

Simultaneous Determination of Reaction Order and Rate Constant of Consecutive Reactions by Union Optimization Algorithm of NGA-TS

Cao Tongcheng*, Zhu Zhongliang, Huang Xiaofeng, Li Tonghua

Contribution from: Department of Chemistry, Tongji University, Shanghai, China, 200092

Received: August 2, 2004

Accepted (in revised form): June 25, 2005

Abstract

Numeric genetic algorithm (NGA) and Tabu search (TS) are both efficient optimization algorithms. They have advantages and shortcomings. A new optimization algorithm, NGA-TS, is proposed to enhance the efficiency of the search algorithm and avoid entrapment into local minima. Target testing factor Analysis (TTFA), combined with NGA-TS, is applied to resolve the two-way kinetic-spectral data obtained in the process of spectroscopic reactions, to obtain a clean spectrum of each component and calculate the kinetic parameters including rate constants and reaction order. By numeric integration the kinetic spectrum of each component can be calculated. The method can also be applied to resolve the reaction system with arbitrary reaction order. With rate constants and reaction order optimized by NGA and NGA-TS respectively, the simulated data matrix for a two-step consecutive reaction has been successfully resolved and it is found that NGA-TS is more efficient than NGA. The data matrices measured for the alkaline hydrolysis of diethyl phthalate and the electrodegradation process of sunset yellow have been solved.

Keywords: TTFA, two-way kinetic-spectral data, reaction order, rate constant, NGA-TS

Résumé

L'algorithme numérique génétique (NGA) et la recherche Tabu (TS) sont tous les deux des algorithmes d'optimisation efficaces qui possèdent leurs avantages et leurs limitations respectives. Nous proposons un nouvel algorithme d'optimisation, NGA-TS, pour améliorer l'efficacité de l'algorithme de recherche et éviter d'être piégé dans des optimisations locales. Nous avons

*Author to whom correspondence should be addressed: ctc@tongji.edu.cn

appliqué l'analyse de facteur du test cible (TTFA), en combinaison avec NGA-TS, pour résoudre les données spectrales double direction obtenues dans le processus de réactions spectroscopiques et acquérir le spectre pur de chaque composante ainsi que les paramètres cinétiques incluant la constante de vitesse et l'ordre de réaction. En utilisant l'intégration numérique pour calculer le spectre cinétique de chaque composante, cette méthode peut être appliquée pour résoudre le système réactionnel avec un ordre de réaction arbitraire. Avec des constantes de vitesse et des ordres de réaction optimisés par NGA et NGA-TS respectivement, nous avons pu résoudre avec succès la matrice de données simulées pour une réaction à deux étapes, trouvé que NGA-TS était plus efficace que NGA et pu obtenir les constantes de vitesse et les ordres de réaction avec exactitude. Les matrices de données mesurées pour l'hydrolyse alcaline du phthalate de diéthyle et le processus d'électrodegradation du jaune orangé ont pu être résolues pour l'obtention de résultats fiables.

Introduction

The on-line spectra measured from a chemical reaction can form a two-way data matrix which contains both the reaction kinetic information and the pure spectrum of each component. The kinetic information such as the reaction orders and rate constants can be evaluated through analysis of the two-way spectroscopic data by means of some chemometrics algorithms when the reaction order is known. Most of the work is directed towards consecutive first-order reaction. Methods such as non-linear least-squares fit (NLLS) (1,2), time shift-trilinear models (3,4), rank annihilation factor analysis (RAFA) (5,6), target testing (TT) (7,8), direct exponential curve resolution algorithm (DECRA) (9), curve resolution (10) and so forth, are commonly applied. But for the reaction whose order is unknown, the determination of

the kinetic process is very difficult and seldom reported. This problem can be solved by using the RAFA method but it can only be used to estimate the first-step reaction order and the rate constant (5,11,12).

In some complex reactions, the kinetic model may be unknown and the order of reaction may not be an integer. Therefore, it is more difficult to solve the two-way data of such reaction by means of the former methods. In this work, numeric integration is used to calculate the kinetic spectra of each species of such a reaction, and TTFA combined with optimization algorithm is used to resolve the reaction order and rate constants of the reaction simultaneously. For a two-step consecutive reaction, there are two reaction orders and two rate constants. Therefore, four parameters need to be optimized. This is an intricate combinatorial optimization problem. The aim of combinatorial optimization is to obtain the best solution from the available solution sets of combinatorial optimization. Some algorithms for combinatorial optimization include simulated annealing (SA)(13), genetic algorithms (GA) (14,15) and tabu search (TS) (16,17), etc. Here, an optimization algorithm, numeric genetic algorithm (NGA) (18) combined with tabu search, is proposed to solve the reaction orders and rate constants of a consecutive reaction.

