Korean J. Chem. Eng., **18**(5), 652-661 (2001)

On the Feasibility and Reliability of Nonlinear Kinetic Parameter Estimation for a Multi-Component Photocatalytic Process

Linxiang Wang* and Chang Nyung Kim†

*Industrial Liaison Research Institute, College of Mechanical & Industrial System Engineering,
Kyung Hee University, Yongin 449-701, Korea

(*Received 23 April 2001 • accepted 18 June 2001*)

Abstract−Nonlinear kinetic parameter estimation plays an essential role in kinetic study in reaction engineering. In the present study, the feasibility and reliability of the simultaneous parameter estimation problem is investigated for a multi-component photocatalytic process. The kinetic model is given by the L-H equation, and the estimation problem is solved by a hybrid genetic-simplex optimization method. Here, the genetic algorithm is applied to find out, roughly, the location of the global optimal point, and the simplex algorithm is subsequently adopted for accurate convergence. In applying this technique to a real system and analyzing its reliability, it is shown that this approach results in a reliable estimation for a rather wide range of parameter value, and that all parameters can be estimated simultaneously. Using this approach, one can estimate kinetic parameters for all components from data measured in only one time experiment.

Key words: Photocatalytic Oxidation, Kinetic Model, Parameter Estimation, Nonlinear Optimization, Hybrid Method

INTRODUCTION

Photocatalysis is a promising approach to efficient destruction of environmental pollutants [Alberici and Jardim, 1997; Herrmann, 1999; Ollis, 2000; Chai et al., 2000], and many experimental investigations have been reported in this field [Anpo et al., 1997; Sirisuk et al., 1999]. Because an experimental approach is usually costly and time consuming, some other investigations with computer simulations have also been carried out, which are based on a mathematical kinetic model for the photocatalytic oxidizing process. Researches have shown that most photocatalytic reactions follow the L-H (Langmuir-Hinshelwood) equation [Fox and Dulay, 1993]. In fact, numerical simulation of the L-H equations is not a difficult task in itself, and here various numerical integral methods can be applied to the equation to solve concentration profiles of each reactant and product with time. The obstacle is that kinetic parameters in the L-H equations, that is, reaction rate constants and adsorption equilibrium constants, are not measurable, and there is no way to deduce an analytical formula to theoretically estimate its value. The only feasible approach to obtaining their values is to estimate them based on experimental data. The reason is that the L-H equations are a set of implicit nonlinear differential equations that are coupled with each other since some reactants are the resultants of others in a multi-component system.

During a long period, kinetic parameters have been estimated by the so-called initial rate method [Levenspiel, 1972], which uses linear regression method, based on the reciprocal form of a single L-H equation. But, this method cannot yield satisfying results owing to the fact that a nonlinear equation is merely replaced by a linearized equation in this method [Mehrab et al., 2000]. When more than one component is being oxidized simultaneously and the L-H equations are coupled with each other, which is very often encountered in photocatalytic reactions, this method is no longer applicable because reciprocals of the L-H equations are still nonlinear.

Taking into account the nonlinearity of the L-H equations, several approaches have been suggested for nonlinear parameter estimation in reaction engineering [Biegler et al., 1986; Kim et al., 1990; Farza et al., 1997; Park and Froment, 1998; Oh et al., 1999; Balland et al., 2000]. For kinetic model governed by the L-H equations, Froment [1987] has shown that nonlinear regression can be applied to perform nonlinear parameter estimation. Mehrab et al. [2000] have adopted the Box-Draper nonlinear regression method to find the best point estimates, in which a variable metric algorithm is employed with an improved gradient calculation. Although the local-convergence methods mentioned above do have a potential to yield a better estimation of kinetic parameters and are expected to be usable in multi-component systems, there is still a rigorous limitation that a rather good initial parameter value should be given. Because an objective function for nonlinear model often contains more than one optimum, a local-convergence method is highly prone to fall into non-global optima [Press, 1986] owing to their downhill (hill-climbing) algorithm. To protect the parameter estimation from regarding a local optimum as a global one, Belohlav et al. [1997] have applied a random search method in nonlinear regression. This approach does work, but is computationally less efficient because of its random search algorithm. Especially when more than one component is oxidized, it hardly results in a satisfying estimation.

To locate the global optimum confidently, various approaches under the term "evolutionary algorithm" have been also investigated recently. Wolf and Moros [1997] have estimated rate constants in oxidizing methane to C_2 hydrocarbons by the Genetic Algorithm (GA); Park and Froment [1998] have used the GA estimated kinetic parameters and tested a heterogeneous catalytic reaction; Balland et al. [2000] have estimated kinetic and energetic parameters in the saponification process of ethyl acetate using the GA.

---

†To whom correspondence should be addressed.
E-mail: cnkim@khu.ac.kr

Although, the kinetic models studied in these works are not given by the L-H equation, their nonlinearities are not as strong as that of the L-H equation, and their final results are not so accurate as those of local convergence method, these works do show us that the GA has a potential to find out, roughly, the location of the real global optimum for the nonlinear estimation problem.

In the present study, we will show it is feasible and reliable to estimate simultaneously all kinetic parameters in the L-H equations by a hybrid genetic-simplex optimization method, even when several oxidation processes are coupled with each other in the reaction. The proposed hybrid optimization method is set up in mating the modified GA with the simplex algorithm. The role of the modified GA is to find a rough estimation for kinetic parameters, and this assures us that subsequent local search will converge to the global optimum when the result of the rough estimation is used as a starting point of the simplex algorithm. The simplex algorithm is a local convergence method used to refine the rough estimation and to make the estimation more accurate. By analyzing the sensitivity of the simulated concentrations with respect to the estimated parameter values, we will show that the hybrid optimization method is able to estimate parameters accurately and reliably. By applying it to a process about which previously published results are available, we will show the hybrid method gives a more accurate estimation.

