European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

A numerical experiment design study on a biodiesel production process

Gaia Franceschini, Sandro Macchietto*
Department of Chemical Engineering, Imperial College
South Kensington Campus, SW7 2AZ London (UK)

Abstract

Advanced model-based experiment design techniques are a reliable tool for rapid development and refining of process models. Through two practical case studies, we demonstrate the validity of this approach as applied to planning of optimal experiments for complex kinetics elucidation. The need for an appropriate use of these tools, in particular a judicious problem formulation, is highlighted. A trade-off is found between the predicted precision of the estimates and important experimental aspects (for example, times and costs of the analytical work). The intelligent application of these techniques allows finding experiments which are suitable for the collection of data for complex kinetic reaction networks.

Keywords: Biodiesel, Model-based Experiment design, Parameter estimation, Kinetics Elucidation

1. Introduction

Building high quality, validated, steady-state or dynamic mechanistic models of process systems is a key activity in process engineering for many applications such as model-based product and process design, control and optimisation. Such models invariably contain adjustable parameters that have to be estimated. Model-based experiment design aims at assisting a modeller/experimenter in devising experiments that will yield the most informative data, in a statistical sense, for use in parameter estimation and model validation. In mathematical terms, given an initial model and assumed parameters, the aim is to minimise the expected inference region of the parameters, i.e. to make the elements of the parameters variance-covariance matrix small. An experiment design calculation thus involves minimising some measure of this matrix by choosing a set of experiment decision variables (length, initial conditions, sampling times, etc.) subject to equality or inequality constraints (Asprey and Macchietto, 2000).
The use of such methodology still requires a modeller/experimenter to choose a-priori: 1) for which parameters to design an experiment and in which order (one experiment for all parameters, one for each parameter, etc.) and 2) which of the available measurements to use in each experiment (all possible variables, just one or two, etc). The aim of this paper is to show how the experimental and analytical work will depend on these

* Author to whom correspondence should be addressed: s.macchietto@imperial.ac.uk

choices, and to stress the need for an intelligent and reasoned use of these tools in order to obtain the best results with the minimum effort. This is done by means of a practical application of industrial interest, the elucidation of kinetics for a biodiesel process.

2. The process and the problem

Biodiesel would be an ideal substitute for conventional diesel fuel if only it was more competitive economically. Efforts have been made to reduce its cost by optimising its production processes (Franceschini *et al.*, 2004). We studied one of these processes, the transesterification of a vegetable oil with methanol in presence of an alkali catalyst under mild pressure conditions in a batch reactor. Two schemes have been proposed for the kinetics involved. According to Noureddini and Zhu (1997), the reaction consists of three reversible steps (Figure 1.a). This kinetics was used in conjunction with the full reactor model of Franceschini *et al*. (2004) which includes phase equilibrium, temperature varying with time, methanol evaporation, energy input and losses (Model 1) and describes the behaviour of this system reasonably well. Komers *et al.* (2002) proposed an alternative kinetic scheme, which takes into account the catalyst concentration and the main saponification side reactions (Figure 1.b). This was used in conjunction with a simple isothermal mono-phase batch reactor model, where only the reactions occur (Model 2). With an Arrhenius form of the reaction rates, Model 1 has twelve parameters for the six reactions; Model 2 has ten reactions and twenty parameters.

(a)

$$TG + M \underset{k_2}{\overset{k_1}{\leftrightarrow}} DG + E$$
$$DG + M \underset{k_4}{\overset{k_3}{\leftrightarrow}} MG + E$$
$$MG + M \underset{k_6}{\overset{k_5}{\leftrightarrow}} G + E$$

(b)

$$TG + M + OH^- \underset{k_{1r}}{\overset{k_1}{\leftrightarrow}} DG + E + OH^-$$
$$DG + M + OH^- \underset{k_{2r}}{\overset{k_2}{\leftrightarrow}} MG + E + OH^-$$
$$MG + M + OH^- \underset{k_{3r}}{\overset{k_3}{\leftrightarrow}} G + E + OH^-$$
$$E + OH^- \xrightarrow{k_4} M + A^-$$
$$TG + OH^- \xrightarrow{k_5} DG + A^-$$
$$DG + OH^- \xrightarrow{k_6} MG + A^-$$
$$MG + OH^- \xrightarrow{k_7} G + A^-$$

Figure 1(a) & 1(b). Kinetic schemes

The symbols in Figure 1 have the following meaning: TG, DG and MG are the tri-, di- and mono-glycerides, respectively; M is the methanol, G the glycerol and E the mixture of esters which constitutes biodiesel. OH^- represents the catalyst's active part and A^- are the anions of the produced soaps.