Since the first presentation by Holland in 1975, GA has attracted a lot of interest. It tries to simulate the evolutionary process of a living species to solve optimization problems. It was first proposed by Holland in 1975 (14), and subsequently improved and developed by De Jong (19) and Davis (20). Several tutorials about GA have been published in journals devoted to different research fields. As examples we cite those by Lucasius and Kateman (21-24), Hibbert (25), Shaffer and Small (26), Wehrens and Buydens (27) and Luke (28). In chemistry, GA has been applied in many research works such as molecular modeling and molecular docking (29,30), regression problem, QSAR, PLS, etc.(31-36). In basic GA, the codes are binary digits. In practical applications, most problems are based on real numbers. The operation is time consuming and consequently its application is limited to some degree. Hibbert estimates the kinetic parameters by using a real number coding and a hybrid GA (37). NGA, in which the real number is used as a chromosome, modifies the basic GA and achieves better results. NGA is efficient for the optimization of consecutive variables and is especially suitable for non-linear problems. Tabu search (TS) is a powerful optimization procedure that has been successfully applied to a number of combinatorial optimization problems (38-42). It has also been applied

to molecular docking and structure-based drug design. It has the ability to avoid entrapment in local optima by employing a flexible memory system.

NGA and TS are two approximate search algorithms which have good commonality and optimization capability and are suitable for the resolution of the complex and large-scale optimization programs. In this work, the NGA-TS algorithm combines the advantage of the NGA and TS. It can improve the capability of searching the global optimization and shorten the search time. A set of synthesis data has been resolved in order to evaluate the performance of this approach. Experimental data from the alkaline hydrolysis of dimethyl phthalate and the electrodegradation of sunset yellow solution are also solved successfully.

Theory

The realization of NGA-TS

Numeric Genetic Algorithm (NGA)

Generally, in NGA, an individual can be represented as a vector or an element in the vector and they are called chromosomes. The evolution of chromosomes is realized by genetic operators, in which there are three processes, namely, crossover, mutation operators and initialization population.

NGA uses vectors of floating-point numbers to represent the individuals in a population. Assume that the number of individuals in the population is N , and individual $X_i \in R^n$, $i=1 \dots N$, the individual X_i is composed of m chromosome, $X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,m}\}$. Basic procedures are as follows:

(1) Initialization population

Parameter N represents population size. Initial vector populations are chosen randomly and should cover the entire parameter space, if nothing else is known about the system.

$$x_{i,j} = x_j(\text{lo}) + \text{random}[0,1] \times (x_j(\text{hi}) - x_j(\text{lo}))$$

$$i = 1, 2, \dots, N, j = 1, 2, \dots, m$$

(2) Crossover

Two individuals X_{r1} and X_{r2} are selected at random, and then the next generation individual V is defined as $v_j = (x_{r1,j} + x_{r2,j})/2$, where $j = \text{random}[1..m]$, and the rest of the chromosomes of V , $v_k = x_{r1,k}$, $k \neq j$, were the same ones as X_{r1} .

(3) Mutation

Two individuals X_{r1} and X_{r2} are selected at random, and then the next generation individual V , $j = \text{random}[1..m]$ is defined, $v_j = x_{r1,j} + \text{random}[-1,1] \times \text{Scale} \times x_{r1,j}$, where

Scale is a parameter, which decreases stepwise from 1.0 to 0.00001, and the rest of the chromosomes of V , $v_k = x_{r1}$, $k, k \neq j$, were the same ones as X_{r1} .

Tabu Search

Tabu Search is a convenient approach for solving combinatorial optimization problems by using a guided, local search procedure to explore the entire solution space without becoming easily trapped in local optima. TS begins by determining an initial solution y . Additional solutions (termed neighbors) are generated by modifying the existing solutions through a sequence of moves. The best new neighbor (y') is used as the starting point for the next iteration unless it is on a Tabu list. Thus, even if no neighbor solutions are better than the initial solution, one of these is still chosen as the starting point for the next iteration. A record of the best solution ever found (y^*) is separately maintained. In addition, the Tabu lists provide an adaptive memory that guides the search by taking advantage of historical information. This memory enables TS to make strategic choices and achieve responsive exploration.

At each iteration, certain moves are classified as Tabu (forbidden) and added to Tabu lists. At the same time, the Tabu property of other solutions will expire and can be removed from the Tabu lists. In this way, Tabu lists are updated continuously and adapt to the current state of the search. Both recency-based and frequency-based management of Tabu lists have been employed. A recency-based Tabu list prevents cycling by immediately putting new solutions on the Tabu list and maintaining them there for a given time period (Tabu tenure), which is typically a number of iterations equal to two times the number of independent variables. This type of list is often called short-term memory. Long-term memory is achieved through frequency-based Tabu lists, which record solutions that have been found most often. Recency-based and frequency-based Tabu lists are not a complete record of the previous solutions—they only record part of the historical information and are updated throughout the search process. Thus, TS makes decisions based on up-to-date knowledge. The use of adaptive memory enables TS to exhibit learning and creates a more flexible and effective search (43).

The union of NGA and Tabu search

The major disadvantage of NGA is its weak climbing ability, that is, when NGA reaches the local optima, it is difficult to skip over the local range and as a result, it becomes stagnant. The main reasons for this phenomenon

are that the probability of mutation for GA is too small, its colony is too small, and consequently the colony loses variety. As a result, the solution obtained is usually the suboptimal solution. If the probability of mutation for GA is increased, the computational time will be too long. In addition, the operation speed of traditional GA is slow, and it must be modified in large scale search.

Compared to GA, TS has a high operation speed. Additionally, TS has a relative dependence on the initial solution. The better the initial solution, the more accurate the Tabu search solution is. Therefore, some algorithms were employed to overcome its shortcomings. For example, heuristic algorithms can bring a better initial solution. Furthermore, TS is a single-single operation, that is, there is only one initial solution in the operation and it only moves a solution to another solution in every generation, while in GA several parallel solutions can be operated in the same time in every generation.