PARAMETER ESTIMATION OF MULTICOMPONENT SYSTEM

In the present study, a process of photocatalytic purification of three VOC components, which has been extensively investigated by Turchi and Rabago [1995], Turchi et al. [1996] and Wolfrum et al. [1997], will be adopted for discussion. In the system, there are three kinetic-significant reactants (acetone, isopropanol and methanol) and one product (carbon dioxide), where acetone is also the product of isopropanol oxidation. The three reactants existing in the same contaminated air stream, say, representing exhaust stream from semiconductor plants, are oxidized simultaneously. The reaction stoichiometry can be depicted as [Turchi et al., 1996]:

Acetone→3 CO_2+xH_2O
Isopropanol→acetone→3 CO_2+xH_2O
Methanol→CO_2+xH_2O

The L-H equations for this process can be expressed as:

$$r_1=\frac{dc_1}{dt}=-\frac{k_1K_1c_1-k_2K_2c_2}{1+\sum_{i=1}^{3}K_ic_i}$$

$$r_2=\frac{dc_2}{dt}=-\frac{k_2K_2c_2}{1+\sum_{i=1}^{3}K_ic_i}$$

$$r_3=\frac{dc_3}{dt}=-\frac{k_3K_3c_3}{1+\sum_{i=1}^{3}K_ic_i}$$

$$r_4=\frac{dc_4}{dt}=\frac{3k_1K_1c_1+k_3K_3c_3}{1+\sum_{i=1}^{3}K_ic_i} \tag{1}$$

where subscript 1 denotes acetone, 2 isopropanol, 3 methanol and 4 carbon dioxide. c is the concentration, k the reaction rate constants and K the adsorption equilibrium constants.

In the previous literature [Turchi et al., 1996; Wolfrum et al., 1997], k_i and K_i, i=1, 2, 3, have been estimated separately from experiments with each individual reactant having its various initial concentrations. In the present study, with only one set of data, the kinetic parameters in the above equations will be estimated by a hybrid optimization method.

To fulfill the nonlinear estimation, let's construct object functions for four components in the system as follows:

$$J_i(\mathbf{k},\mathbf{K})=\sum_{m=1}^{p}[c_i(m)-\hat{c}_i(m)]^2 \tag{2}$$

where vector $\mathbf{k}=\{k_1, k_2, \ldots, k_m, \ldots\}$ denotes all reaction rate constants and vector $\mathbf{K}=\{K_1, K_2, \ldots, K_m, \ldots\}$ denotes all adsorption equilibrium constants. $c_i(m)$ is experimental concentration of the *i*-th component at the *m*-th instance with time, while $\hat{c}_i(m)$ is the simulated concentration of the *i*-th component at the *m*-th instance. Apparently, the object function is a square summation of difference between the simulated and experimental concentrations (Least Square Estimation). To make the problem solvable, it is convenient to combine all object functions into a single total object function by the following weighted average (or summation) method.

$$J=\sum_{i=1}^{n}w_iJ_i/\sum_{i=1}^{n}w_i \tag{3}$$

where w_i is the weight for the *i*-th component. Thus, the kinetic parameter estimation problem is expressed as a nonlinear optimization problem as follows:

$$\begin{cases}\min\limits_{(k,K)} J\\ \dfrac{dc_i}{dt}=f_i(k_1,K_1,c_1,k_2,K_2,c_2,\Lambda)\end{cases} \tag{4}$$

Apparently, since the L-H equations are a set of implicit nonlinear ordinary differential equations, the object function must be a multipeak function. Therefore, the only practical way to solve this problem is numerical search method.

HYBRID OPTIMIZATION METHOD

It has been understood that the GA has the potential to locate the global optimum but its final result may not be accurate enough, while the local-convergence method has the potential to locate the local optimum accurately but is highly prone to fall into non-global optimum. Therefore, it is natural to expect that a hybrid algorithm set up by mating the GA with the local convergence method should be a promising approach for nonlinear kinetic parameter estimation. In the following paragraph a brief description will be given about the GA and the simplex algorithm, which are hybridized for the present study. Detailed information about these algorithms can be found in many previous investigations [Cheney and Kincaid, 1985; Winston, 1991; Gen and Cheng, 1997].

The GA maintains a population of individuals, say P(n), for generation n and each individual consists of a set of genes, where each gene stands for a parameter to be estimated. One individual represents one potential solution to the problem at hand. Each individual is evaluated to give some measure of its fitness. Some individuals undergo stochastic transformation by means of genetic operations to form new individuals. There are two transformations: crossover, which creates new individuals by combining parts from two individuals, and mutation, which creates a new individual by making changes in a single individual. New individuals, called offspring S(n), are then evaluated. A new population is formed by selecting fitter individuals from the parent population and the offspring population. After some generations, the algorithm converges to the best individual, which hopefully represents an optimal or sub-optimal solution to the problem. A general structure of the GA is as follows:

```
begin
    n=0;
    initialize P(n);
    evaluate P(n);
    while (not termination condition) do
    begin
        recombine P(n) to yield S(n);
        evaluate S(n);
        select P(n+1) from P(n) and S(n);
        n=n+1;
    end
end
```

To be applied to this concrete problem, some operations of the GA are modified here.

1. Modified Genetic Algorithm

In the present study, variables to be optimized are reaction rate constants $\mathbf{k}=\{k_1, k_2, \ldots, k_m, \ldots\}$, and adsorption equilibrium constants $\mathbf{K}=\{K_1, K_2, \ldots, K_m, \ldots\}$, which are all float values, and this fact makes the coding procedure different from that of the traditional GA, whose individuals are all coded by binary figures.

1-1. Representation

Using real possible values of the reaction rate constants and the adsorption equilibrium constants as its gene, the *j*th individuals $\mathbf{V}_j$ can be encoded as:

$$\mathbf{V}_j\text{: }\{k_1^j, k_2^j, \ldots, k_n^j, K_1^j, K_2^j, \ldots, K_n^j\}$$

$$lk \le k_i^j \le uk;\ lK \le K_i^j \le uK$$

where k_i^j is the *j*-th possible value of reaction rate constant of the *i*th component, while K_i^j is the *j*-th possible value of adsorption equilibrium constant of *i*-th component in the current generation. *l*k and uk are the lower and upper limit, respectively, for reaction rate constants, and *l*K and uK are the lower and upper limit, respectively, for adsorption equilibrium constants. It is also feasible to set different limits for each variable.

1-2. Crossover and Mutation

In order to explore the search space, some randomly chosen individual pairs are recombined by crossover operation, which is clearly sketched by the following two in- dividuals

$$\mathbf{V}_j\text{: }\{k_1^j, k_2^j, \ldots, k_n^j, K_1^j, K_2^j, \ldots, K_n^j\}$$

$$\mathbf{V}_{j+1}\text{: }\{k_1^{j+1}, k_2^{j+1}, \ldots, k_n^{j+1}, K_1^{j+1}, K_2^{j+1}, \ldots, K_n^{j+1}\}$$

New individuals resulting from the above may be:

$$\mathbf{V}_j^*\text{: }\{k_1^j, k_2^j, \ldots, k_n^j, K_1^j, K_2^{j+1}, \ldots, K_n^{j+1}\}$$

$$\mathbf{V}_{j+1}^*\text{: }\{k_1^{j+1}, k_2^{j+1}, \ldots, k_n^{j+1}, K_1^{j+1}, K_2^j, \ldots, K_n^j\}$$

The position where an individual is cut off for recombination is randomly chosen, and the number of individuals chosen to be recombined is set by crossover probability. After crossover, all entries (genes) in all individuals are given a chance of undergoing mutation operation with a certain mutation probability. If a gene is selected for mutation operation, it will be assigned a random value in the given range.

