Talanta 93 (2012) 72–78

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Artificial neural networks combined with experimental design: A “soft” approach for chemical kinetics

Filippo Amato[a], José Luis González-Hernández[b], Josef Havel[a,c,d,∗]

[a] Department of Chemistry, Faculty of Science, Masaryk University, Kampus Bohunice, Kamenice 5/A14, 625 00 Brno, Czech Republic
[b] Department of Physical Chemistry, University of Salamanca, 37008 Salamanca, Spain
[c] R&D Center for Low-cost Plasma and Nanotechnology Surface Modifications, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic
[d] Department of Physical Electronics, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

ARTICLE INFO

Article history:
Received 12 October 2011
Received in revised form 12 January 2012
Accepted 19 January 2012
Available online 28 January 2012

Keywords:
Chemical kinetics
Soft-modelling
Artificial Neural networks
Experimental design
Rate constants
Multicomponent analysis
Optimization

ABSTRACT

The possibilities of artificial neural networks (ANNs) “*soft*” computing to evaluate chemical kinetic data have been studied. In the first stage, a set of “*standard*” kinetic curves with known parameters (rate constants and/or concentrations of the reactants), which is some kind of “normalized maps”, is prepared. The database should be built according to a suitable experimental design (ED). In the second stage, such data set is then used for ANNs “learning”. Afterwards, in the second stage, experimental data are evaluated and parameters of “other” kinetic curves are computed without solving anymore the system of differential equations.

The combined ED-ANNs approach has been applied to solve several kinetic systems. It was also demonstrated that using ANNs, the optimization of complex chemical systems can be achieved even not knowing or determining the values of the rate constants. Moreover, the solution of differential equations is here not necessary, as well. Using ED the number of experiments can be reduced substantially. Methodology of ED-ANNs applied to multicomponent analysis shows advantages over classical methods while the knowledge of kinetic reactions is not needed. ANNs computation in kinetics is robust as shown evaluating the effect of experimental errors and it is of general applicability.

1. Introduction

The fundamental task in chemical kinetics is the determination of (i) the number and kind of chemical species, (ii) the reaction mechanism, (iii) rate constant values and (iv) concentrations of the species as a function of time. Such results form what is usually called “*hard*” kinetic model of the system. Once the mechanism is known, the rate of the reaction can be expressed by differential equations and the values of rate constants are obtained by integration.

In the last 50 years several computer programs, based on general least-squares method, have been developed to manage and treat different types of chemical data. Pioneering contribution in chemical equilibria was given by Sillén with his family of LETAGROP [1] programs. These programs and many others developed since, such as modern HYPERQUAD [2] or OPIUM (http://web.natur.cuni.cz/~kyvala/opium.html) for example, allow to compute just stability constants. First computer program for kinetics was LETAGROP KINET developed by Sillén [3], later on modified [4], and many other programs based on general least squares method for kinetics were developed since (SPECFIT [5], KILET [6]), KINAGDC [7], etc.).

Nevertheless, computation in kinetics is faced with several difficulties. The behaviour of chemical kinetic systems is often highly nonlinear and the determination of kinetic model may be failing as consequence of undistinguishability and/or non-unique identifiability [8]. Even in the case in which the mechanism is known in detail, the analytic solution of differential equations can be mathematically impossible [9]. Usually several experimental simplifications are introduced and kinetic rate constant values are estimated by numerical methods [9].

Hard approach seems to be a not efficient way to obtain information of strictly practical use in problems of applied chemistry. Another possibility is the so called “*soft-modelling*” or “*model-free*” approach. The main advantage of “*soft*” methods is the possibility to model the kinetic behaviour of complex chemical systems without any *a priori* knowledge. Soft-modelling approach is often based on powerful mathematical tools known as artificial neural networks (ANNs).

It seems that the first evaluation of ANNs applicability in chemical kinetics was given by Pérez-Bendito and Silva [10]. Several

---

∗ Corresponding author at: Department of Chemistry, Faculty of Science, Masaryk University, Kampus Bohunice, Kamenice 5/A14, 625 00 Brno, Czech Republic. Tel.: +420 549 492 698.
E-mail address: havel@chemi.muni.cz (J. Havel).