Comparing between GA and TS, the two algorithms can be combined by using the search capability of GA and climbing capability of TS, which can unite the advantages of both algorithms while avoid their disadvantages. The improvement can enhance its optimization ability sharply.

In this paper, a NGA-TS hybrid algorithm is built by combining the NGA and TS and applying it to solve the optimization problem.

The general structure of the NGA-TS adopted in this paper is described as follows:

Step 1: Initialization: an initial population is constructed of different sized solutions.

Step 2: NGA Procedure: scale = 1.0 is set, and then the NGA operations are started followed by Crossover, Mutation and Memory.

Step 3: TS Procedure: when the optimal solution is not changed in the defined generations, i.e. NGA reaches to the local optima, the algorithm turns to the TS step. In TS, a TS scale is defined, and it can be regulated automatically to ensure a solution is as good as possible. When the TS scale reaches a very low value (such as 0.0001), the algorithm returns to the NGA and the best chromosome instead of the newer of the TS is used. Then, the NGA continues to search and the NGA scale decrease.

Step 4: NGA and TS run alternatively.

Step 5: End: the search ends until the NGA scale reaches an extremely small value or the sum of generations reaches at given times.

TTFA

Assume that spectra collected during a kinetic process can be arranged against measurement time and form a data matrix, \mathbf{Y} . The size of the matrix is $nt \times nw$, where nt denotes the number of measurement time and nw is the number of the corresponding wavelengths. The data matrix \mathbf{Y} can be decomposed to the pure absorption spectral matrix \mathbf{S} and the kinetic spectra matrix \mathbf{Q}^T :

$$\mathbf{Y} = \sum_{i=1}^{nc} s_i \mathbf{q}_i^T + \mathbf{E}_1 = \mathbf{S} \mathbf{Q}^T + \mathbf{E}_1 \quad (1)$$

where nc is the number of the absorbing species, s_i and \mathbf{q}_i^T represent the pure absorption spectrum and kinetic spectrum of the i th component, respectively. Superscript T denotes the transpose of a matrix or a vector. \mathbf{E}_1 is the error matrix.

By using Principal Component Analysis (PCA) (44), the data matrix \mathbf{Y} can be decomposed to two orthogonal matrices, a score matrix \mathbf{T} and a loading matrix \mathbf{P}^T , that is:

$$\mathbf{Y} = \mathbf{T} \mathbf{P}^T + \mathbf{E}_2 \quad (2)$$

The relation between \mathbf{T} and \mathbf{S} or \mathbf{P}^T and \mathbf{Q}^T is as follows:

$$\mathbf{S} = \mathbf{T} \mathbf{R}^{-1} \quad (3)$$

$$\mathbf{Q}^T = \mathbf{R} \mathbf{P}^T \quad (4)$$

where $\mathbf{R}(nc \times nc)$ is a target transformation matrix and it can be computed from the present target (the kinetic spectrum or the absorption spectrum of each component). When the pure absorption spectrum of each component is unknown, since the chemical reaction is in accord with a fixed kinetic model, the testing target $\mathbf{Q}_{\text{test}}^T (nc \times nt)$ can be formed by the kinetic spectrum of each component which is computed according to the supposed kinetic equation. Then, based on the principle of Least-Squares Regression, the target transformation matrix \mathbf{R} can be computed according to:

$$\mathbf{R} = \mathbf{Q}_{\text{test}}^T \mathbf{P} \quad (5)$$

Through the target transformation matrix \mathbf{R} , the pure spectral matrix $\mathbf{S}_{\text{test}} (nw \times nc)$ can be computed according to Eq. (3). So the residual matrix \mathbf{Y}_{res} of the original data matrix can be obtained according to:

$$\mathbf{Y}_{\text{res}} = \mathbf{Y} - \mathbf{S}_{\text{test}} \mathbf{Q}_{\text{test}}^T \quad (6)$$

When a series of different parameter of the supposed kinetic equation is chosen, a different target transformation matrix $\mathbf{Q}_{\text{test}}^T$ can be calculated. A different residual matrix \mathbf{Y}_{res} will be obtained.

Here the Sum of Square (SSQ) of the residual matrix \mathbf{Y}_{res} is regarded as the objective function. When it reaches the minimum, the supposed pure spectra \mathbf{S}_{test} and kinetic

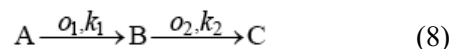
spectra $\mathbf{Q}_{\text{test}}^T$ are considered to be equal to the actual pure spectra and kinetic spectra. The kinetic model of the reaction, including the reaction order and rate constants, is determined. To compare this with the noise of measurement conveniently, SSQ is always transformed to Residual Standard Deviation (RSD), according to:

$$\text{RSD} = \sqrt{\frac{\text{SSQ}}{nw(nc)}} \quad (7)$$

The key step of this method is the target's optimization by TT, which is called TTFA.

The computation of the kinetic spectra of two-steps consecutive reaction

Suppose that the following two-steps consecutive reaction is considered:



where A, B and C represent the reactant, the intermediate and the final product, respectively, and o_1 , o_2 and k_1 , k_2 are reaction orders, rate constants for these two-step reactions.