1-3. Evaluation and Selection

The evaluation function plays the role of the environment in natural evolution, and it rates individuals in terms of their fitness. For the minimization problem here, the fitness of each individual is defined over object function values as:

$$F_j = \frac{\sum_{j=1}^{m} J(\mathbf{v}_j)}{J(\mathbf{v}_j)} \tag{5}$$

where F_j is the fitness of the *j*-th individual v_j, m is the number of individuals in the present generation, and $J(\mathbf{V}_j)$ is the object function value of the *j*-th individual. It is apparent that the individual having the smallest object function value will have the highest fitness. In constructing the next generation, the probability of the selection of the *j*th individual is calculated by the following equation:

$$P_j = \frac{F_j}{\sum_{j=1}^{m} F_j} \tag{6}$$

The selection is implemented by adopting a roulette wheel approach [Gen and Cheng, 1997].

2. Simplex Algorithm

From the mathematical point of view, the simplex algorithm is a relatively simple algorithm, but is effective for many optimization problems, especially in case it is difficult to deduce an analytical gradient formula [Cheney, Kincaid, 1985]. Its principle is as follows:

When the simplex algorithm starts with a given starting point in an m dimensional space, it will choose arbitrarily one point different from the starting point along each dimensional axis in a small neighborhood of the starting point, and define a simplex in the search space with the given point (starting point) together with the m points chosen along each dimensional axis (m is the number of parameters to be optimized, i.e., the dimension of search space, and the simplex has m+1 points). Then a downhill (or hill-climbing) method is applied to update the simplex iteratively and to make the algorithm finally achieve the optimum. Assume that $\mathbf{v}_j$ is the worst one among the m+1 points $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m+1}$, which means $\mathbf{v}_j$ has the biggest object function value (in case of minimizing); a new point can be created as follows:

$$\mathbf{v}^* = (1-\lambda)\mathbf{v}_j + \lambda \frac{\sum_{i=1}^{m+1} \mathbf{v}_i - \mathbf{v}_j}{2n} \tag{11}$$

where λ is the step size which can be optimized in each iteration. By replacing the worst point $\mathbf{v}_j$ with the newly created point $\mathbf{v}^*$, the simplex can be updated. In the next iteration, the new worst point in the new simplex is identified and replaced by a new point. If this updating operation is repeated, all the m+1 points will come closer to the local optimum, and the step size will become smaller [Winston, 1991]. When the step size becomes smaller than a given small value (convergence criteria), each one of the m+1 points can be regarded as the optimum.

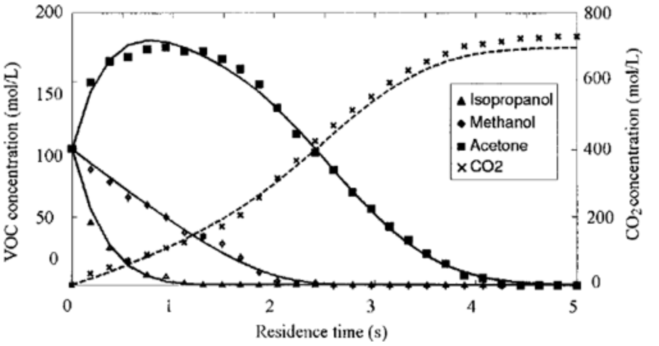

Fig. 1. Comparison of experimental observations with the simulated concentrations based on the estimated parameters by the hybrid optimization method (All experimental data are the same as those in Fig. 2).

3. Solution Strategy

In the simplex algorithm, the simplex is updated by replacing the single worst point (the current point) in the search space iteratively. During a single iteration, a new point is selected from the neighborhood of the current point. If the new point provides a smaller value of the object function, the current point is deleted and the new point will be used in the process of the simplex algorithm; otherwise, another neighbor is selected and tested. It is clear that this search strategy provides local optimum values only, and these values depend on the selection of the starting point. Considering the fact that the optimization problem in the present study is implicit and nonlinear, which means that the object function generally has more than one extrema, we have to produce a reasonable starting point to ensure the simplex algorithm converges to a global optimum rather than a local optimum. On the other side, the GA is a compromise between an accurate local convergence method and a robust random method; it combines elements of directed and stochastic search. It has been shown by many investigations that the GA will finally converge to the best individual roughly with a random initial population, but it is hard to improve its accuracy. In our hybrid method for kinetic parameter estimation, the genetic algorithm is used not to find the final best solution to the problem, but to yield a rough guess of parameters, which will be used as the starting point of the simplex algorithm; therefore. it is unnecessary to let the GA operate a long time till it finally converges to the best solution.

In our hybrid optimization method, the GA is adopted firstly with its initial individuals created randomly in a given bound of parameter values. About 2000 generations are evolved, and the fittest individual in the population (from the first generation to the last generation) is chosen as the favorite point. Here, the number of generations used in the GA is an empirical value based on the experience of the author because there is no general value for this purpose. Successively, the simplex algorithm is started to refine the solution by using the point given by the GA as its starting point. When the step size of the simplex algorithm is smaller that 1.0e-6, the search process is stopped and the optimization is regarded as converged.

PERFORMANCE OF THE HYBRID METHOD

By the hybrid optimization method mentioned in the above section, all kinetic parameters in the system defined by Eq. (1) have been estimated simultaneously by only one set of data, which is read carefully from the published experimental figure [Turchi et al., 1996]. For the sake of convenience, the units of concentration, reaction rate constants and adsorption equilibrium constants are all changed into SI system. In this test with the modified GA, the population has 200 individuals, the crossover probability is set 0.2 and mutation probability is set 0.01. The convergence criterion for the simplex algorithm is that its step size is smaller than 1.0×10^{-6}, which means all points of the final stage of the simplex algorithm come close enough and wander in a very small region (approximated by a polyhedron whose diameter is smaller than 1.0×10^{-6}). With the assumption that $0<k_i<10000$ and $0<K_i<10$, i=1, 2, 3 (it is a reasonable guess judging from our knowledge of reaction engineering, and if these bounds do not work, we can easily change these bounds), the hybrid optimization method finally yields the estimated values as follows:

k_1=140.8013766138399, K_1=0.02038825898700673
k_2=7275.7 08194557086, K_2=0.01252516703250473
k_3=68.65429354582977, K_3=0.2569758737264554

