Sequential Experimental Design Strategy for Rapid Kinetic Modeling of Chemical Synthesis

Sébastien Issanchou, Patrick Cognet, and Michel Cabassud
Laboratoire de Génie chimique—UMR 5503 CNRS/INPT (ENSIACET)/UPS, BP 1301, 31106 Toulouse Cedex 1, France

DOI 10.1002/aic.10439
Published online April 29, 2005 in Wiley InterScience (www.interscience.wiley.com).

In this work, an optimum experimental design strategy has been developed to increase the reliability of chemical reaction engineering models when few experiments are to be performed. The strategy has been used for heterogeneous liquid–liquid reactions carried out in a batch reactor. The specific case of alkaline hydrolysis of n-amylacetate has been selected to illustrate the proposed methodology, allowing experiments to be carried out to identify the associated kinetic parameters. The kinetic parameters that arise from the model system of ordinary differential equations have been estimated by weighted least squares and a D-optimum sequential design has been implemented to reduce the volume of the joint confidence region of the parameters. Estimates and their uncertainty obtained at the end of the strategy, as well as the generalization ability of the prediction model, are presented and discussed. © 2005 American Institute of Chemical Engineers *AIChE J,* 51: 1773–1781, 2005
Keywords: experimental design, kinetics, mass transfer, multiphase reactors, parameter identification

Introduction

In the field of fine and specialty chemistry, rapid process development is essential to accommodate fast and constantly changing fluctuations of the market. Process development activity is primarily concerned with the laboratory determination of optimal operating conditions for candidate reactions and its subsequent implementation on a larger scale. Nonetheless, contrary to the traditional chemical industry where detailed studies are performed to elucidate the mechanisms and the kinetics of reactions, the fine chemicals industry cannot generally support such in-depth studies because of cost and time considerations. As a result, the design of fine chemicals production units has been based for some time on an empirical approach, which consists in laboratory-scale measurements of the reaction time and to directly use these results in large-scale batch units (Lundström et al., 1999). To determine the optimal production conditions (temperature, pressure, initial concentrations of reactants and catalyst, etc.), a complementary approach consists in using linear and quadratic models whose parameters can be quickly and precisely estimated by the construction of factorial and centered composite designs of experiments (Benoist et al., 1994). Classically, the model is then used to predict an optimum over a bounded region of the experimental space. Even if this approach leads to a reduction of the experimental effort needed for the process modeling and scale-up, however, it will also lead to serious drawbacks because (1) the model may be valid over only a confined region of the experimental space, (2) it may not properly take into account the dynamics of a batch production and mass-transfer effects, and (3) the laboratory optimal conditions may not be directly valid when scaling-up.

An innovative approach consists in using chemical reaction engineering theory–based models and in building a nonlinear optimum experimental design to improve the determination of its parameters. Box and Lucas (1959) pioneered the development of optimal experimental strategies for parameter estimation of nonlinear mechanistic models. The basic idea of optimum experimental design is to act on the joint confidence region of parameter estimates, which is defined as an ellipsoid

Correspondence concerning this article should be addressed to P. Cognet at Patrick.Cognet@ensiacet.fr.

in the parameter space. The minimization of the volume of the ellipsoid leads to the most common optimality criterion—the D-optimality criterion—although many others have been suggested and investigated. A review of criteria for optimum experimental design can be found in Walter and Pronzato (1990, 1994), whereas Kiefer (1974) reported a family of functions that generalize the classical optimum design criteria.

There is a vast literature stressing applications of optimal design theory to kinetic modeling of homogeneous and catalytic chemical or biochemical reactions (see, for example, Argawal and Brisk, 1985; Atkinson and Hunter, 1968; Baltes et al., 1994; Sedrati et al., 1999) in a batch reactor. However, despite the great industrial importance of liquid–liquid processes, the rapid modeling of liquid–liquid batch systems has not yet been intensively investigated. Moreover, except for some contributions (Juusola et al., 1972; Mezaki, 1969; Reilly et al., 1977; Rippin et al., 1980), only a few experimental applications of this method have been reported with respect to chemical kinetic studies.

In this work, a mathematical model for a liquid–liquid stirred tank batch reactor is developed. A sequential design strategy is implemented to reduce the experimental effort needed during the identification step. This strategy is applied and tested on a case study, the heterogeneous alkaline hydrolysis of *n*-amylacetate, for which experiments can be easily carried out.

Mathematical Model of the Liquid–Liquid Batch Reactor

The general model of the liquid–liquid batch reactor, developed in this work, is based on the partial mass balance equations derived in both the organic and the aqueous phase. Herein, we assume that *nc* compounds are involved in *nr* chemical reactions that take place only in the continuous aqueous phase.

The mathematical modeling of mass transfer between the organic and the aqueous phases is described by using Fick's law and the film theory of Lewis and Whitman (1924). We further assume that there is no mass-transfer resistance in the organic phase and that the concentration profiles in the aqueous film are linear. By introducing the thermodynamic equilibrium of species at the interface, the molar flow density of a diffusive compound i from the organic phase to the aqueous phase is then given by

$$N_i = k_i^c\left(\frac{C_i^{d,I}}{m_i} - C_i^c\right) \tag{1}$$

By denoting ν_{ji} and α_{ji} as the stoichiometric coefficient and the partial order of constituent i in reaction j, respectively, the generation rate of a compound i by chemical reaction is given by

$$r_i = \sum_{j=1}^{nr} \nu_{ji} k_j \prod_{k=1}^{nc} (C_k^c)^{\alpha_{jk}} \tag{2}$$

The evolution of the concentration of compound i in the continuous aqueous phase with respect to time is then given by

$$\frac{dC_i^c}{dt} = r_i + \frac{1}{V^c}\left(AN_i - C_i^c\frac{dV^c}{dt}\right) \tag{3a}$$

with $C_i^c(t=0) = C_i^{c,0}$, whereas in the organic dispersed phase, it is

$$\frac{dC_i^d}{dt} = \frac{1}{V^d}\left(-AN_i - C_i^d\frac{dV^d}{dt}\right) \tag{3b}$$

with $C_i^d(t=0) = C_i^{d,0}$

We assume that the volume variation in both phases arises only from mass transfer effects, that is,

$$\frac{dV^c}{dt} = \sum_{i=1}^{nc} \vartheta_i N_i \tag{4a}$$

with $V^c(t=0) = V^{c,0}$.

$$\frac{dV^d}{dt} = -\sum_{i=1}^{nc} \vartheta_i N_i \tag{4b}$$

with $V^d(t=0) = V^{d,0}$.