In this work, NGA-TS is used to optimize o_1 , o_2 and k_1 , k_2 simultaneously. Then, the target testing matrix $\mathbf{Q}_{\text{test}}^T$ can be calculated according to the four optimized parameter. The RSD of the residual matrix \mathbf{Y}_{res} obtained by TTFA is regarded as the fitness function of the NGA-TS method.

The key step of the process is the calculation of the supposed kinetic spectra $\mathbf{Q}_{\text{test}}^T$ according to different kinetic model. Suppose the initial concentration of the reactant A is set to unity, while that of the intermediate B and final product C are zero. The relative concentration for each species A, B and C is represented by q_j ($j=A, B, C$). The rate equation of the mentioned reaction can be written as:

$$r_A = -\frac{dq_A}{dt} = k_1 q_A^{o_1} \quad (9)$$

$$r_B = \frac{dq_B}{dt} = k_1 q_A^{o_1} - k_2 q_B^{o_2} \quad (10)$$

The relative concentration for species A (q_A) can be computed by integrating the Eq. (9)

$$q_{A,j} = \exp\left\{\frac{\ln[(o_1 - 1)k_1 t_j + 1]}{1 - o_1}\right\} \quad (o_1 \neq 1) \quad (11-a)$$

$$q_{A,j} = \exp(-k_1 t_j) \quad (o_1 \neq 1) \quad (11-b)$$

where j stand for the j th reaction time. However, it is difficult to integrating Eq. (10). Here, the numeric integration is used to obtain q_B .

Suppose $q_{B,j}$ represents the relative concentration of B at the j th reaction time and $q_{B,1}$, i.e. the relative concentration of B at the beginning of the reaction, is equal 0. When $j=2, \dots, nt$, the rate of the $(j-1)$ th reaction time can be calculated according to Eq. (10):

$$r_{B,j-1} = k_1 q_{A,j-1} - k_2 q_{B,j-1}^2 \quad (12)$$

$q_{B,j}$ can be calculated approximately as following:

$$q_{B,j} = q_{B,j-1} + r_{B,j-1}(t_j - t_{j-1}) \quad (13)$$

To improve the accuracy of the result, the average reaction rate between the $(j-1)$ th and j th reaction time can be calculated based on $q_{B,j}$, i.e. the result of Eq. (13), and used to compute a new $q_{B,j}$, according to:

$$\overline{r_{B,j}} = (r_{B,j-1} + r_{A,j} - k_2 q_{B,j}^2) / 2 \quad (14)$$

$$q_{B,j} = q_{B,j-1} + \overline{r_{B,j}}(t_j - t_{j-1}) \quad (15)$$

It is necessary to take into consideration that Δt , the interval of the $(j-1)$ th and j th reaction time, should be divided into some smaller intervals if Δt is long. The smaller interval should be used to compute new $q_{B,j}$ to ensure the accuracy of $q_{B,j}$.

$q_{C,j}$, i.e. the relative concentration of C at the j th reaction time, can be obtained as the following:

$$q_{C,j} = 1 - q_{A,j} - q_{B,j} \quad (16)$$

Finally, the testing target matrix $\mathbf{Q}_{\text{test}}^T$ can be formed by the vectors of the relative concentration for each species A, B and C, i.e. q_A , q_B and q_C .

Experimental

Synthetic data

To evaluate the performance of the method, two sets of simulated two-way data were created based upon Eq. (8). The absorption spectra of each species were produced

by Lorenz function based on the method mentioned in the literature (5). The maximum absorption wavelength λ_{max} , the maximum absorbance A_{max} and the half peak width b are listed in Table 1. Considering the experimental error, noise of Gaussian distribution was added to the synthetic data. The standard deviation of absorbance was determined by converting from the transmittance,

Table 1. Parameters used for synthetic spectrum.

Component	A_{max}	$\lambda_{\text{max}}/\text{nm}$	b
A	1.00	250	0.05
B	0.95	260	0.06
C	0.90	270	0.07

Table 2. Kinetic and measuring parameters used for synthetic data.

Parameter	Value
Reaction order for 1st step	1.00
Reaction order for 2nd step	0.60
Rate constants k_1	0.05500
Rate constants k_2	0.01200
Range of wavelength/nm	240-300
Interval of wavelength/nm	1
Number of wavelength	61
Range of reaction time/s	0-600
Number of reaction time	20
Standard deviation S_T	0.001

S_T (6). The kinetic and measured parameters are listed in Table 2. By using the parameters listed in Table 1 and 2, the original data matrices, \mathbf{Y}_s , with dimension (61×20) can be obtained.

Experimental data from alkaline hydrolysis of diethyl phthalate

A 1.50 mL aliquot of 0.44 mol/L NaOH (aq) solution was transferred into a 1 cm quartz cell. A mini stir bar, whose speed was set to 500 rpm, was also put into the quartz cell. Then, 1.50 mL of 9.44×10^{-4} mol/L of diethyl phthalate (DEP), that had been kept in a double-layer thermostatic beaker at $25.0 \pm 0.1^\circ\text{C}$, was rapidly added and the hydrolysis reaction started. The concentration of NaOH was 0.22 mol/L and that of DEP was 4.72×10^{-4} mol/L in the mixed solution. Simultaneously, the kinetic software of Agilent 8453 UV-VIS spectrophotometer was started to measure the absorption of the sample at the chosen time. The temperature during the reaction was $25.0 \pm 0.1^\circ\text{C}$, which was held constant by a Peltier

constant temperature equipment. Consider during lower wavelength, the CO_2 could cause interference by being absorbed into the NaOH solution, and during high wavelength, each species would have no absorption. The wavelength range could be chosen and was set to 255-300 nm (46 data points, interval of 1 nm) and the data matrix \mathbf{Y}_H with size of 46×33 was formed. Due to the fact that the spectral features changed more at the very beginning of the reaction, the interval of sampling time was set to 10 seconds during the first 120 seconds and then the cycle time was increased 5% each time until the end of the reaction, which was 2000 seconds.