Comparison is made between the simulated concentrations obtained from our estimated kinetic parameters and the experimental concentrations in Fig. 1. Because the exact values of the kinetic parameters are unknown, the only way we can show the performance of the hybrid optimization method is to compare our estimated results with those in other published investigations. Here, for the comparison the estimated results made by Turchi et al. [1996] and Wolfrum et al. [1997] are reproduced in Fig. 2, where their parameter estimation was based on single component data (separate estimation for each component). For clarification, data are arranged in the same style in the two figures, where it is very clear that our hybrid optimization method gives a more accurate estimation of the kinetic parameters in the system, because the simulated concentrations based on the hybrid optimization method fit more accurately with the experimental data than those based on separate estimation though the mathematical kinetic models are the same. Furthermore, only 27 experimental concentrations of each reactant are used for our estimation, while it has been reported in the previous literature that a series of experiments for each component are

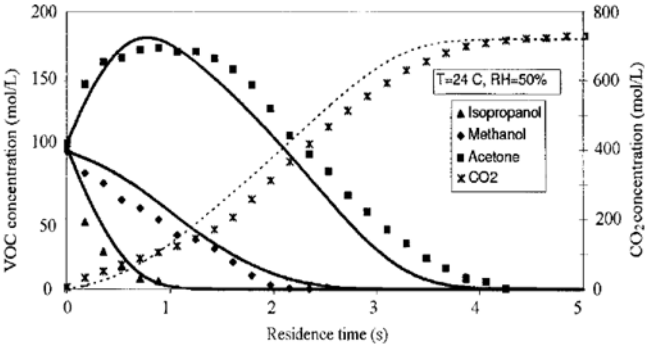

Fig. 2. Comparison made by Turchi et al. [1996] between experimental observations of photocatalytic process given in Eq. (1) and simulated concentrations based on parameter estimation by single component data (T denotes temperature, RH denotes relative humidity).

necessary for their parameter estimation.

This application indicates that the hybrid optimization is feasible for simultaneous estimation of kinetic parameters in the L-H equations for a multi-component system, and this method does give a more accurate estimation than what the separate estimation method does.

RELIABILITY OF THE ESTIMATION

In order to show the reliability of the kinetic parameter estimation, let's analyze the sensitivity of the simulated concentrations with respect to the estimated kinetic parameters, which determines error transfer during the estimating process. To clarify the problem, a single component system is used for discussion, whose L-H equation is:

$$r=\frac{dc}{dt}=\frac{kKc}{1+Kc} \tag{6}$$

where c is concentration, k is the reaction rate constants and K is the adsorption equilibrium constants.

1. Sensitivity

It is known that there must be some measurement error introduced into the measured concentration during experiments, and the measurement error will affect the estimation results. The ratio of error in the estimated parameters to that introduced in measurements is determined by the sensitivity of the estimated parameter with respect to the value of the object function. Because the experimental concentration profile with time in the given system is fixed in the process of parameter estimation (it was used as input), we can consider the sensitivity of the estimated parameters with respect to the simulated concentrations instead. It is not a difficult task to obtain the sensitivity of simulated concentrations from the L-H equations.

$$s_{k,m}=1/\frac{d\hat{c}(m)}{dk}\approx 1/\Delta t\sum_{j=1}^{j=m}\frac{K\hat{c}(j)}{1+K\hat{c}(j)} \tag{7}$$

$$s_{K,m}=1/\frac{d\hat{c}(m)}{dK}\approx 1/\Delta t\sum_{j=1}^{j=m}\frac{k\hat{c}(j)}{[1+K\hat{c}(j)]^2} \tag{8}$$

where $s_{k,m}$ is the sensitivity of k with respect to the m-th simulated concentrations, while $s_{K,m}$ is sensitivity of K with respect to the m-th simulated concentrations. $\hat{c}(j)$ is the simulated concentrations at the j-th instance with time and Δt is time interval between concentration sampling.

Generally, $k>>1$ while $K<<1$; it makes $s_{k,m}>s_{K,m}$ in most cases. This fact means that error in the estimated reaction rate constant is larger than that in the estimated adsorption equilibrium constant, where both the errors are caused by the experimental concentrations error. However, it does not mean that the estimated K is more reliable than estimated k, because their absolute magnitudes are different. If "relative sensitivity" is to be used, we can easily find the error transfer:

$$rs_{k,m}=1/\frac{d\hat{c}(m)/c_{ref}}{dk/k}\approx 1/\Delta t\sum_{j=1}^{j=m}\frac{kK\hat{c}(j)}{[1+K\hat{c}(j)]c_{ref}} \tag{9}$$

$$rs_{K,m}=1/\frac{d\hat{c}(m)/c_{ref}}{dK/K}\approx 1/\Delta t\sum_{j=1}^{j=m}\frac{kK\hat{c}(j)}{[1+K\hat{c}(j)]^2c_{ref}} \tag{10}$$

where $rs_{k,m}$ and $rs_{K,m}$ are the relative sensitivity of the estimated k and K with respect to the concentrations, respectively, and c_{ref} is the reference concentration (we use the initial concentration as the reference here). Because K is small, $1+K\hat{c}(j)$ is a little larger than 1, and $rs_{k,m}<rs_{K,m}$, we can deduce that the estimated k has a little smaller relative error than the estimated K when the same experimental concentrations are used, which means the estimated reaction rate constant is a little more reliable than the estimated adsorption equilibrium constants.

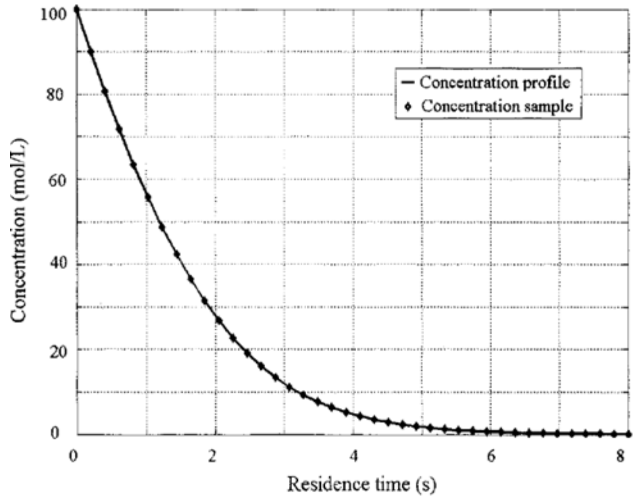

Fig. 3. Simulated concentration of the system given in Eq. (6) and its sampling.