Starting from data from previous experimental work, a reasonable good fit was obtained using Model 1. However, simulation results showed that more informative data are required (in particular during the early stages of reaction) for a more precise estimation of the twelve kinetic parameters. For Model 2, the question is whether the estimation of the four additional kinetic constants is at all possible with our available experimental and analytical setup. Saponification reactions are secondary and their rates are slow. Indeed, a first attempt at estimating these parameters showed it is impossible to estimate them with sufficient precision using the data of just one experiment reported in Komers

et al. (2001). Two sets of optimally designed experiments were therefore devised to address these questions.

3. Methods

Modelling, simulation, parameter estimation and experiment design were carried out using gPROMS, interfaced with two databases (Infodata and DIPPR) and a program (Multiflash) for the calculation of the thermo-physical properties of the reaction mixture.

The dynamic Models 1 and 2 with initial parameters obtained as described in Section 2 were used to design for each model a new set of experiments for precise parameter estimation. First, a dynamic sensitivity analysis was carried out to identify the parameters respect to which the model is more sensitive and the more suitable intervals to collect the data for their estimation. For example, Figure 2.a and Figure 2.b, respectively, show the results for Model 1, where the most important parameters are found to be three: the pre-exponential factors of the third reaction step (A_5 and A_6) and of the first reverse reaction (A_2).

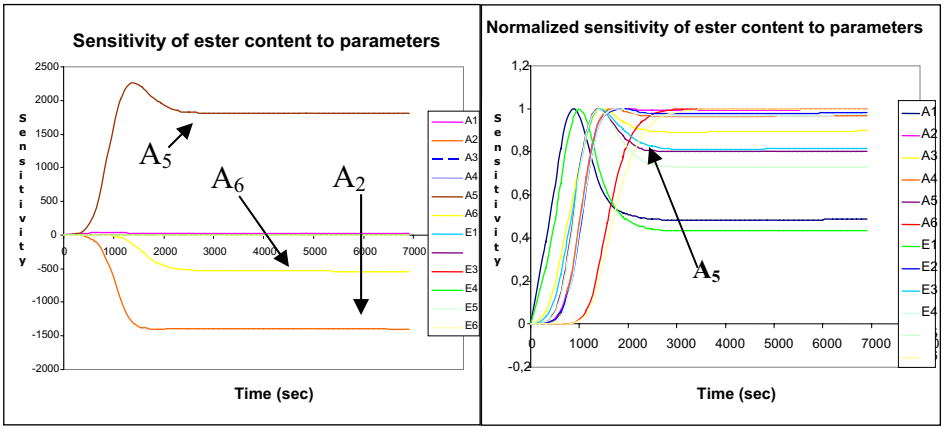

Figure 2. Sensitivity analysis results for Model 1. A perturbation of 5% was applied on each parameter and the ester content is the response used.

These results were used to plan the experiment design problem. First, the parameters to be estimated (individually, in pairs, etc.) and measurements to use (content of ester, glycerol and monoglycerides) were selected. Then, the number of samples and initial guesses for the sampling times were chosen (all in the range around the maximum of the corresponding normalised sensitivity curve). This is important: for example, using sixteen randomly chosen sampling points gives worse optimal estimates than with eight sampling points chosen according to the sensitivity analysis (t-values of 31.9 and 42.09, respectively for parameter A_1 of Model 1 – see section 4). This way, an initial set of optimally designed experiments was found. These first results were then improved by introducing constraints to ensure that the optimal experiments satisfied many practical limitations of the apparatus. For instance, our reactor can be heated, but not cooled, so the temperature profile slope can only be positive or null. In this way, we obtained a set of experiments actually realizable in our laboratory. The next step was to analyse this

set of predicted optimal experiments, taking into account for each the costs and times required to obtain the experimental data. The chemical analyses are time consuming and so we tried to reduce as much as possible the number of samples and variables that have to be measured. However, the predicted accuracy of the estimated parameters will decrease and so a trade-off has to be found between the expected precision of the estimation and the overall amount of experimental and analytical work.