Finally, to account for the dependency of the model parameters on medium temperature, we assume that both kinetic constants k_j ($j = 1, nr$) and partition coefficients m_i ($i = 1, nc$) follow an Arrhenius law, that is,

$$k_j = k_j^0 \exp\left(-\frac{T_{k,j}}{T}\right) \tag{5a}$$

where $T_{k,j} = E_{k,j}/R$ and

$$m_i = m_i^0 \exp\left(-\frac{T_{m,i}}{T}\right) \tag{5b}$$

whereas we assume that mass-transfer coefficients k_i^c ($i = 1, nc$) do not depend on the temperature.

Correlation

When the mass-transfer coefficient is not included in the set of identified parameters, it is estimated by a correlation developed by Alwan et al. (1983) during the study of *n*-amylacetate extraction with soda. This correlation (Eq. 6) is derived from a similar correlation established by Nagata et al. (1971) for estimating heat-transfer coefficients and shows the influence of the Reynolds number, the Schmidt number, and the geometric dimensions of the reactor. Interfacial area and mass-transfer coefficients of diffusing species can first be estimated by using the correlation of Chen and Middleman (1967) for the mean Sauter diameter and the correlation of Wilke and Chang (1955) for the diffusion coefficients, as follows:

$$k_i^c = 0.90\,\frac{D_i}{d_r}\left(\frac{N_a d_a^2}{\eta}\right)^{2/3}\left(\frac{\eta}{D_i}\right)^{1/3}\left(\frac{d_a}{d_r}\right)^{-0.3}\left(\frac{w}{d_r}\right)^{0.45}(n_b)^{0.2}\left(\frac{L}{d_r}\right)^{-0.6} \tag{6}$$

Then, in its operative form, the kinetic model of the liquid–liquid batch reactor that has been implemented consists ofn a system of $2(nc + 2)$ ordinary differential equations (ODE): Eqs. 3a, 3b, 4a, and 4b.

Parameters of the model that must be identified are the following:

- k_j^0: frequency factor of chemical reaction rate constant of reaction j
- $T_{k,j}$: modified activation energy of chemical reaction rate constant of reaction j
- m_i: distribution coefficient of component i
- m_i^0: frequency factor of distribution coefficient of component i
- $T_{m,j}$: modified activation energy of distribution coefficient of component I
- k_i^c: mass transfer coefficient of component i in the continuous phase

Among the $(2nr + 2nc)$ parameters, mass-transfer coefficients can be calculated during the first stage of the strategy according to Eq. 6, as stated later in Results and Discussion.

Parameter Estimation and Optimum Experimental Design: Theoretical Background

Parameter estimation and parameter form variance–covariance matrix estimation

First, let $\mathbf{y^m}(\mathbf{p}, \mathbf{x})$ be the vector of the computed outputs of a given model and $\mathbf{y^e}$ the vector of observed data on the system we want to model. Assuming that the model has an exact structure and that there are no errors in the independent variables $\mathbf{x}$, the stochastic of the model is then given by

$$\mathbf{y^e} = \mathbf{y^m}(\mathbf{p_t}, \mathbf{x}) + \mathbf{e_t} \tag{7}$$

where $\mathbf{p_t}$ is the true value of the parameters and $\mathbf{e_t}$ is the vector of the true realizations of a stochastic variable ε, which is assumed to follow a normal law $N(\mathbf{0}, \Sigma)$.

The goal of the least-square estimator of $\mathbf{p_t}$ with weighting factor $\mathbf{\Sigma}^{-1}$ is to find out the particular value $\mathbf{p}^*$ of $\mathbf{p}$ that minimizes j_{wls} in the following expression

$$j_{wls} = \sum_{i=1}^{n} [\mathbf{y^e} - \mathbf{y^m}(\mathbf{p}, \mathbf{x})]^T \Sigma^{-1} [\mathbf{y^e} - \mathbf{y^m}(\mathbf{p}, \mathbf{x})] \tag{8}$$

Under some regularity conditions, this asymptotic weighted least-square estimate of $\mathbf{p_t}$ is an unbiased and efficient estimate that is normally distributed, that is,

$$\mathbf{p}^* \rightarrow N[\mathbf{p_t}, \mathbf{M_f}^{-1}(\mathbf{p_t}, \mathbf{x})] \tag{9}$$

where $\mathbf{M_f}$ is the Fisher information matrix that can be computed by

$$\mathbf{M_f}(\mathbf{p}, \mathbf{x}) = \sum_{i=1}^{n} \left[\frac{\partial \mathbf{y^m}(\mathbf{p}, \mathbf{x})}{\partial \mathbf{p}}\right]^T \Sigma^{-1} \left[\frac{\partial \mathbf{y^m}(\mathbf{p}, \mathbf{x})}{\partial \mathbf{p}}\right] \tag{10}$$

From the Cramer–Rao theorem, $\mathbf{M_f}^{-1}(\mathbf{p_t})$ is indeed the lower bound for the asymptotic variance–covariance matrix $\mathbf{V_p}$ of any unbiased estimate of $\mathbf{p_t}$ (Sorenson, 1980; Zacks, 1981) and, as a result, it is often used as an appropriate statistical estimation of $\mathbf{V_p}$, even if the data are finite in number.