Because the relative sensitivity is a function of the kinetic parameters (to be estimated) and the concentrations, this fact makes it difficult to abstract a general formula of sensitivity analysis for all reaction processes. Instead, we use a concrete numerical test to show how the error introduced into the experimental concentrations will affect the accuracy of the estimated parameters. Let k=100, K=0.01, and the initial concentration is c(0)=100 in the system given by Eq. (6). It is easy to simulate this system and the obtained concentrations are sketched in Fig. 3. To test the reliability of kinetic parameter estimation by the hybrid optimization method, we sample the simulated concentration at 30 different instances along time span evenly from 0 to 9 seconds, which are marked as discrete small circles in the figure. Here we pretend that we know nothing about the values of k and K, and they have to be estimated with the sampled concentrations. In an ideal case (no error exists in the experimental concentrations), the estimated k and K by the hybrid optimization method are

k=100.0000000360223
K=0.009999999993527

The obtained parameters are almost the same as their real values, and the test indicates the hybrid optimization method has the potential to recover kinetic parameters accurately. Now, let's introduce a random error into the input concentrations:

$$c_{input}=c_{exact}+\varepsilon \tag{11}$$

where c_{exact} is concentration without any error (calculated with given parameter values), ε is the mean zero random error introduced into the concentrations with $|\varepsilon|<2$ in this test, and c_{input} is concentration

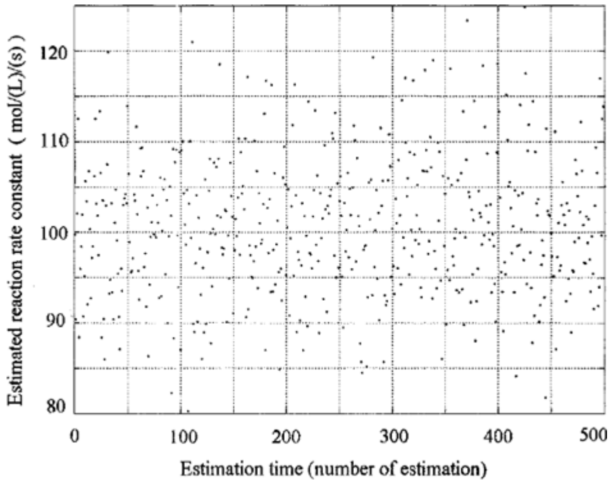

Fig. 4. Distribution of the estimated reaction rate constant with 500 times estimation including random error.

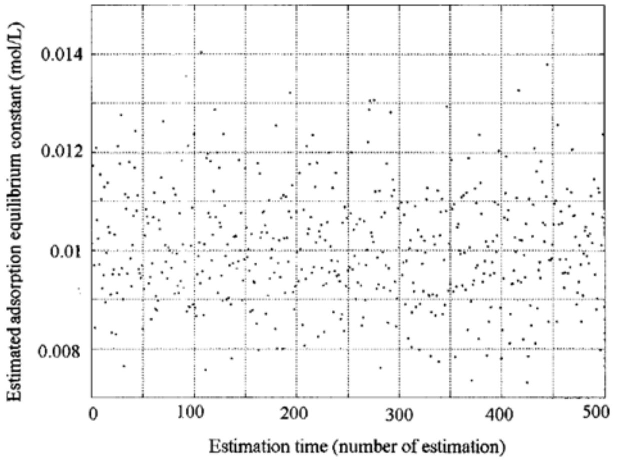

Fig. 5. Distribution of the estimated adsorption equilibrium constants with 500 times estimation including random error.

used for parameter estimation.

Considering the randomness of the introduced error, we run the estimation 500 times to abstract an approximate statistical characteristic. The distributions of estimated reaction rate constant are marked in Fig. 4, while distributions of estimated adsorption equilibrium constant are in Fig. 5. It is very clear from the figures that the estimated k and K are distributed in small bounded regions, respectively, with their average values:

$k=100.5615342848486$
$K=0.010061512537634$

Here, the estimated parameters are very close to their real values. This result shows that the errors in the estimated parameters are bounded, and if enough data are available, the parameters can be recovered accurately.

If only one set of data is available, the reliability of the parameter estimation can be shown by the distribution of the relative error defined in the following equation:

$$e=(x-\bar{x})/\bar{x} \tag{12}$$

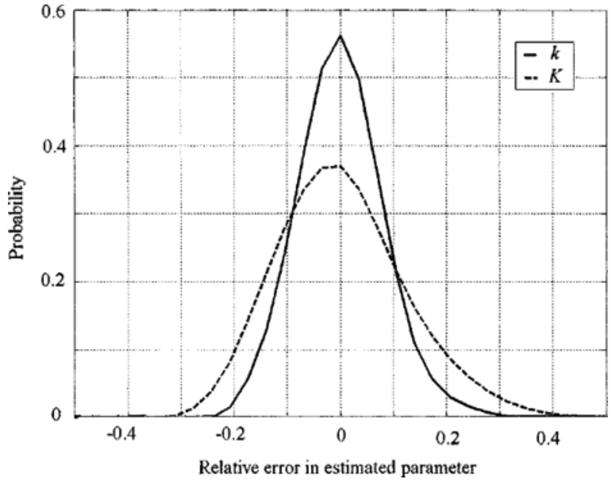

Fig. 6. Distributions of the probabilities of the estimated parameters with respect to relative error when random error is introduced into the system (based on 500 estimations).

where e is the relative error and x denotes either k or K. $\bar{x}$ denotes the average of x. Here, to show the reliability of the estimated results, we use a probability density function with respect to the relative error of the estimated results, where the probability density function is obtained by counting the number of points and by smoothing the resulting probabilities to a curve. The probability density functions of k and K are sketched together in Fig. 6 for convenience of comparison.

From the probability density functions, we can easily draw two bottom lines. The first one is that each of estimated k and K has a high and narrow peak around the zero relative error point with bounded relative errors. It assures us that, when only one set of data is used, the estimations are reliable even though some experimental error exists in the concentration. The second one is that the reaction rate constant can be estimated more reliably than the adsorption equilibrium constants, because the probability density function of the former has a higher peak around the zero relative error point, and this agrees with the result of the qualitative sensitive analysis carried out before.