4. Experiment design results

4.1 Model 2: effect of operating conditions and measured variables

The decision variables to be optimised were: the initial concentration of oil, methanol and catalyst, the process time and the sampling times. Of the many combinations attempted of parameter to be estimated, measured variables and number of samples, the set of four optimal experiments A, B, C, D in Table 2 permit an excellent identification of the complete saponification kinetics with a small amount of analytical work (just one additional experiment, not shown, being sufficient to estimate all other reaction constants).

Table 2. Optimal experimental conditions (saponification parameters)

	EXP.A	**EXP.B**	**EXP.C**	**EXP.D**	**EXP.E**
Parameters estimated	k_4	k_5	k_6	k_7	k_5 & k_6
Process time (min)	30	28	51	74	105
Amount of oil/ MeOH (mol/l)	0.87/4.06	0.99/0.99	0.83/4.97	0.83/4.97	0.83/4.97
Amount of catalyst (mol/l)	0.0187	0.0145	0.0103	0.011	0.00953
Number of samples	7	7	7	7	7
Variables measured	G	E	G	E, G	E, G
Number of analyses (EP+GP)[a]	7+7	7+7	7+7	14+14	14+14
Precision (t-value)	7.442	15.05	13.66	6.0225	2.175/3.412
Reference t-value	1.943	1.943	1.943	1.771	1.783

[a] The sample spontaneously separates in an ester phase (EP) and in a glycerol phase (GP)

After the sensitivity analysis, the first experiment design study attempted showed that it is impossible to identify all four saponification reaction parameters with just one experiment, albeit optimal. The reason is that very different operating conditions are required to identify each parameter. For example, for parameter k_4, methanol has to be slightly in excess (the ratio *r* between the amounts of the two reactants is 3.65÷4.7). On the other hand, k_7 requires the maximum allowed alcohol excess in three quarters of the experiments and an almost stochiometric ratio in the others, depending on what is measured. Parameters k_5 and k_6 are the only two that could be estimated in pair. They need a very small reactants ratio (~1) in half of the experiments and the maximum allowed alcohol excess for the others. This very large difference in operating conditions can be explained if we consider the two experiment design studies applied to these two parameters singularly (Exp.B and Exp.C in Table 2). Indeed, k_5 needs a very small amount of methanol except for one case (see the paragraph end), whereas k_6 requires an alcohol excess of 100%. If we look at the t-values for the couple k_5 & k_6 (Exp.E in Table 2), we can see that when the alcohol is in excess the precision of the estimation is improved for the latter, whereas with a ratio of 1:1 it is easier to estimate k_5. In

addition, we found that very different results can be obtained according to the variables which can be measured in the experiment. For example, for k_5 a very small amount of methanol is required. Indeed, the less methanol is in the reactor the less the transesterification reactions extent, which gives higher yields for the saponification ones. Yet, when we measure the glycerol content (case E3, not documented here), a very large excess of alcohol is used, probably because glycerol is not directly involved in the reaction studied. So, obtaining a satisfactory estimation of this parameter needs a high amount of this compound to be produced in order to measure it.

4.1.1 Trade off

Different options were available for the two parameters k_5 and k_6. Two experiments, one for each parameter, could be designed or a single experiment for the couple. If we compare the t-values for cases B, C and E (2.2/3.4 vs. 15/14), we can see that experiments B and C give a very higher precision. In both cases (B+C and E), the number of analyses is the same because, if we use only one experiment, we have to measure two variables in order to obtain good results, whereas, if we use two different experiments, it is just sufficient to measure only one. In the end, we chose to carry out two distinct experiments for the estimation of one parameter at a time, because the total duration of both experiments is shorter than the single experiment for the pair, while still giving a higher parameter precision.

4.2 Model 1

Here, the decision variables were: the experiment duration, the initial state of the experimental equipment (initial oil and methanol charges and temperature), the temperature profile during the experiment and the sampling times at which measurements are taken. Of the many tried, a set of six experiments that lead to identification of all parameters was established. The experimental conditions are shown in Table 3 together with the predicted precision of the estimate.