In a dynamic model described by a set of ODE, the computation of the derivatives of the model outputs with respect to the parameters must be preceded by the computation of the sensibility equations that are also given as a set of ODE. Details on the formation of these sensibility equations can be found elsewhere (Hosten and Emig, 1975).

Confidence region and experimental design

The most commonly used confidence regions are np-dimensional ellipsoids whose centers are at the estimate $\mathbf{p}^*$. They define a bounded closed subset of the parameter space that contains the true value of the parameters $\mathbf{p_t}$ with a given probability, irrespective of the data sample obtained for the identification.

In the general case of a nonlinear model with respect to the parameters, it is possible to define a linearized confidence region by linearizing the model in the vicinity of the estimate $\mathbf{p}^*$. To this approximation corresponds the convergent approximation of $\mathbf{V_p}$ by $\mathbf{M_f}^{-1}$ and the so-called first-order confidence region can be computed by

$$(\mathbf{p} - \mathbf{p}^*)^T \mathbf{M_f} (\mathbf{p} - \mathbf{p}^*) = c \tag{11}$$

where c represents a constant depending on the parameters of the Fisher distribution.

This equation represents an ellipsoid centered at $\mathbf{p}^*$. Its characteristics are dependent on the scalar values of $\mathbf{M_f}$. In particular, the volume of the joint confidence region is proportional to the square root of the determinant of $\mathbf{M_f}^{-1}$ and the lengths of its axis are proportional to the square roots of the eigenvalues of $\mathbf{M_f}^{-1}$.

The optimum experimental design concept consists in choosing, from among all admissible operating conditions, those that will lead to an optimal modification of the confidence region of the estimates. The D-optimal criterion used in this work searches for the optimal operating conditions $\mathbf{x^D}$ that satisfy

$$\mathbf{x^D} = \arg\min_{\mathbf{x}\in\mathbf{X}} \det(\mathbf{M_f}^{-1}) \tag{12}$$

which will thus lead to the optimal reduction of the volume of the asymptotic confidence region of the estimates.

Sequential strategy

When the model is nonlinear with respect to the parameters, $\mathbf{M_f}$ is dependent on the $\mathbf{p}$ value. Therefore, the estimation of $\mathbf{V_p}$ by the inverse of the Fisher information matrix should be computed with the true value of the parameters $\mathbf{p_t}$, which is a priori unknown. To avoid the dependency of the optimum design criterion on $\mathbf{p_t}$, the most common way is to use a full

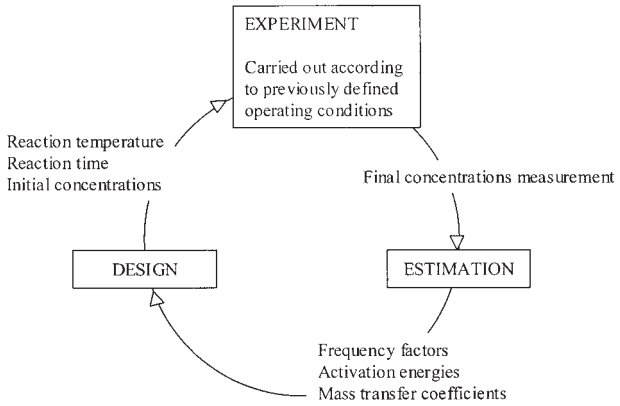

Figure 1. Sequential experimental design strategy.

sequential strategy (Figure 1) that consists in alternating parameter identification and design of experiments to improve the nominal value of $\mathbf{p}$, which is used instead of $\mathbf{p_t}$. First, a few nonoptimal experiments are performed and an estimate $\mathbf{p}^{*,1}$ is computed. This value is then used to compute the first optimal experiment. Once this experiment has been performed, a new estimate $\mathbf{p}^{*,2}$ is computed with all available data and is used to determine the next experiment.

The objective of this work is to show the influence of the choice and the number of experiments. Therefore, no specific convergence criterion has been introduced. A criterion could be introduced later based, for example, on the difference between two consecutive estimation values of the parameters. When this difference is inferior to a previously fixed tolerance, the procedure is stopped.

The next section presents the application of this strategy to an experimental application: the alkaline hydrolysis of *n*-amylacetate.

Application Example: Alkaline Hydrolysis of *n*-Amylacetate

Reaction mechanism and stoichiokinetic model

We consider the case of alkaline hydrolysis of *n*-amylacetate (A) with soda (B) in a batch reactor. Because of the partial miscibility of the ester in water, the reaction is carried out in a heterogeneous liquid–liquid medium. The ester has to enter the aqueous phase before it can react with soda. We assume that soda cannot enter the organic phase and, then, the reaction is supposed to occur only in the aqueous phase. This reaction leads to the production of *n*-pentanol (C) and sodium methanoate (D). Sodium methanoate is supposed to stay in the aqueous phase, whereas we assume that *n*-pentanol can enter the organic phase.

The stoichiometric reaction scheme is given by

$$A^d \rightarrow A^c \tag{13a}$$

$$A^c + B^c \rightarrow C^c + D^c \tag{13b}$$

$$C^c \rightarrow C^d \tag{13c}$$

Several mechanisms have been investigated and observed for the alkaline hydrolysis of esters with soda and a review can be found in Ingold (1953). However, it has been shown that, in the case of *n*-amylacetate (March, 1992), the main mechanism that governs the reaction kinetics is a bimolecular nucleophilic substitution followed by a fast proton transfer (presented in Figure 2).