2. About the Object Function

In the kinetic parameter estimation, the role of the object function is to make the simulated concentrations as close to the exact concentrations (no experimental concentration error) as possible. If the exact concentrations are available (which is not the case that we can expect), the ideal object function should be

$$J=\sum_{m=1}^{p}[c_{exact}(m)-c_{simulated}(m)]^2 \tag{13}$$

where $c_{exact}(m)$ is the ideal exact concentrations while $c_{simulated}(m)$ is the simulated concentrations, and m denotes m-th instance with time. Apparently, the function's minimum is zero, which means there is no difference between the simulated and the exact concentrations when the function is minimized, and in this case the estimated parameters should be accurate, as shown in the above numerical test. But, in practice, the experimental errors always exist and are unknown, which makes the exact concentrations unavailable; therefore, the object function given by Eq. (2) is commonly used instead

But, this substitution (replace) causes the simulated concentration to deviate lightly from the exact one. To show it clearly, the object function is rewritten here as:

$$J=\sum_{m=1}^{p}[c_{exact}(m)+\varepsilon(m)-c_{simulated}(m)]^2 \tag{14}$$

in which the experimental concentration is decomposed into two components: the one is the exact concentration and the other is the experimental error $\varepsilon(m)$. If the simulated concentrations and the exact concentrations are the same, the object function defined in Eq. (14) will have the value of $\sum_{m=1}^{p}\varepsilon(m)^2$. Because we cannot find this situation in a real application, it is assumed in the estimation algorithm that the simulated concentration is the same as the exact one when the object function defined in Eq. (14) is minimized. In fact, it is not what we can get in a real application.

Let's consider the following inequality:

$$\sum_{m=1}^{p}[c_{exact}(m)+\varepsilon(m)-c_{simulated}(m)]^2\le \sum_{m=1}^{p}[c_{exact}(m)-c_{simulated}(m)]^2+\sum_{m=1}^{p}\varepsilon(m)^2 \tag{15}$$

It indicates that the minimum of object function defined in Eq. (14) should be smaller than $\sum_{m=1}^{p}\varepsilon(m)^2$, where $c_{simulated}(m)$ are not equal to $c_{exact}(m)$. Of course, the estimated parameters are also influenced by the errors in the experimental concentrations.

Here, we can consider the "overshoot" of the object function, which can be defined as the value by which the minimum of the object function defined in Eq. (14) is smaller than $\sum_{m=1}^{p}\varepsilon(m)^2$. Because the overshoot of the object function is generally small, the error in the estimated parameters is also small and tolerable. On the other hand, the "overshoot" is not caused by the optimization algorithm itself, but by the fact that no detailed information about the experimental error is available, which forces us to use the object function defined in Eq. (14) instead of the one defined in Eq. (13). From the information point of view, it is reasonable that we cannot recover the kinetic parameters exactly when some information about the reaction is lost in the experiment.

3. Further Verification

To show further the reliability of the hybrid optimization method for kinetic parameter estimation, here the fitness of the estimated concentrations (calculated based on the estimated kinetic parameters) to the given exact concentrations is tested, where the initial values used to calculate the estimated concentrations are different from those used to estimate kinetic parameters. Let's consider the following two component system:

$$\frac{dc_A}{dt}=-\frac{k_AK_Ac_A}{1+K_Ac_A+K_Bc_B}$$
$$\frac{dc_B}{dt}=-\frac{k_BK_Bc_B-k_AK_Ac_A}{1+K_Ac_A+K_Bc_B} \tag{16}$$

where subscripts A and B denote components A and B, respectively. Given the kinetic parameters as $k_A=100$, $k_B=40$, $K_A=0.04$ and $K_B=0.01$, the exact concentration profile with different initial values of components A and B can be easily calculated as lines in Fig. 7, where three sets of initial concentrations are chosen for the test. In Fig. 7(a), the initial concentrations are $c_A(0)=180$ and $c_B(0)=180$, in Fig. 7(b) $c_A(0)=120$, $c_B(0)=120$, while in Fig. 7(c) $c_A(0)=60$, $c_B(0)=60$. To estimate the kinetic parameters, the exact concentration profiles are sampled evenly from the time range of 0 to 20 seconds with

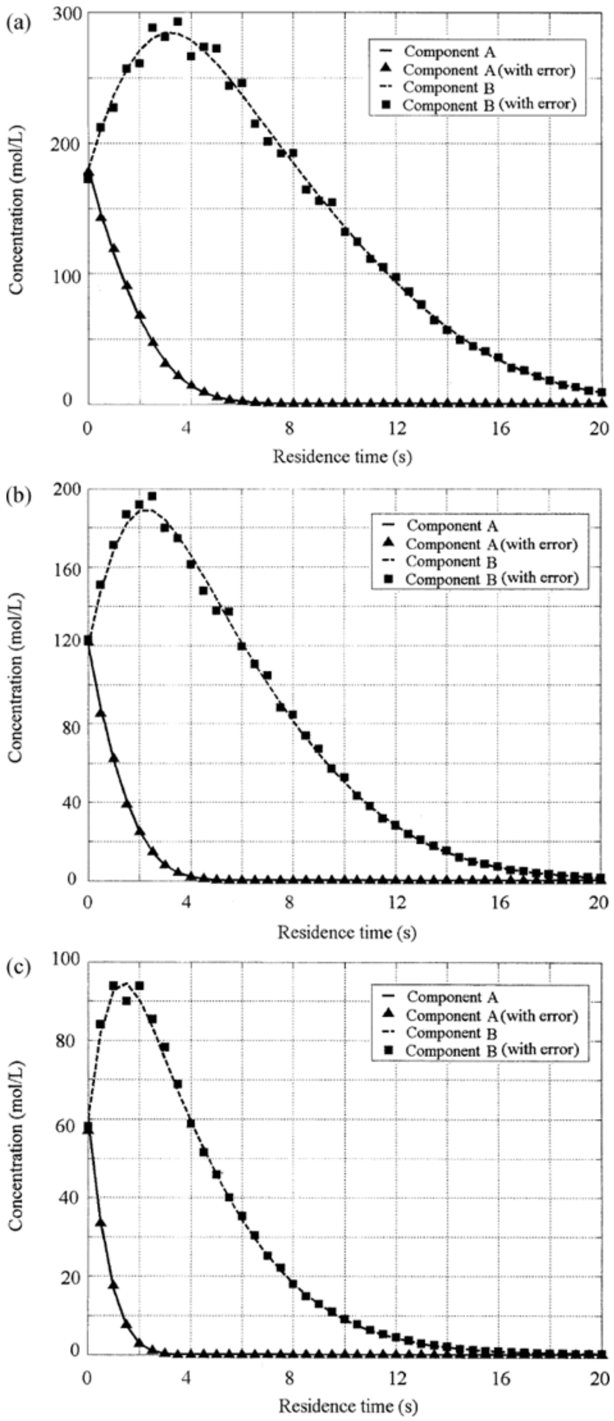

Fig. 7. An artificial experiment of a two-component system. (Exact concentrations are denoted by lines, while sampled concentrations with random error are denoted by discrete markers).
(a) Initial concentrations of $c_A(0)=180$ and $c_B(0)=180$ (mol/L).
(b) Initial concentrations of $c_A(0)=120$ and $c_B(0)=120$ (mol/L).
(c) Initial concentrations of $c_A(0)=60$ and $c_B(0)=60$ (mol/L).

an interval of 0.5 second. A bounded random relative error, whose maximum will not exceed 5% of the concentrations, is superposed on the exact concentration to simulate the experiment error. For comparison, concentration samples are also sketched in the figure as discrete markers.