Table 3. Optimal experimental conditions (transesterification parameters)

	EXP.F	**EXP.G**	**EXP.H**	**EXP.I**	**EXP.L**	**EXP.M**
Parameters estimated	A_1	A_2 & A_6	A_3 & A_4	A_5	E_1,E_3 & E_5	E_2,E_4 & E_6
Process time (min)	19	94	71	64	27	76
T max (K)	310	368	340	332	316	341
Amount of oil/ MeOH (mol/l)	3/23.76	3/14.57	3.3/16.96	3.2/14.9	3.95/12	3.95/12
Number of samples	8	12	8	14	8	8
Variables measured	E	E	E	E	MG, E	G, E
Number of analyses (EP+GP)	8+8	12+12	8+8	14+14	8+8	8+8
Precision (t-value)	36.59	13.4/8.5	9.64/5.2	24.89	96/76/49	12.6/21/52
Reference t-value	1.895	1.812	1.943	1.771	1.771	1.771

4.2.1 Trade off

Different options were available for the three most important parameters of the model (A_2, A_5, A_6). We could design an experiment for each parameter, or one experiment for all, or an experiment for a couple plus a single experiment for the remaining parameter. The couple A_2 & A_5 could not be estimated in a single experiment, because even with twelve samples and three variables measured the results are not statistically satisfactory. Parameter A_5 is the most difficult to estimate and would require a very large number of

samples. With only eight samples –sufficient for the other parameters- the estimate is not precise enough or the temperature has to be unrealistically high. In the end, we selected two optimal experiments for the estimation of these parameters, one for the couple A_2 & A_6 and one for parameter A_5 alone. Indeed, as we can see in Table 4, there is an evident trade-off between the number of experiments, the number of analyses and the predicted precision of the estimation.

Table 4. Comparison between the optimal experiments for the most important parameters

	A_2+A_5+A_6	**A_2, A_5, A_6**	**A_2, A_5+A_6**	**A_2+A_6, A_5**
Number of experiments	1	3	2	2
Number of analyses (EP)	45	7+14+7=28	7+13=20	12+14=26
Number of analyses (GP)	45	7+14+7=28	7+13=20	12+14=26
Number of analyses (Total)	90	56	40	52
Precision: t-value (A_2)	4.243	13.29	13.29	13.4
Precision: t-value (A_5)	5.104	24.89	15.28	24.89
Precision: t-value (A_6)	5.734	13.93	7.938	8.497

5. Conclusions

Model-based experiment allows in principle to optimise experimental effort. However, we demonstrated here via a practical example that a judicious problem formulation (choice of constraints, measured and estimated variables) is essential to obtain good results. The use of a preliminary sensitivity analysis is compulsory, for instance, for the initial choice of sampling points. Our study highlights also an important trade-off between precision of the estimates, number of samples, experiments and analytical effort required. Taking into account the most critical variables (in our case the number of analyses, but for a biological system it could be the length of the experiment or the total amount of substances available), this trade-off analysis, which can be performed prior to experimentation, allows a modeller/experimenter to establish the next set of optimal experiments.

We trust the messages from this study are useful, generic to many situations and worthy of further refinement. They indicate possible future developments for the optimal experiment design method, aimed at systematising the choices highlighted. The actual data obtained from our optimal experiments will be described in a follow-up paper.

References

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

Franceschini G., A. Bertucco and S. Macchietto, 2004, Simulation and optimisation of a biodiesel production process. Convegno GRICU 2004: Nuove Frontiere di Applicazione delle Metodologie dell'Ingegneria Chimica, Porto d'Ischia (Napoli), 12-15 settembre 2004.

Noureddini H. and D. Zhu, 1997, Kinetics of Transesterification of Soybean Oil, Journal of the American Oil Chemist's society (JAOCS), 74, 1457-1463.

Komers K., R. Stloukal, J. Machek and F. Skopal, 2001, Biodiesel from rapeseed oil, methanol and KOH 3. Analysis of composition of actual reaction mixture, Eur. J. Lipid Sci. Technol., 103, 363-371.

Komers K., F. Skopal, R. Stloukal and J. Machek, 2002, Kinetics and mechanism of the KOH–catalyzed methanolysis of rapeseed oil for biodiesel production, Eur J Lipid Sci Tech, 104,728-737.