The first step is the rate determining one and, consequently, the overall reaction rate can be described by a second-order van't Hoff's law, the order being one with respect to ester and alkali, that is,

$$r_A = r_B = -r_C = -r_D = -k_1^0 \exp\left(-\frac{T_{k,l}}{T}\right) C_A^c C_B^c \tag{14}$$

Finally, for the transformation carried out batchwise, the model can be expressed as

$$\frac{dC_A^c}{dt} = \frac{1}{V^c}\left[k_A^c A_e\left(\frac{C_A^d}{m_A} - C_A^c\right) - C_A^c \frac{dV^c}{dt}\right] - kC_A^c C_B^c \tag{15}$$

$$\frac{dC_B^c}{dt} = -kC_A^c C_B^c - \frac{C_B^c}{V^c}\frac{dV^c}{dt} \tag{16}$$

$$\frac{dC_C^c}{dt} = \frac{1}{V^c}\left[k_C^c A_e\left(\frac{C_C^d}{m_C} - C_C^c\right) - C_C^c \frac{dV^c}{dt}\right] + kC_A^c C_B^c \tag{17}$$

$$\frac{dV^c}{dt} = \upsilon_A k_A^c\left(\frac{C_A^d}{m_A} - C_A^c\right) + \upsilon_C k_C^c\left(\frac{C_C^d}{m_C} - C_C^c\right) \tag{18}$$

$$\frac{dV^d}{dt} = -\upsilon_A k_A^c\left(\frac{C_A^d}{m_A} - C_A^c\right) - \upsilon_C k_C^c\left(\frac{C_C^d}{m_C} - C_C^c\right) \tag{19}$$

Experimental equipment and chemicals

The experiments used to identify the parameters of the *n*-amylacetate alkaline hydrolysis model were carried out in a 750-mL glass reactor equipped with a glycol-filled heating jacket to control the bulk temperature (cf. Figure 3). A temperature sensor immersed in the solution ensured that the bulk temperature remained constant.

The stirrer was a six-blade Rushton turbine and four baffles were placed in the vessel.

Organic chemicals [that is, *n*-amylacetate (purity $\geq$ 98.5%), pentanol (purity $\geq$ 99%), and butanol (purity $\geq$ 99.5%)], used

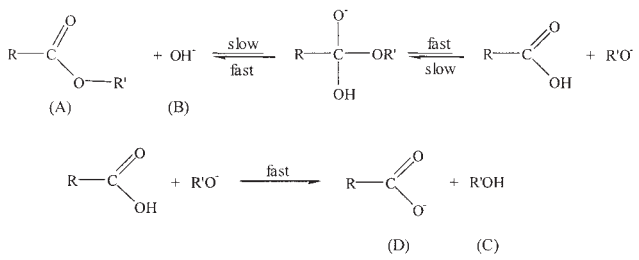

Figure 2. Mechanism of the alkaline hydrolysis of *n*-amylacetate.

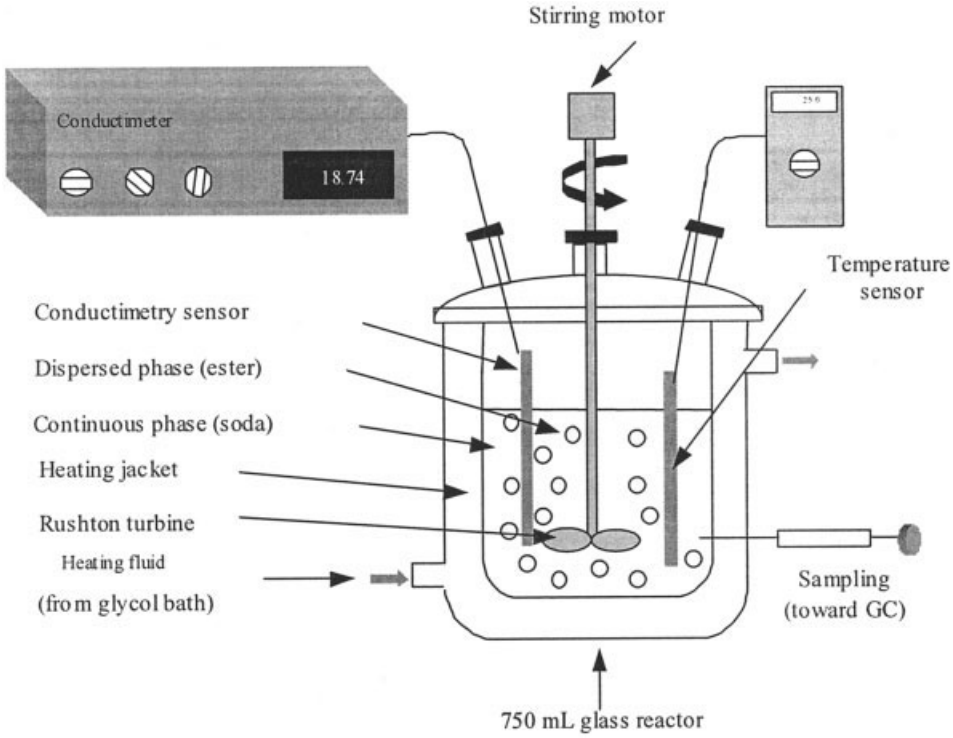

Figure 3. Experimental setup.

as reagents and/or calibration standards, were all analytical grade. Soda solutions were prepared from sodium hydroxide pellets (purity $\geq$ 99%). Concentrated HCl ($\sim$5 M) was used to stop the reaction.

Experimental procedure and analytical methods

Before starting an experiment, 600 mL of soda solution was introduced into the reactor. The stirrer was turned on and the bulk temperature was brought to the specified isothermal reaction temperature by filling the heating jacket with glycol. When the reaction temperature was reached, 20 mL of pure ester was rapidly added to the reactor, giving the initial time of the reaction ($t = 0$). Because the reaction took place in a diluted medium, the low exothermicity of the reaction did not cause the bulk temperature to vary more than a few tenths of a degree from its initial value.

At specified measurement times, about 5 mL of the aqueous phase was sampled and the reaction was stopped by adding concentrated HCl. The pH of the solution was checked and a 2-mL sample was taken. Finally, a known small amount of *n*-butanol was added and the sample was analyzed by gas chromatography (GC).

A conductimetry sensor directly immersed in the dispersion was used to follow the concentration of soda in the aqueous phase.

