaction supernetwork containing all feasible chemical reactions and mass transfers, which allows the strategy to be applied for the heterogeneous reaction system. The application of the strategy has been highlighted with an experimental case study involving NaOH-catalyzed ethanolysis of sunflower oil.

The strategy has also the potential to be applied in other reaction systems. These applications differ in model structure development, but the subsequent steps are identical. For example, the Eley-Rideal mechanism consists of three steps: (i) adsorption of one reactant on the vacant active site of the catalyst; (ii) chemical reaction between the adsorbed reactant and another reactant to produce the adsorbed product; (iii) desorption of the product from the catalyst surface. All the above steps can be described in the same manner as conventional chemical reactions by considering the adsorbed reactant/product as pseudo-species, referring to the representation of mass transfer between two phases in this study. The model structure for a given gas-solid catalytic reaction system in a given reactor can be constructed after: (i) developing of reaction network within the reaction supernetwork including all feasible chemical reactions, adsorptions and desorptions; (ii) defining the rate law for each of the involved chemical reactions, adsorptions and desorptions; (iii) indicating the (characteristics of) given reactor type. Therefore, in order that the strategy as well as the strategy-based software can be applied in more reaction systems, it is necessary to expand the database of reaction types, rate laws, and reactor models.

Furthermore, there are a number of limitations to the computer-aided software in its current version. Following the failure of the SMUTs, a new reaction network used for model structure development is currently proposed through empirical method by the user, but a better prediction quality of the new candidate set of model structures is not guaranteed without model parameter estimation, which takes much computation time. Therefore, in the future, more attention should be paid to how to identify the best reaction network with minimum computation time and minimum human intervention, namely how to develop a promising approach consisting of (i) developing reaction network through theoretical method by optimizer, (ii) estimating model parameters through stepwise method, which decreases significantly computation time, (iii) identifying the best model structure through direct method. Software implementation of the functions, enabling the user to customize the performances (such as the selectivity of a product with respect to a reactant, i.e. the product of the current conventional performances) and personalize the linear or nonlinear equality or inequality constraints between the model parameters or operating conditions in general form, are also planned in the future works.

Acknowledgments

The authors gratefully acknowledge CNRS and Lorraine Region for the financial support granted to this study.

Nomenclature

Abbreviations

A Alcohol

AMA Alkali-Metal Alkoxide

AMH Alkali-Metal Hydroxide

AMSFA Alkali-Metal Salt of Fatty Acid

BSTR Batch Stirred-Tank Reactor

CSTR Continuous Stirred-Tank Reactor

CTR Continuous Tubular Reactor

DAEMD Data Acquisition Experiments for Model Discrimination

DAEPPI Data Acquisition Experiments for Parameter Precision Improvement

DG Di-Glyceride

DL Di-Linolein

DO Di-Olein

DP Di-Palmitin

E Ethanol

EL Ethyl Linoleate

EO Ethyl Oleate

EP Ethyl Palmitate

FFA Free Fatty Acid

FAAE Fatty Acid Alkyl Ester

FAEE Fatty Acid Ethyl Ester

G Glycerol

M_F Fisher Information Matrix

MG Mono-Glyceride

ML Mono-Linolein

MO Mono-Olein

MP Mono-Palmitin

MUT Model Under Test

NaOH Sodium hydroxide

PE Preliminary Experiment

RN Reaction Network

SBSTR Semi-Batch Stirred-Tank Reactor

SMUT Set of Models Under Test

STR Stirred-Tank Reactor

TG Tri-Glyceride

TL Tri-Linolein

TO Tri-Olein

TP Tri-Palmitin

TR Tubular Reactor

W Water

VEPO Validation Experiment for Performance Optimization

Greek symbols

γ kinetic order

Δh molar enthalpy

$\Delta \hat{\theta}$ vector of parameter confidence intervals

$\Sigma_{\hat{\theta}}$ parameter variance-covariance matrix

$\Sigma_{\hat{y}}$ model prediction variance-covariance matrix

Σ_y variance-covariance matrix of the residuals

$\overline{\Sigma_y}$ average variance-covariance matrix of the residuals

θ vector of model parameters

$\hat{\theta}$ vector of parameter estimates

Θ space of model parameters

ξ matrix operating conditions

Ξ space of operating conditions

τ experiment duration

ν stoichiometric coefficient

χ molar fraction

Latin symbols

A frequency factor

c concentration

C criterion

C_p heat capacity

E_a activation energy

f model function

F flow rate

G Jacobian matrix

h function of parameter variance-covariance matrix

J mass-transfer rate

\mathbf{J} vector function of performances

\mathbf{J}_n vector function of normalized performances

k chemical reaction rate constant

K overall mass-transfer coefficient

m distribution coefficient

M molar mass

M_T transformation matrix

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

Q sensitivity matrix

r chemical reaction rate

S heat-transfer surface

t sampling time / time / t-value

T temperature

U global heat-transfer coefficient

v molar volume

V volume

\mathbf{x} vectors of time-dependent state variables

$\dot{\mathbf{x}}$ vectors of first derivatives of time-dependent state variables