Experimental data from electrodegradation of sunset yellow

100 mL of 0.100 g/L sunset yellow solution (pH=11.0, adjust with NaOH solution) was added to the electrolytic cell (double-layer thermostatic beaker). The concentration of the supporting electrolyte, Na_2SO_4 , was 10.0 g/L. A reticular $\text{Ti}/\text{SnO}_2\text{-Sb}_2\text{O}_3\text{-PtO}$ electrode was used as an anode, while two sheets of pure titanium served as the cathode. The current density was 20 mA/cm². Electrolysis was started and a stopwatch was used to time the reaction. Nine samples were taken from the electrolyzing solution during 40 min at 5 min intervals. Each time 2.00 mL of electrolyte was diluted into 10.00 mL of purified water. The reference solution can be prepared by treating the solution similarly without sunset yellow. UV-spectra of the samples were acquired in a 1 cm quartz cell using Agilent 8453 UV-VIS spectrometer. The wavelength range was 240 - 550 nm and the interval was 5 nm. The data matrix, \mathbf{Y}_E , was formed with a dimension of 63×9 .

Results and Discussion

Synthetic data

The reliability of this algorithm can be discriminated by resolving the synthetic data. By using TFA combined with NGA and NGA-TS method, respectively, the synthetic data matrix \mathbf{Y}_s can be resolved and five sets of kinetic parameters and RSD_{\min} can be obtained as shown in Table 3. From the data listed in Table 3, it can be seen that the kinetic parameters solved by using NGA-TS correspond well to given values, which are listed in Table 2. However, sometimes the parameters solved by using NGA do not correspond. Their solved RSD_{\min} are also not the same. This indicates that at each time the results obtained by using NGA may converge at a different value. It can be concluded that the results

of NGA may be entrapped in a local optima while that of NGA-TS can converge at the global optima, i.e. the correct kinetic parameter, at each time.

From the resolution of the synthesis data, it can be seen that NGA-TS is a better optimization algorithm and can overcome the shortcoming of NGA, namely, being easily entrapped in a local optima. It is feasible to determine the reaction orders and rate constants of complex reaction by using TFA combined with NGA-TS.

A contour map of global search results are shown in Fig.1. It should be noted that there is a large region that the value of objective function was nearly unchanged when o_1 is between 1.2 and 2.0 and o_2 is approximately close to 0.6. This is the local optima and the search results are sometimes entrapped in it, as shown in Table 3.

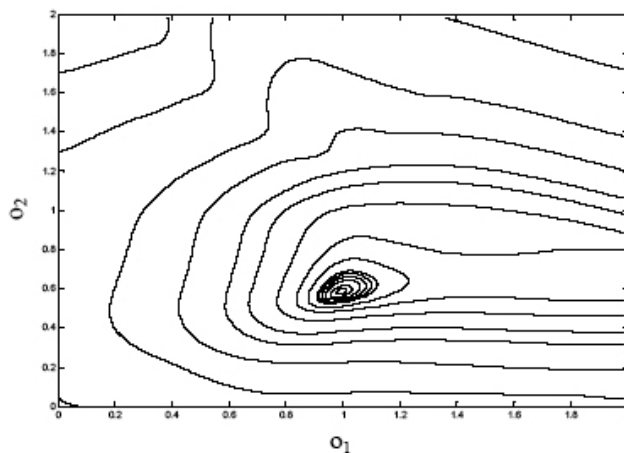


Figure 1. Contour curves of global search for synthetic data.

Resolution to the data from the alkaline hydrolysis of dimethyl phthalate

Fig. 2 illustrates the process of alkaline hydrolysis of dimethyl phthalate. PCA was operated on the data matrix \mathbf{Y}_H . The result is listed in Table 4. After extracting three principal components, the residual standard deviation (RSD) of the residual matrix is 0.00014, and meets the noise of the instrument. According to the principle of the PCA, the existence of three absorptive species is proved in the system, namely, the dimethyl phthalate, the intermediate methyl phthalate, and the final hydrolysis product phthalate.

TFA was operated combined with NGA-TS on the data matrix \mathbf{Y}_H . The results are listed in Table 5. The kinetic parameter results are as follows: $o_1=0.973$; $o_2=0.960$; $k_1=0.0229$; $k_2=0.000938$. The RSD of the residual matrix is equal to 0.000447 when the resolution is

Table 3. The solved results of synthetic data by NGA and NGA-TS.