The concentrations of ester and alcohol in the aqueous phase were determined by GC. The internal standard method using *n*-butanol as a standard was performed and the calibration curve showed a very satisfactory linear relationship between both area ratios and concentration ratios over the entire useful range.

Results and Discussion

Numerical strategies

The ODE system formed by the $2(nc + 2)$ equations standing for the evolution of the concentrations and the volumes and the ODE system formed by the $[2(nc + 2)]np$ equations needed for the computation of the sensibilities were solved simultaneously by a Gear method (Hindmarsh, 1980).

The minimization of the weighted least-squares criterion was performed by using a Levenberg–Marquardt algorithm (Marquardt, 1963), available in the IMSL libraries. Each time an estimation was performed, the variance of the measurement noise for each constituent was computed and statistical tests were used to select the appropriate weights (that is, the form of the Σ matrix) in the next design and estimation phases.

Finally, to ensure the determination of the D-optimal operating conditions during the design phases, a nonlinear sequential quadratic programming (SQP) method (Schitkowski, 1986), available in the IMSL libraries, was used.

As previously mentioned, the strategy is separated in two steps. During the first step, interfacial area and mass-transfer coefficients of diffusing species are first estimated (Eq. 6).

During the second step of the strategy, mass-transfer coefficients are estimated simultaneously with the other parameters of the model (that is, k_1^0, $T_{k,1}$, m_A^0, $T_{m,A}$, m_C^0, $T_{m,C}$) and the reduction of their uncertainty is included in the optimum experimental design criterion as well.

Implementation of the strategy

In this work, V^c, V^d, and N_a are fixed to the following values: $V^c = 600 \times 10^{-6}$ m^3, $V^d = 20 \times 10^{-6}$ m^3, and $N_a = 240$ rpm.

Table 1. Admissible Experimental Domain

	$C_B^{c,0}$ (mol m^{-3})	T (K)	t_r (s)
Lower bound	40	288	120
Upper bound	120	308	5400

The experimental conditions, which we have to optimize in a D-optimal sense, are the initial concentration of B ($C_B^{c,0}$), the temperature of the reactor (T), and the measurement time (t_r) at which the concentrations of A, B, and C are determined in the aqueous phase. The admissible experimental domain that was retained is given in Table 1.

Three initial experiments, given in Table 2, were performed. From the experimental data obtained from these experiments, nominal values of the parameters were estimated, thus allowing us to start the sequential procedure.

Table 3 shows the D-optimal experiments that have been determined and carried out sequentially. For each experiment, values given in parentheses correspond to the computed D-optimal operating conditions, whereas the other values correspond to the applied operating conditions. It has been shown (Issanchou, 2002) that the strategy efficiency will not be overly distorted by choosing experimental conditions that are slightly different from the computed values.

The initial concentration of soda is always maximal: more information is obtained at the maximum value of soda concentration because experiments starting at an initial soda concentration of 40 mol m^{-3} are included in those starting with the maximal soda concentration. The temperature of the reactor is either minimal or maximal. It can be observed that measurement times are in accordance with the kinetics of the phenomena. At the lower bound of the temperature (slow kinetic), measurement times are about 2200 s, whereas they are about 1000 s when the temperature is at the upper bound (fast kinetic). The strategy includes an experiment with a very short measurement time when mass-transfer coefficients are to be estimated, which favors their precise estimation. On the other hand, very long measurement times lead to a precise estimation of distribution coefficients as well as a reduction of correlation between mass-transfer coefficients and distribution coefficients, which means extra-diagonal terms of $\mathbf{V_p}$. In all cases, it has been shown (Issanchou, 2002; Issanchou et al., 2003) that designed experiments are such that the derivatives of the observed model outputs with respect to the parameters might be important, that is, designed experiments lead to a greater quantitative identifiability of the parameters.

After each experiment, the model parameters are identified and then used to compute the next experiment according to the D-optimal strategy. It was previously shown (Issanchou et al., 2003) that the strategy rapidly decreases the confidence intervals of parameters. Herein, the objective is to focus on the final model and to demonstrate its accuracy. Ten experiments have been considered, a number that corresponds to a stabilization of the identified values of the parameters, together with narrow confidence intervals, as shown in Table 4.

Table 2. Initial Experiments

Experiment	$C_B^{c,0}$ (mol m^{-3})	T (K)	t_r (s)
1	120.6	289.0	3000
2	120.6	298.5	2400
3	121.0	307.0	1800

Table 3. D-Optimal Experiments

Experiment	$C_B^{c,0}$ (mol m^{-3})	T (K)	t_r (s)
4	120.3 (120)	308.5 (308)	990 (983)
5	121.4 (120)	288.7 (288)	2180 (2179)
6	120.1 (120)	288.1 (288)	2180 (2184)
7	120.9 (120)	308.5 (308)	930 (930)
8	120.9 (120)	288.1 (288)	2210 (2210)
9	122.3 (120)	308.5 (308)	330 (330)
10	120.7 (120)	308.0 (308)	5400 (5400)

Measurement variance

Figure 4 shows the evolution of the measurement standard deviation for ester, soda, and pentanol, during the first step of the strategy. In this first step, mass-transfer coefficients values were calculated according to Eq. 6. Identification of mass-transfer coefficients was added for the two remaining D-optimal–designed experiments.

Estimated parameters and confidence intervals

Estimates and their 95% confidence intervals (CI) are reported in Table 4, after three initial experiments and after the 10 D-optimal–designed experiments.

The D-optimal strategy leads to a substantial reduction of the uncertainty of all the parameters of the model and the confidence intervals were reduced by a 10% order of magnitude. Whatever the physicochemical constants are, their activation energies are always better identified than their frequency factors because the model outputs are less sensitive to the later ones.

To evaluate the quality of the estimation, estimates have been compared with values found in the literature, that is, assuming quite similar operating conditions (cf. Table 5).