X conversion

\mathbf{y} vector of measured response variables

$\hat{\mathbf{y}}$ vector of measurable variables predicted by the model

Y selectivity

Subscripts

e refers to the e^{th} experiment

f refers to the f^{th} sample

i refers to the i^{th} species

j refers to the j^{th} reaction

k refers to the k^{th} phase

l refers to the l^{th} phase

p refers to the p^{th} performance

q refers to the q^{th} reactor

r refer to the r^{th} measured response variable

s refer to the s^{th} measured response variable

u refers to the u^{th} parameter

0 refers to initial

Superscripts

exch exchange

in inlet of the reactor

int initial

max maximal

min minimal

opt optimal

out outlet of the reactor

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. [https://doi.org/10.1016/0098-1354\(90\)87003-8](https://doi.org/10.1016/0098-1354(90)87003-8).

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. [https://doi.org/10.1016/S0009-2509\(98\)00270-X](https://doi.org/10.1016/S0009-2509(98)00270-X).

Asprey, S.P., Macchietto, S., 2000. Statistical tools for optimal dynamic model building. *Computers & Chemical Engineering*, 24(2), 1261-1267. [https://doi.org/10.1016/S0098-1354\(00\)00328-8](https://doi.org/10.1016/S0098-1354(00)00328-8).

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. [https://doi.org/10.1016/S0169-7439\(99\)00057-X](https://doi.org/10.1016/S0169-7439(99)00057-X).

Bonvin, D., Rippin, D.W.T., 1990. Target factor analysis for the identification of stoichiometric models. *Chemical Engineering Science*, 45(12), 3417-3426. [https://doi.org/10.1016/0009-2509\(90\)87147-K](https://doi.org/10.1016/0009-2509(90)87147-K).

Brendel, M., Bonvin, D., Marquardt, W., 2006. Incremental identification of kinetic models for homogeneous reaction systems. *Chemical engineering science*, 61(16), 5404-5420. <https://doi.org/10.1016/j.ces.2006.04.028>.

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. <https://doi.org/10.1016/j.ces.2007.10.010>.

Buzzi-ferraris, G., Forzati, P., 1983. A new sequential experimental design procedure for discriminating among rival models. *Chemical Engineering Science*, 38(2), 225–232. [https://doi.org/10.1016/0009-2509\(83\)85004-0](https://doi.org/10.1016/0009-2509(83)85004-0).

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. <https://doi.org/10.1016/j.chemolab.2008.08.002>.

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.

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. <https://doi.org/10.1016/j.biortech.2014.01.028>.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science*, 63(19), 4846-4872. <https://doi.org/10.1016/j.ces.2007.11.034>.

Galvanin, F., Ballan, C.C., Barolo, M., Bezzo, F., 2013. A general model-based design of experiments approach to achieve practical identifiability of pharmacokinetic and pharmacodynamic models. *Journal of pharmacokinetics and pharmacodynamics*, 40(4), 451-467. <https://doi.org/10.1007/s10928-013-9321-5>.

Hii, C.J., Wright, A.R., Willis, M.J., 2014. Utilizing a genetic algorithm to elucidate chemical reaction networks: An experimental case study. *International Journal of Chemical Engineering and Applications*, 5(6), 516. DOI: 10.7763/IJCEA.2014.V5.439.

Hunter, W.G., Reiner, A.M., 1965. Designs for discriminating between two rival models. *Technometrics* 7(3), 307–323. DOI: 10.2307/1266591.

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. https://doi.org/10.1007/978-3-540-73920-3_10.

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. <https://doi.org/10.1016/j.fuproc.2014.01.017>.

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. <https://doi.org/10.1016/j.apenergy.2014.02.046>.

Mathieu, F., Commenge, J.M., Falk, L., & Lomel, S., 2013. Technologies comparison for iterative data acquisition strategies. *Chemical Engineering Science*, 104, 829-838. <https://doi.org/10.1016/j.ces.2013.09.053>.

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. <https://doi.org/10.1137/090757009>.