Computational		NGA			
Time	σ_1	σ_2	k_1	k_2	RSD_{\min}
1	1.82	0.637	0.0322	0.0159	0.00285
2	0.997	0.597	0.0551	0.0120	0.00261
3	1.46	0.661	0.0419	0.0147	0.00290
4	1.88	0.668	0.0304	0.0166	0.00286
5	0.777	0.666	0.0501	0.0116	0.00412
Computation		NGA-TS			
Time	σ_1	σ_2	k_1	k_2	RSD_{\min}
1	0.996	0.596	0.0551	0.0120	0.00261
2	0.997	0.597	0.0551	0.0120	0.00261
3	0.996	0.596	0.0551	0.0120	0.00261
4	0.996	0.596	0.0551	0.0120	0.00261
5	0.996	0.596	0.0551	0.0120	0.00261

Table 4. The PCA results for data matrix measured in the alkaline hydrolysis of DEP.

N	EV	EV_n/EV_{n+1}	RSD
1	20.338	12.79	0.03212
2	1.59	3.05	0.01008
3	0.52073	42.06	0.00025
4	0.01238	15.93	0.00004
5	0.000777	1.19	0.00004
6	0.000655		0.00004

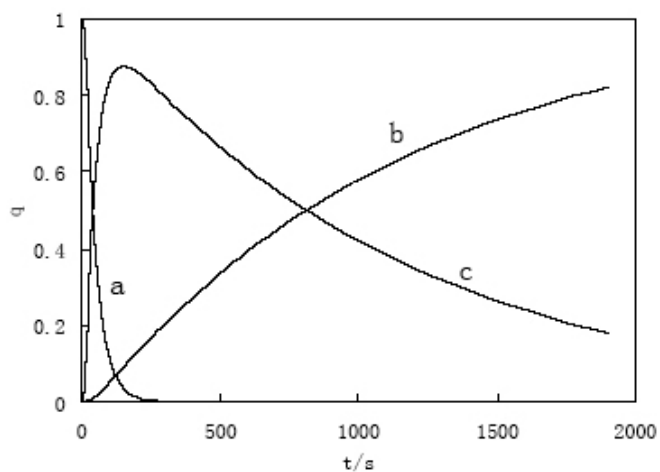


Figure 2. Absorption spectra of DEP in alkaline hydrolysis process.

Table 5. The solved kinetic parameters for alkaline hydrolysis of DEP.

Parameter	Value
σ_1	0.973
σ_2	0.960
k_1	0.0229
k_2	0.000938
RSD_{\min}	0.000447

optimized. This indicates that the hydrolysis of dimethyl phthalate in excess NaOH solution complies with the first-order consecutive reaction model and the unit of the rate constants are in s^{-1} . The solved absorption spectra of each species are illustrated in fig. 3.

Resolution to the data from the electrodegradation of sunset yellow solution

Fig. 4 illustrates the process of the electrodegradation of sunset yellow solution. Operate PCA on the data matrix Y_E . The results are listed in Table 6. According to the principle of the PCA, the existence of two absorptive species is proved in the system. Since the final product of phenol degradation is CO_2 , which is not absorbed in the measured wavelength range, it can be assumed that there exists one absorptive intermediate during the degradation. It should be noted that RSD of this experiment is equal to 0.0012 and much bigger than that of alkaline hydrolysis of dimethyl phthalate, i.e. 0.00014. This is mostly caused by the different ways of measurement used in the two experiments. This experiment adopts the off-line measurement while another experiment adopts on-line measurement, which has less error resulted by several factors, such as errors in sampling, variance of different quartz cell and so on.

TTFA was operated combined with NGA-TS on the data matrix Y_E . The results are listed in Table 7. The kinetic parameter results are as follows: $\alpha_1 = 0.781$; $\alpha_2 = 0.767$; $k_1 = 0.0946$; $k_2 = 0.0450$. The RSD of the residual

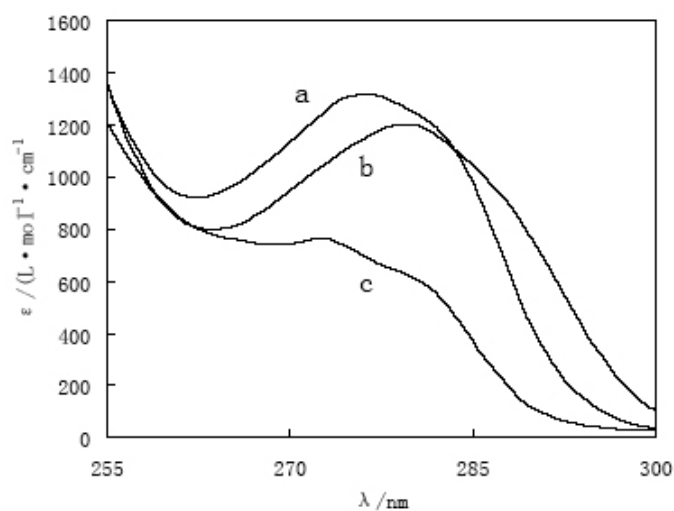


Figure 3. Kinetic and absorption spectrum of each component in the alkaline hydrolysis of DEP. a) DEP; b) monoethyl phthalate; c) phthalate.