Choosing the concentration samples in Fig. 7(a) as input, the hybrid optimization method estimates the four kinetic parameters as

k_A=94.64923483415211, k_B=42.39506915461477
K_A=0.03972790629826880, K_B=0.008864935373366311

If the concentration samples in Fig. 7(b) are chosen as input, the estimated parameters will be:

k_A=111.2879861550922, k_B=38.65915807659286
K_A=0.03671313198303160, K_B=0.001036427370199021

While the concentration samples in Fig. 7(c) are used, the estimated result will be:

k_A=98.99980578291010, k_B=37.01748896537582
K_A=0.04292663036394419, K_B=0.001120936475488305

It is clear that all three sets of the estimated parameters are close to the exact parameter values with a very small error, which indicates that the hybrid optimization method is independent on the initial concentrations.

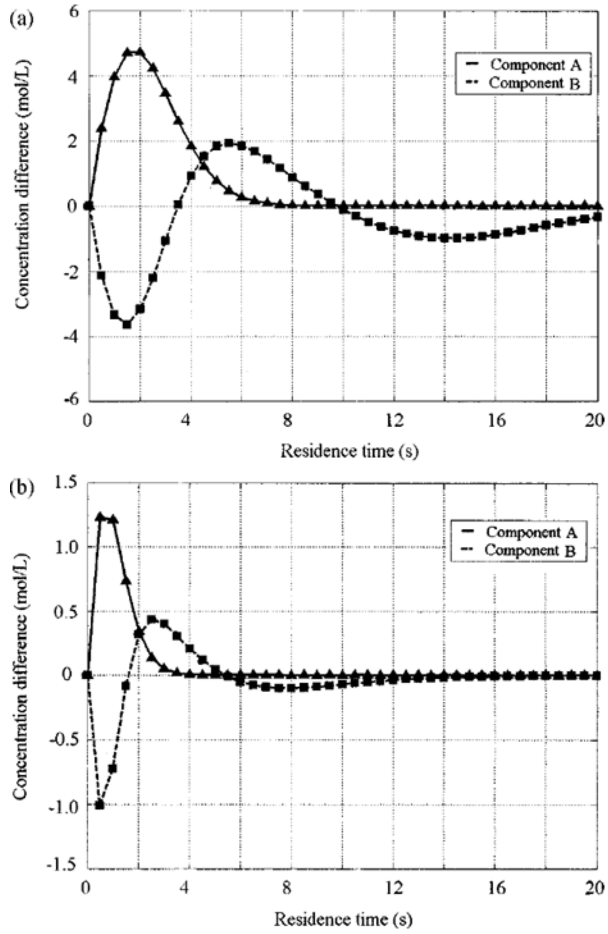

Fig. 8. (a) The concentration difference between the exact one [given in Fig. 7(a) and (c)] and the currently calculated one.
(a) Initial concentrations of $c_A(0)$=180 and $c_B(0)$=180 (mol/L).
(b) Initial concentrations of $c_A(0)$=60 and $c_B(0)$=60 (mol/L).

To check the reliability of the hybrid optimization method in another way, one of the three sets of estimated parameters mentioned above is chosen to calculate the concentration profile with different initial values. Here, the estimated parameters related to the initial concentrations $c_A(0)$=120, $c_B(0)$=120 are chosen, but the concentration profiles are calculated with the initial values $c_A(0)$=180, $c_B(0)$= 180 and $c_A(0)$=60, $c_B(0)$=60. The goodness of the approximation to the exact concentrations is sketched in Fig. 8, where the goodness is denoted by a concentration difference between the exact one (given in Fig. 7(a) and (c)) and the currently calculated one. The concentration differences shown in Fig. 8(a) are calculated with $c_A(0)$=180, $c_B(0)$=180, while in Fig. 8(b) initial values are $c_A(0)$=60, $c_B(0)$=60.

When the initial concentrations are $c_A(0)$=180, $c_B(0)$=180 in this figure, the maximal difference is only about 5; when $c_A(0)$=60, $c_B(0)$=60, the maximal difference is only about 1.2. The maximal relative errors in both cases are less than 3%, which is less than the relative errors introduced into the concentration samples. It is clear in Fig. 8 that the concentration differences between the exact and the estimated value are very small with both sets of initial concentrations. The small concentration difference indicates that the estimated kinetic parameter is a good approximation to the real exact value. Considering the fact that the initial concentrations used for comparing concentration difference are different from that used to estimate kinetic parameters, it can be concluded that the hybrid optimization method has the capability to estimate kinetic parameters reliably with only one set of concentration samples, provided that the mathematical model is a reasonable description of a focused reaction. Of course, when many sets of data are available, the estimation can be improved by some related statistical method.

CONCLUSION

For a photocatalytic reactions, either single component or multicomponent systems, it has been customary to estimate kinetic parameters separately based on single component data, even when the L-H equations are coupled where more than one reactants are involved. In this study, for a multi-component photocatalytic system, it has been illustrated that simultaneous parameter estimation is feasible in solving the relevant multi-objective optimization problem by a hybrid genetic-simplex method. In this method the genetic algorithm is used to find roughly an optimum in a rather wide range, while the simplex algorithm is used sequentially to refine the rough optimum and make it accurate. Applying it to a real reaction, we found the estimated results by the hybrid optimization method are more accurate than those of existing investigations. By sensitivity analysis and numerical verification, it has been also shown that the estimation is reliable even when only one set of experimental data (with unknown error) is available. This investigation proposes an effective approach to abstract kinetic parameters from very few experimental data.

ACKNOWLEDGEMENT

This work was supported by Korea Research Foundation Grants

(KRF-1999-005-E00025).