Michalik, C., Brendel, M., Marquardt, W., 2009. Incremental identification of fluid multi-phase reaction systems. *AIChE journal*, 55(4), 1009-1022. <https://doi.org/10.1002/aic.11738>.

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. <https://doi.org/10.1016/j.cep.2007.01.016>.

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

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. <https://doi.org/10.1016/j.fuproc.2014.09.008>.

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. <https://doi.org/10.1016/j.ces.2012.10.014>.

Violet, L., Loubière, K., Rabion, A., Samuel, R., Hattou, S., Cabassud, M., Prat, L., 2016. Stoichiokinetic model discrimination and parameter identification in continuous microreactors. *Chemical Engineering Research and Design*, 114, 39-51. <https://doi.org/10.1016/j.cherd.2016.07.025>.

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. DOI: 10.1021/ie504343d.

Appendix 1: Description of the software

Module A: Project initialization

The software requires the user to enter the basic information about the studied liquid-phase synthesis.

In the submodule A1, the user has 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.

In the submodule A2, the user has to: (i) indicate the number of considered reactors; (ii) select for each reactor its configuration (from STR and TR), process type (from batch, semi-batch and continuous), thermal model (from adiabatic, isothermal and isoperibolic); (iii) 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.

In the submodule A3, the user has to: (i) indicate the number of all possible species and the number of all possible phases for the reaction system; (ii) 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; (iii) indicate for each species whether it presents in each possible phase. Then, in the background, the following information is automatically generated: (i) operating conditions in general form (illustrated in Table 16), that characterize all possible reactors; (ii) time-dependent state variables, from which the measurable variables are defined.

Table 16: Operating conditions in general form. (× = needs to be defined)

Operating condition	Variable	Unit	BSTR	SBSTR	CSTR	CTR
Experiment duration / Sampling time	t	[s]	×	×	×	
Initial temperature	T^{int}	[K]	×	×	×	
Initial load of reactants, solvents and catalysts	n_i^{int}	[mol]	×	×	×	
Feed temperature	T^{in}	[K]		×	×	×

Feed flow rate	F^{in}	[m ³ /s]	×	×	×
Feed duration	t^{in}	[s]	×		
Concentration of reactants, solvents and catalysts in the feed	c_i^{in}	[mol/m ³]	×	×	×

In the submodule A4, the user has to: (i) indicate the number of performances; (ii) define the performance from a suggestion list including the conventional performances, such as: 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 time (and maximal reaction temperature, for the reactor whose thermal mode is not isothermal); (iii) define for each performance its objective and veto values.

Module B: Model structure development

In this module, the software requires the user to enter a reaction supernetwork (B1) and to propose the reaction networks under test (B2). Then, the involved parameters are automatically generated (see Section 2.3.1), the user has to define the corresponding kinetic orders (B2) and to select the parameter to be actually estimated by indicating for each parameter whether it is independent, if not, to select its dependence from a suggestion list (B3). Once the reaction networks are available, the corresponding model structures for each considered reactor are automatically constructed in the background using the default reaction rate expressions (see Section 2.2.2).

Module C: Experimental pre-design

In the submodule C1, the user has to specify the bounds on the operating conditions in general form (see Table 16) and the linear or nonlinear equality or inequality constraints among them. Note that the same operating conditions can be represented through three different manners: the practical operating conditions (see Table 7), the practical operating variables (see Table 9) and the operating

conditions in general form (see Tables 8, 11 12 and 13). An example illustrating the relationship among them is given in Table 17.

Table 17: Relationship among practical operating conditions, practical operating variables and operating conditions in general form.

Operating condition	Practical operating condition	Practical operating variable	Operating condition in general form
Temperature	30 °C	30 °C	303.15 K
Initial load of oil	200 g	200 g	0.2279 mol
Initial load of NaOH	0.2 g	0.1 g per 100 g of oil	0.0050 mol

In the submodule C2, the user has to select the measurable variables, which can be divided into two types: conventional measurable variables (i.e. measurable time-dependent state variables) and unconventional measurable variables (i.e. measurable products of time-dependent state variables). Considering that model structure analysis is not integrated in the software, the user has to ensure that given the defined variables to be actually measured, all the candidate model structures proposed at B2 should be identifiable and discriminable.

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 PEs, the user has to not only input the experimental data into the software but also define the measurement errors (on the basis of precisions of measurement instruments and his expertise). The DAEPPi is designed by the software, the user has to input the measured responses into the software.

Module E: Model parameter development

The software first requires the user to specify 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 bound, is identified.