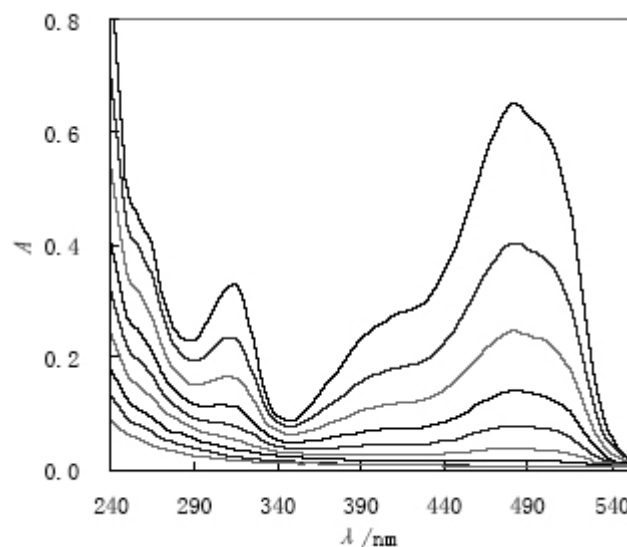


Figure 4. Absorption spectra of sunset yellow in the electrodegradation process.

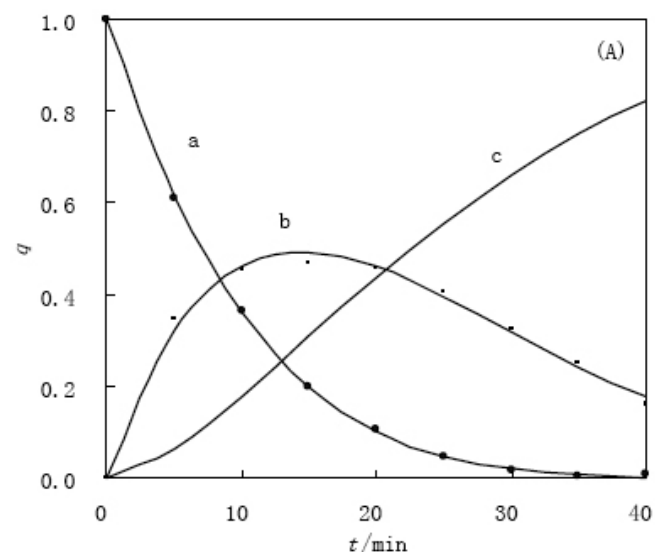


Figure 5(A). Kinetic spectrum of each component in electrodegradation of sunset yellow: a) sunset yellow; b) intermediate; c) end product.

matrix is equal to 0.00265 when the resolution is optimized. The solved kinetic spectra of each species are shown in Fig. 5(A). The solved absorption spectra of sunset yellow solution and the apparent intermediate are shown in Fig.5(B).

Table 6. The PCA results for data matrix measured in the electrodegradation of sunset yellow.

N	EV	EV_n/EV_{n+1}	RSD
1	4.147	6.00	0.0255
2	0.6914	25.80	0.0012
3	0.0268	2.35	0.0007
4	0.0114	1.18	0.0005
5	0.0097		0.0003

Table 7. The solved kinetic parameters for electrodegradation of sunset yellow.

Parameter	Value
α_1	0.781
α_2	0.767
k_1	0.0946
k_2	0.0450
RSD_{\min}	0.00265

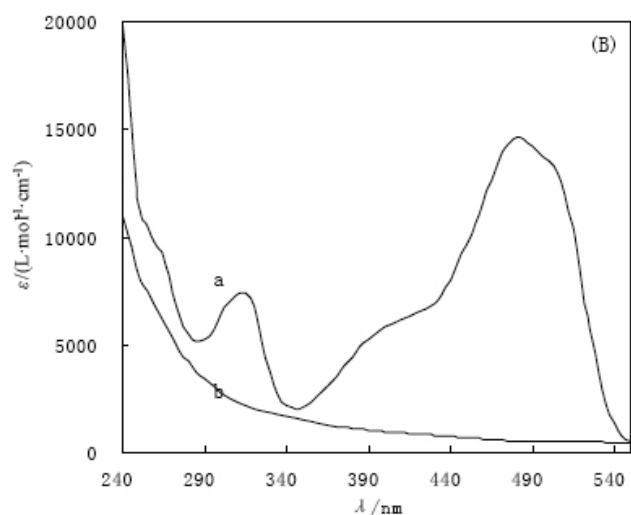


Figure 5(B). Absorption spectrum of each component in electrodegradation of sunset yellow: a) sunset yellow; b) intermediate; c) end product.

Conclusion

TTFA, combined with NGA-TS, can be applied to resolve the two-way kinetic-spectral data measured from spectroscopic reactions and acquire some information of the reaction system such as rate constants, reaction orders, and pure absorption spectra of each species and so on, when the pure absorption spectra of each species and the kinetic model of the reaction are unknown. The numeric integration is applied to compute the kinetic

spectra of each species by using the solved rate constants and reaction order. This method can be applied to solve the reaction system which has arbitrary reaction order, including a non-integer reaction order. By using this method, a simulated data matrix for a two-step consecutive reaction with an unknown kinetic model has been successfully determined and two actual reaction systems were also solved to obtain a reliable result. Obviously, TTFA-NGA-TS can also be applied to determine other reaction models by using the proper method to compute the kinetic spectra.

Here, a new combinational optimization algorithm, NGA-TS, is applied to obtain the global best point quickly. There is sharply improvement on skipping local optima comparing to NGA. Through the results of the synthesis system, it can be concluded that the NGA-TS is an efficient search algorithm and it can reach the best global point with stability.

Acknowledgements

The authors are grateful to the National Natural Science Foundation of China for financial support (Project 20275026).