NOMENCLATURE

c_i : concentration of component *i* [mol/L]
$c_{exact}(m)$: the ideal (without error) concentration at the *m*th instance [mol/L]
$c_i(m)$: the *m*th observation of concentration of the component *i* [mol/L]
$c_{simulated}(m)$: the simulated concentration at the *m*th instance [mol/L]
$\hat{c}_i(m)$: the *m*th simulated concentration of component *i* [mol/L]
e : the relative error
F_j : the fitness of *j*th chromosome
J : object function
J_i : object function of component *i*
k : collection of all reaction rate constant
k_i : reaction rate constant of component *i* [mol/(L)(s)]
k_i^j : the *j*th possible value of k_i
K : collection of all adsorption equilibrium constants
K_i : adsorption equilibrium constant of component *i* [L/mol]
K_i^j : the *j*th possible value of K_i
*l*k : lower limit of the reaction rate constant
*l*K : lower limit of the adsorption equilibrium constant
P_j : possibility of the *j*th chromosome
r_i : reaction rate of component *i* [mol/(L)(s)]
uk : upper limit of the reaction rate constant
UK : upper limit of the adsorption equilibrium constant
$\mathbf{V}_i$: the *i*th chromosome
w_i : average weight of component *i*
x : one of estimated reaction rate constants or adsorption equilibrium constants
$\bar{x}$: average of x

Greek Letter

$\varepsilon(m)$: random error at the *m*th instance

REFERENCES

Alberici, R. M. and Jardim, W. F., "Photocatalytic Destruction of VOCs in the Gas-Phase Using Titanium Dioxide," *Applied Catlysis B: Environment*, **14**, 55 (1997).

Anpo, M., Matsuoka, M., Hano, K., Mishima, H., Ono, T. and Yamashita, H., "Photocatalytic Decomposition of N_2O on Cu+/Y-Zeolite Catalysts Prepared by Ion-Exchange," *Korean J. Chem. Eng.*, **14**, 498 (1997).

Balland, L., Estel, E., Cosmao, J. M. and Mouhab, N., "A Genetic Algorithm with Decimal Coding for Estimation of Kinetic and Energetic Parameter," *Chemometrics and Laboratory Intelligent Systems*, **50**, 121 (2000).

Belohlav, Z., Zamostny, P., Kluson, P. and Volf, J., "Application of Random-Search Algorithm for Regression Analysis of Catalytic Hydrogenations," *Canadian Journal of Chemical Engineering*, **75**, 735 (1997).

Biegler, L. T., Damiano, J. J. and Blau, G. E., "Nonlinear Parameter Estimation: A Case Study Comparison," *AIChE Journal*, **32**, 29 (1986).

Chai, Y. S., Lee, J. C. and Kim, B. W., "Photocatalytic Disinfection of *E. coli* in a Suspended TiO_2/UV Reactor," *Korean J. Chem. Eng.*, **17**, 633 (2000).

Cheney, W. and Kincaid, D., "Numerical Mathematics and Computing," Cole Publish Co., New York (1985).

Farza, M., Hammouri, H., Othman, S. and Busawon, K., "Nonlinear Observers for Parameter Estimation in Bioprocesses," *Chemical Engineering Science*, **52**, 4251 (1997).

Fox, M. A. and Dulay, M. T., "Heterogeneous Photocatalysis," *Chemical Review*, **93**, 341 (1993).

Froment, G. F., "The Kinetics of Complex Catalytic Reactions," *Chemical Engineering Science*, **42**, 1073 (1987).

Gen, M. and Cheng, R. W., "Genetic Algorithms and Engineering Design," Wiley, New York (1997).

Goldberg, D. E., "Genetic Algorithms in Search, Optimization and Machine Learning," Addison-Welsly, New York (1989).

Herrmann, J. M., "Heterogeneous Photocatalysis: Fundamentals and Application to the Removal of Various Types of Aqueous Pollutants," *Catalysis Today*, **53**, 115 (1999).

Kim, I. W., Liebman, M. J. and Edgar, T. F., "Robust Error-in-Variables Estimation using Nonlinear Programming Techniques," *AIChE Journal*, **36**, 985 (1990).

Levenspiel, O., "Chemical Reaction Engineering," Wiley, New York (1972).

Mehrab, M., William, A. A., Murray, M. Y. and Park, M. R., "Non-linear Parameter Estimation for a Dynamic Model in Photocatalytic Reaction Engineering," *Chemical Engineering Science*, **55**, 4885 (2000).

Michalewicz, Z., "Genetic Algorithm+Data Structure=Evolution Programs (second edition)," Springer-Verlag, New York (1994).

Oh, S. C., Sohn, S. H., Yeo, Y. K. and Chang, K. S., "A Study on the Prediction of Ozone Formation in Air Pollution," *Korean J. Chem. Eng.*, **16**, 144 (1999).

Ollis, D. F., "Photocatalytic Purification and Remediation of Contaminated Air and Water," *C. R. Acad. Sci. Paris, Se´rie IIc, Chimie/Chemistry*, **3**, 405 (2000).

Park, T. Y. and Froment, G. F., "A Hybrid Genetic Algorithm for the Estimation of Parameters in Detailed Kinetic Models," *Computers in Chemical Engineering*, **22**, 103 (1998).

Press, W. H., Flnnery, B. P., Teukolsky, S, A. and Vetterling, W. T., "Numerical Recipes," Cambridge University Press, New York (1986).

Sirisuk, A., HillJr, C. G. and Anderson, M. A., "Photocatalytic Degradation of Ethylene over Thin Films of Titania Supported on Glass Rings," *Catalysis Today*, **54**, 159 (1999).

Turchi, C. S. and Rabago, R., "Bench scale Testing of Photocatalytic Oxidation (PCO) to Destroy Volatile Organic Compound (VOC) Emissions," *SEMATECH Technology Transfer Report #95042791A-ENG* (1995).

Turchi, C. S., Rabago, R. and Jassal, A., "Destruction of Volatile Organic Compound (VOC) Emissions by Photocatalytic Oxidation (PCO): Field Test Results," *SEMATECH Technology Transfer Report #96083084A-ENG* (1996).

Winston, W. L., "Introduction to Mathematical Programming: Application & Algorithm," PWS-Kent, Boston (1991).

Wolf, D. and Moros, R., "Estimation of Rate Constants of Heterogeneous Catalytic Reactions without Supposition of Rate Determining

Surface Steps - An Application of a Genetic Algorithm," *Chemical Engineering Science*, **52**, 1189 (1997).

Wolfrum, E. J., Rabago, R. and Jassal, A., "Destruction of Volatile Organic Compound (VOC) Emissions by Photocatalytic Oxidation (PCO): Final Reports," SEMATECH Technology Transfer Report #97013236A-ENG (1997).