Module F: Model evaluation

Model evaluation is performed in this module. According to the model evaluation results, the various modules will be repeated until a set of adequate and accurate models allowing fitting all experiments is identified. The software is developed based on the strategy, each software sub(module) has its strategy refaction, i.e. one or several steps. Specific description about this module 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.

Module G: Results

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

Appendix 2: Chromatograms

Figure 10 and Figure 11 show respectively the chromatographic analyses of FAEEs and glycerides in the 3rd sample of the 4th series of PEs. These results proved respectively that

- EO, EL and EP are the dominating esters in the sample. Therefore, TO, TL and TP are the dominating TGs in the sunflower oil;
- Glyceride amounts could not be distinctively determined due to the superposition of their peaks.

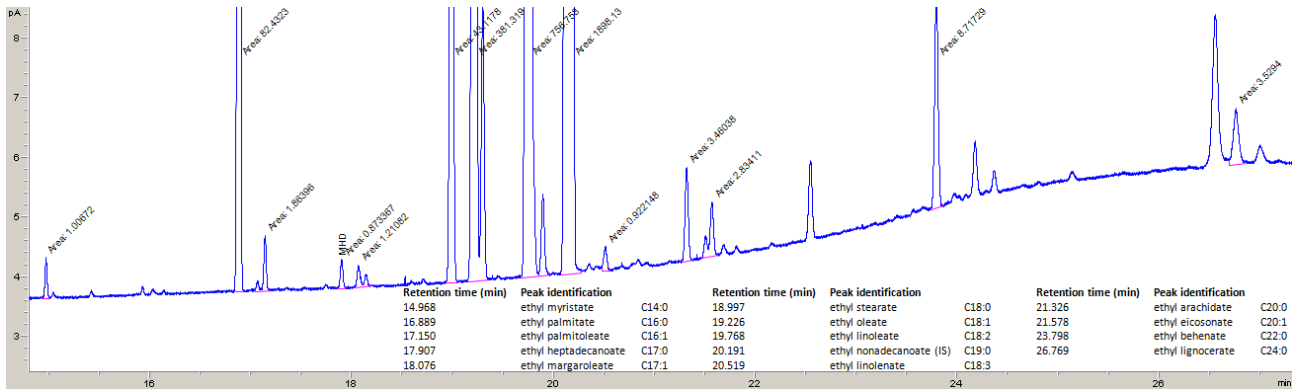


Figure 10: Chromatographic analysis of fatty acid ethyl esters in the 3rd sample of the 4th series of PEs.

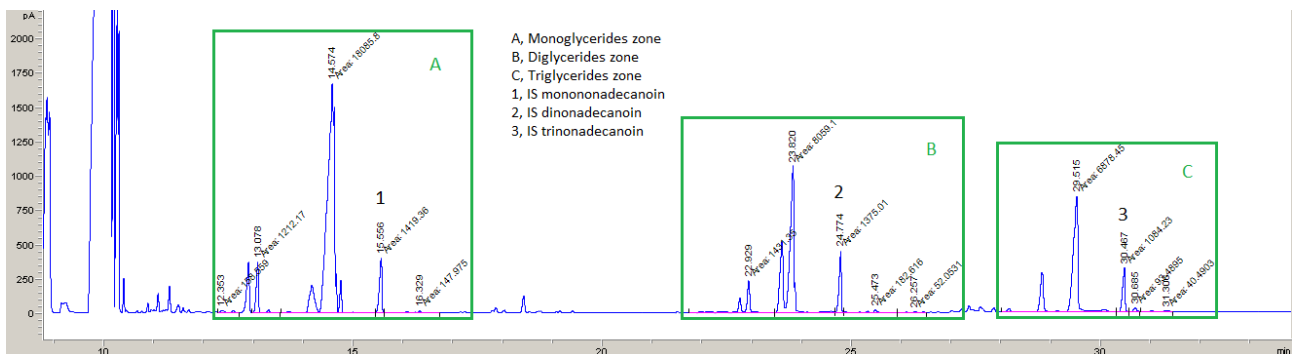


Figure 11: Chromatographic analysis of glycerides in the 3rd sample of the 4th series of PEs.

Appendix 3: Measured responses

The measured responses of the 1st DAEPPi, the 2nd DAEPPi and the VEPO are shown in Table 18.

Table 18: Measured responses of the DAEPPis and the VEPO.

Experiment	N° sample	C _{EO}	C _{EL}	C _{EP}	C _{MGs}	C _{DGs}	C _{TGs}
		[mol/m ³]	[mol/m ³]	[mol/m ³]	[mol/m ³]	[mol/m ³]	[mol/m ³]

1 st 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

2 nd 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
------	---	--------	---------	--------	-------	------	------