References

1. R.X. Cai., X.G. Wu, Z.H. Liu, *Analyst*, **124**, 751 (1999).
2. A. Juan, M. Maeder, M. Martı́nez, *Chemometr. Intell. Lab.*, **541**, 25 (2000).
3. S. Bijlsma, D.J. Louwerse, W. Windig *et al. Anal. Chim. Acta*, **376**, 339 (1998).
4. S. Bijlsma, D.J. Louwerse., A.K. Smilde, *J. Chemometr.*, **13**, 311 (1999).
5. Z.L. Zhu, L. Yu, W.F. Song, T.H. Li, *Chem. J. Chin. Univ.*, **22**, 734 (2001).
6. Z.L. Zhu, J. Xia, J. Zhang, T.H. Li, *Anal. Chim. Acta*, **454**, 21 (2002).
7. E. Furusjo, L.G. Danieleson, *Anal. Chim. Acta*, **373**, 83 (1998).

8. E. Furusjo, L.G. Danieleson, *Chemometr. Intell. Lab.*, **50**, 63 (2000).
9. W. Windig, B. Antalek, *Chemometr. Intell. Lab.*, **37**, 241 (1997).
10. S. Bijlsma, D.J. Louwerse, W. Windig, A.K. Smilde, *Anal. Chim. Acta*, **376**, 339 (1998).
11. Z.L. Zhu, W.Z. Cheng, Y. Zhao, *Chemometr. Intell. Lab.*, **64**, 157 (2002).
12. Z.L. Zhu, W. Li, J. Xia, *Anal. Chim. Acta*, **527**, 203 (2002).
13. S. Kirkpatrick, C.D. Gelatt., M.P. Vecchi, *Science*, **220**, 551 (1983).
14. J.H. Holland, "Adaptation in Natural and Artificial Systems", MIT Press, 1975.
15. D.E. Goldberg, "Genetic Algorithm in Search Optimization and Machine Learning", Addison-Wesley Publishing, 1989.
16. F. Glover, *ORSA J on Comp.*, **1**, 190 (1989).
17. F. Glover, *ORSA J on Comp.*, **29**, 4 (1990).
18. P.S. Cong, T.H. Li, *Anal. Chim. Acta*, **293**, 191 (1994).
19. D. Jong, An Analysis of the Behavior of a Class of Genetic Adaptive Systems, Ph.D. Thesis, University of Michigan, Ann Arbor, MI., 1975.
20. L. Davis, Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York, NY, 1991.
21. C.B. Lucasius, G. Kateman, *Chemometr. Intell. Lab.*, **19**, 1 (1993).
22. C.B. Lucasius, G. Kateman, *Chemometr. Intell. Lab.*, **25**, 99 (1994).
23. C.B. Lucasius, G. Kateman, *Comput. Chem.*, **18**, 127 (1994).
24. C.B. Lucasius, G. Kateman, *Comput. Chem.*, **18**, 137 (1994).
25. D.B. Hibbert. *Chemometr. Intell. Lab.*, **19**, 277 (1993).
26. R.E. Shaffer, G.W. Small, *Anal. Chem.*, **69**, 236 (1997).
27. R. Wehrens, L. Buydens, *Trends Anal. Chem.*, **17**, 193 (1998).
28. B.T. Luke, J. Devillers, Eds., "An overview of genetic methods. In Genetic Algorithms in Molecular Modeling", Academic Press, New York, 1996, p.35.
29. J. Devillers, Eds., "Genetic algorithms in computer-aided molecular design. In Genetic Algorithms in Molecular Modeling", Academic Press, New York, 1996, p.1.
30. A. Sundaram, V. J. Venkatasubramanian, *Chem. Info. Comput. Sci.*, **38**, 1177 (1998).
31. T.H. Li, M. He, P.S. Cong, *Chemometr. Intell. Lab.*, **45**, 177 (1999).
32. T.J. Hou, J.M. Wang, X.J. Xu. *Chemometr. Intell. Lab.*, **45**, 303 (1999).
33. K. Barry, D. Charles, K. Robert, *et al.*, *Chemometr. Intell. Lab.*, **66**, 51 (2003).
34. D. Frank, K. Birgit, G. Guenter, *Chemometr. Intell. Lab.*, **65**, 67 (2003).
35. K.L. Tang, T.H. Li., *Chemometr. Intell. Lab.*, **64**, 55 (2002).
36. R. Leardi, *J. Chemometr.*, **15**, 559 (2001).
37. D.B. Hibbert, *Chemometr. Intell. Lab.*, **19**, 319 (1993).
38. J.A. Bland, G.P. Dawson, *Tabu Search Des. Optim.*, **23**, 195 (1991).
39. F. Glover, *Ann. Oper. Res.*, **41**, 3 (1993).
40. F. Glover, *Manag. Decis. Econ.*, **11**, 365 (1990).
41. F. Glover, H.J. Greenberg, *Eur. J. Oper. Res.*, **39**, 119 (1989).
42. F. Glover, *Comput. Oper. Res.*, **13**, 533 (1986).
43. B. Lin, D.C. Miller, *Comput. Chem. Eng.*, **28**, 1451 (2004).
44. E.R. Malinowski, D.G. Howery, "Factor Analysis in chemistry", Wiley-interscience, New York, 1980, p. 72.