Values seem to be of the same order of magnitude. The kinetic constant of the reaction estimated in this work is slightly greater than the value estimated by Alwan et al. (1983) and Hiraoka et al. (1990), which is probably explained by the fact that these authors did not take into account the diffusion of synthesized pentanol toward the organic phase. In the same way, a mass-transfer resistance of pentanol in the organic film may explain that the mass-transfer coefficient of pentanol is smaller than the correlated value.

Table 4. Estimates and 95% Confidence Intervals (CI) after 3 and 10 Experiments

	After 3 Experiments		After 10 Experiments	
Parameter	Estimate	CI (%)	Estimate	CI (%)
k_1^0 (m^3 mol^{-1} s^{-1})	114,150	494	700,525	45
T_{k_1} (K)	6207	24	6769	2
m_A^0 (—)	4051	358	162,514	43
$T_{m,A}$ (K)	628	169	1697	7
m_C^0 (—)	804,969	866	270,146	66
$T_{m,C}$ (K)	3154	83	2841	7
$10^5\ k_A^c$ (m s^{-1})	3.1	Correlated	10.8	8
$10^5\ k_C^c$ (m s^{-1}	3.6	Correlated	0.3	10

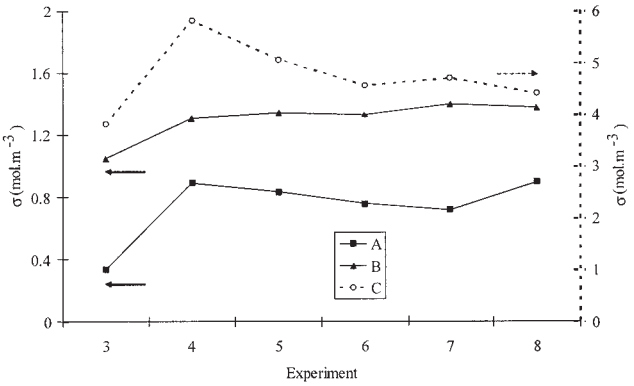

Figure 4. Evolution of standard deviation of measurements during the strategy.

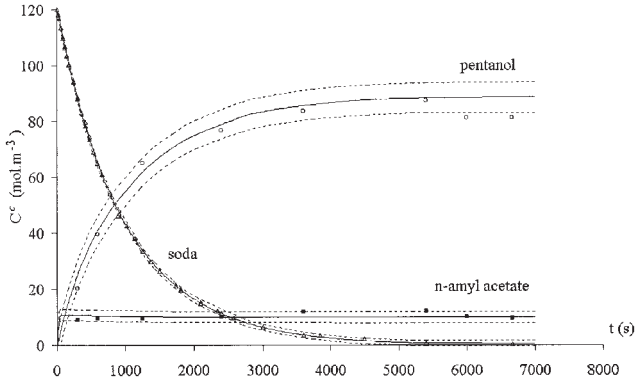

Figure 5. Prediction model for $T = 298$ K and $C_B^0 = 120$ mol m^{-3}.

Model validation

Except for one experiment that was carried out at $T = 298$ K during the initialization step, all experiments were carried out at the lower and upper bounds of temperature. In the same way, all experiments were done with a maximal initial concentration of soda. To verify the ability of the model to properly account for the effect of both these factors, three additional experiments were carried out at $T = 298$ K and $C_B^{c,0} = 120$ mol m^{-3} and one additional experiment was carried out at $T = 298$ K and $C_B^{c,0} = 40$ mol m^{-3}. Figures 5 and 6 show, respectively, for these two sets of experimental conditions, the comparison of experimental data with simulated data together with computed 95% prediction intervals.

Both figures show a very satisfactory prediction of the system outputs over the entire range of observation times. Furthermore, the computed prediction intervals seem to be very reliable, which denotes a good estimation of the model parameters, the variance of the noise, and the variance of the parameter estimates. Finally, it should be noted that better results might be obtained at $C_B^{c,0} = 40$ mol m^{-3} by taking into account the influence of the ionic strength, which was neglected in the model.

Example of a dynamic productivity criterion optimization

Let us suppose now that we want to optimize the productivity of pentanol present in the aqueous phase by determining the isothermal reactor temperature T, the initial concentration of soda $C_B^{c,0}$, and the reaction time t_r that maximize the following criterion:

$$P = \frac{C_C^c}{t_r + t_0} \tag{15}$$

with $t_0 = 1800$ s corresponding to extra reaction times.

The SQP algorithm was used to determine the operating conditions that optimize this criterion in the same experimental space, in which the value of pentanol concentration in the aqueous phase was computed by the model we obtained at the end of the experimental strategy. The operating conditions we obtained are $T = 308$ K, $C_B^{c,0} = 120$ mol m^{-3}, and $t_r = 940$ s. An experiment was then performed at $T = 308$ K and $C_B^{c,0} = 120$ mol m^{-3} and we observed the evolution of the predicted criterion and its experimental measured values vs. time (Figure 7).

Consequently, in that case study, the model we obtained at the end of the strategy provided an accurate determination of the "optimal operating conditions" of the process. Furthermore, the use of a knowledge-driven model might be very powerful to search for optimal operating conditions when scaling-up.

Conclusions

An optimum experimental design strategy for precise and rapid kinetic modeling of liquid–liquid batch reactions has been developed. The case of alkaline hydrolysis of *n*-amylacetate has been used to illustrate the methodology. A D-optimal design has been determined and performed to reduce the volume of the confidence ellipsoid of the parameter estimates of the proposed model. Experiments have been designed in a fully sequential way to increase the robustness of the strategy by using a nominal value of the parameters as accurate as possible.

Table 5. Comparison of Estimates with Literature Data

Parameter	Estimate	Literature	Commentary
k (25°C) (m^3 mol^{-1} s^{-1})	0.096	0.082–0.085*	*Ref: Alwan et al. (1983); Hiraoka et al. (1990)
k (30°C) (m^3 mol^{-1} s^{-1})	0.139	0.113*	Ionic strength $I \approx 40$–50 mol m^{-3} Diffusion of pentanol in the organic phase not taken into account
m_A (—) ($I = 120$ mol m^{-3})	547	540**	**Computed for $T = 25$°C Influence of ionic strength given in Alwan et al. (1983); Nanda and Sharma (1966)
m_C (—)	20	37†	†Computed from water solubility given in Howard and Meylan (1997)
$10^5\ k_A^c$ (m s^{-1})	10.8	3.1‡	‡Correlated values (Eq. 6)
$10^5\ k_C^c$ (m s^{-1})	0.3	3.6‡	

It has been observed from the analysis of the computed design that only specific areas of the admissible experimental domain are able to provide very informative data to precisely estimate the model parameters. In this particular case, use of the D-optimum design has provided accurate estimates, even if only a small number of experiments have been carried out. Furthermore, the model has shown interesting properties in generalization and given a reliable prediction of the system state under several operating conditions.

Finally, the strategy should be an innovative alternative to the use of black-box models that, although they effectively reduce the experimental effort, show little performance in extrapolation.

Notation

A = interfacial area, $m^2\ m^{-3}$
c = constant in Eq. 10
C_i = concentration of component i, mol m^{-3}
d_a = stirrer diameter, m
d_r = reactor diameter, m
D_i = diffusion coefficient of component i, $m^2\ s^{-1}$
$\mathbf{e}$ = vector of residuals
I = ionic strength, mol m^{-3}
j_{wls} = weighted least-squares criterion
k_j = chemical reaction rate constant of reaction j, $(\text{mol m}^{-3})^{(1-\alpha jk)}\ s^{-1}$
k_j^0 = frequency factor of chemical reaction rate constant of reaction j, $(\text{mol m}^{-3})^{(1-\alpha jk)}\ s^{-1}$
k_i^c = mass transfer coefficient of component i in the continuous phase, m s^{-1}
L = liquid height, m
m_i = distribution coefficient of component i
m_i^0 = frequency factor of distribution coefficient of component i
$\mathbf{M}_f$ = Fisher information matrix
n = number of experiments
n_b = number of blades of the stirrer
nc = number of components
np = number of parameters
nr = number of chemical reactions
N_a = stirrer speed, s^{-1}
N_i = diffusion molar flow density of component i, mol $m^2\ s^{-1}$
$\mathbf{p}$ = vector of parameters
P = productivity criterion in Eq. 15, mol $m^{-3}\ s^{-1}$
r_i = chemical reaction generation rate of component i
t = time, s^{-1}
t_0 = constant in Eq. 15, s^{-1}
t_r = measurement time, s^{-1}
T = temperature, K

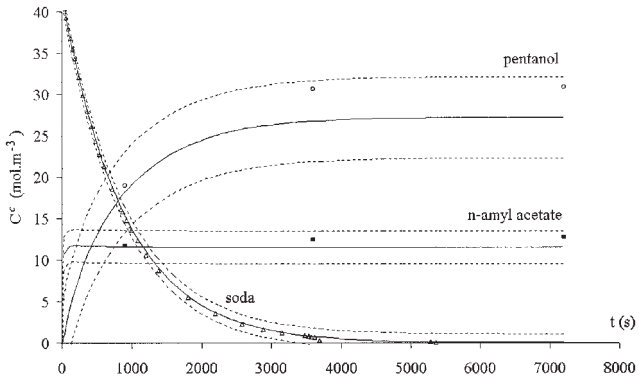

Figure 6. Prediction model for T = 298 K and C_B^0 = 40 mol m^{-3}.

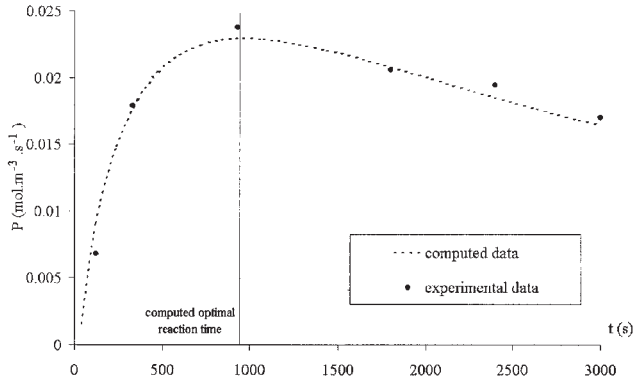

Figure 7. Evolution of the productivity criterion vs. time.

$T_{k,j}$ = modified activation energy of chemical reaction rate constant of reaction j, K
$T_{m,j}$ = modified activation energy of distribution coefficient of component I, K
V = volume, m^3
$\mathbf{V}_p$ = variance–covariance matrix of the parameters
w = stirrer width, m
$\mathbf{x}$ = vector of operating conditions
$\mathbf{X}$ = admissible experimental domain
$\mathbf{y}^e$ = vector of observed experimental outputs
$\mathbf{y}^m$ = vector of computed model outputs

Greek letters

α_{ji} = partial order of component i in reaction j
ε = vectorial stochastic variable
η = kinematic viscosity of the solution, $m^2\ s^{-1}$
ν_{ji} = stoichiometric coefficient of component i in reaction j
σ = standard deviation
Σ = variance–covariance matrix of errors
ϑ_i = molar volume of component i, $m^3\ mol^{-1}$

Superscripts

A = relative to *n*-amylacetate
B = relative to soda
C = relative to *n*-pentanol
D = relative to sodium methanoate
t = true values

Superscripts

* = estimated values
0 = initial condition
c = continuous phase
d = dispersed phase
D = D-optimal
I = interface

Literature Cited

Alwan, S., S. Hiraoka, and I. Yamada, "Extraction of *n*-Amylacetate with Alkaline Hydrolysis in Aqueous Phase," *Chem. Eng. Commun.*, **22**, 317 (1983).

Argawal, A. K., and M. L. Brisk, "Sequential Experimental Design for Precise Parameter Estimation," *Ind. Eng. Chem. Process. Des. Dev.*, **24**, 203 (1985).

Atkinson, A. C., and W. G. Hunter, "The Design of Experiments for Parameter Estimation," *Technometrics*, **10**(2), 271 (1968).

Baltes, M., R. Schneider, C. Sturm, and M. Reuss, "Optimal Experimental Design for Parameter Estimation in Unstructured Growth Models," *Biotechnol. Prog.*, **10**(5), 480, (1994).

Benoist, D., Y. Tourbier, and S. Germain-Tourbier, *Plans d'Expériences: Construction et Analyse*, Lavoisier, Paris (1994).
Box, G. E. P., and H. Lucas, "Design of Experiments in Non-Linear Situations," *Biometrika*, **46**, 77 (1959).
Chen, H. T., and S. Middleman, "Drop Size Distribution in Agitated Liquid–Liquid Systems," *AIChE J.,* **13**, 989 (1967).
Hindmarsh, A. C., "LSODE and LSODI, Two New Initial Value Ordinary Differential System Solver," *ACM SIGNUM*, **15**, 11 (1980).
Hiraoka, S., Y. Yamada, Y. Tada, H. Mori, N. Narita, H. Suzuki, and T. Aragaki, "Measurement of Continuous Phase Mass Transfer Coefficient at Droplet Surface in Liquid—Liquid Mixing Vessel by Chemical Reaction Method," *J. Chem. Eng. Jpn.*, **2**, 166 (1990).
Hosten, L. H., and G. Emig, "Sequential Experimental Design Procedures for Precise Parameter Estimation in Ordinary Differential Equations," *Chem. Eng. Sci.*, **30**, 1357 (1975).
Howard, P. H., and W. M. Meylan, *Handbook of Physical Properties of Organic Chemicals*, CRC Press, Boca Raton, FL (1997).
Ingold, C. K., *Structure and Mechanism in Organic Chemistry*, Cornell Univ. Press, Ithaca, NY (1953).
Issanchou, S., *Stratégie expérimentale pour la determination de modèles cinétiques en milieu liquide–liquide*, PhD Thesis, Toulouse, France: Institut National Polytechnique de Toulouse (2002).
Issanchou, S., P. Cognet, and M. Cabassud, "Precise Parameter Estimation for Chemical Batch Reactions in Heterogeneous Medium," *Chem. Eng. Sci.*, **58**, 1805 (2003).
Juusola, J. A., D. W. Bacon, and J. Downie, "Sequential Statistical Design Strategy in an Experimental Kinetic Study," *Can. J. Chem. Eng.*, 50, 796 (1972).
Kiefer, J., "General Equivalence Theory for Optimum Designs (Approximate Theory)," *Ann. Stat.*, **2**(5), 849 (1974).
Lewis, W. K., and W. G. Whitman, "Principles of Gas Absorption," *Ind. Eng. Chem.*, **16**, 1215 (1924).
Lundström, L., T. Salmi, J. Lehtonen, L. P. Lindfors, S. Toppinen, and J. Aittamaa, "Modelling of Specialty Chemicals Production in Liquid–Liquid Reactors—A Case Study: Synthesis of Diols," *Chem. Eng. Sci.*, **54**, 1 (1999).
March, J., *Advanced Organic Chemistry*. 4th Edition. Wiley, New York, NY (1992).
Marquardt, D. W., "An Algorithm for Least Squares Estimation on Nonlinear Parameters," *J. Soc. Ind. Appl. Math.*, **11**, 431 (1963).
Mezaki, R., "Precise Parameter Estimation in Reaction Modelling-Sensitivity in Choosing Experimental Conditions," *Can. J. Chem. Eng.*, **47**, 192 (1969).
Nagata, S., M. Nishikawa, and T. Takimoto, "Jacket Side Heat Transfer Coefficient in Mixing Vessel," *Kagaku Kogaku*, **35**, 924 (1971).
Nanda, A. K., and M. M. Sharma, "Effective Interfacial Area in Liquid–Liquid Extraction," *Chem. Eng. Sci.*, **21**, 707 (1966).
Reilly, P. M., R. Bajramovic, G. E. Blau, D. R. Branson, and M. W. Sauerhoff, "Guidelines for the Optimal Design of Experiments to Estimate Parameters in First Order Kinetic Models," *Can. J. Chem. Eng.*, **55**, 614 (1977).
Rippin, D. W. T., L. M. Rose, and C. Schifferli, "Non-Linear Experimental Design with Approximate Models in Reactor Studies for Process Development," *Chem. Eng. Sci.*, **35**(1), 356 (1980).
Schitkowski, K., "A Fortran Subroutine Solving Nonlinear Programming Problems," *Ann. Oper. Res.*, **5**, 485 (1986).
Sedrati, Y., M. Cabassud., M.-V. Le Lann, and G. Casamatta, "Sequential Experimental Design Strategy for Kinetic Parameter Estimation," *Comput. Chem. Eng. Suppl.*, S427 (1999).
Sorenson, H. W., *Parameter Estimation, Principles and Problems*, Marcel Dekker, New York, NY (1980).
Walter, E., and L. Pronzato, "Qualitative and Quantitative Experiment Design for Phenomenological Models—A Survey," *Automatica*, **26**, 195 (1990).
Walter, E., and L. Pronzato, *Identification de modèles paramétriques à partir de données expérimentales*, Masson, Paris (1994).
Wilke, C. R., and P. Chang, "Correlation of Diffusion Coefficients in Dilute Solutions," *AIChE J.*, **1**, 26 (1955).
Zacks, S., *Parametric Statistical Inference, Basic Theory and Modern Approaches*, Pergamon Press, London (1981).

Manuscript received Nov. 14, 2003, and revision received Oct. 21, 2004.