0039-9140/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2012.01.044

authors applied ANNs to model real chemical systems, for example in order to predict: (i) the behaviour of industrial reactors [11], (ii) reaction rate values [12,13], (iii) rate constant values [14,15], (iv) analyte concentrations in multicomponent kinetic determinations [16–18] and (v) kinetics of drug release [19]. The results obtained [11–19] for specific systems show clearly the relevance of ANNs applications in this field. ANNs in combination with ED have also been extensively used in capillary electrophoresis where kinetic processes are also involved [20].

The aim of this work is to study applicability of ANNs in chemical kinetics, to generalize ANNs use and to evaluate their power, robustness but also limitations. In optimization involving kinetics to study the role of experimental design and to give general guidelines for ANN applications in experimental practice of chemical kinetics.

2. Theoretical aspects

2.1. General considerations

Chemical systems are generally characterized by t physico-chemical parameters, either microscopic as rate constants or macroscopic as concentrations. The whole set of parameters may be indicated by:

$$\{x_r\}_{r=1,2,\ldots t} \tag{1}$$

Each parameter x_r is related to the values of all the others:

$$x_s = f(x_j) \tag{2}$$

where x_s parameters are usually experimentally accessible. An example useful for our further discussion is given by the generalized Lambert–Beer–Bouguer's equation:

$$A = b\sum_{l=1}^{k} \varepsilon_i c_i = f(\varepsilon_1, \varepsilon_2, \ldots \varepsilon_k, c_1, c_2, \ldots c_k) \tag{3}$$

where ε_i and c_i are the molar absorptivity and concentration value of the i-th chemical species respectively.

Relationships among different sets of parameters can be too complex or some of them even unknown.

In chemical kinetics, concentration and absorbance values are usually accessible by experimental measurements. We can write:

$$c = f(k_{\pm 1}, k_{\pm 2}, \ldots k_{\pm j}, t) \tag{4}$$

$$A = g(c) = h(k_{\pm 1}, k_{\pm 2}, \ldots k_{\pm j}, t) \tag{5}$$

where $k_{\pm i}$ are kinetic rate constants and t is the time.

According to *hard-modelling* approach, the analytical expression for *f*, *g* and *h* should be found. However, another way is to approximate such functions using *soft-modelling*. The result is an approximated description of the behaviour of the system with respect to output parameter values. This is relevant for the purposes of applied chemistry. As a powerful *soft-modelling* tool, artificial neural networks are often used. They are able to accomplish function approximation at sufficient degree of accuracy [21].

2.2. Artificial neural networks

An artificial neural network is the transposition of the brain structure into a simplified formal architecture in which a series of units, called neurons, is organized into layers. The network is formed by linking each neuron of a layer to every neuron in the subsequent one as shown in Fig. 1.

The neural network uses sets of input and output parameter values usually written as row vectors. A couple of correspondent input and output row vectors is joined into a new vector called "*example curve*" or simply "*curve*". All the curves so obtained are grouped in a "*training matrix*". Data in each curve are values of x_j and x_s parameters (Eq. (2)) which can be either experimental or calculated.

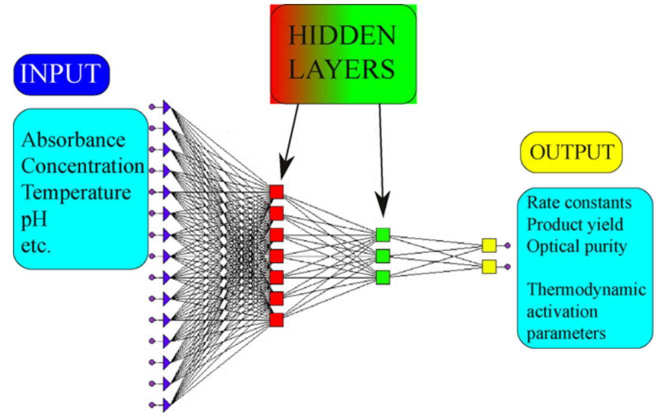

Fig. 1. A general structure of ANN architecture with inputs, outputs and 2 hidden layers for applications in kinetics.

We indicate x_j as *output* and x_s as *input* parameter respectively.

The first layer of the network contains neurons that receive input data values from the rows of the training matrix. This information is transmitted from the i-th neuron of a layer to the j-th neuron of the subsequent one after weighing with a weight w_{ij}. Layers following the input one are called "*hidden*". In each neuron of a hidden layer the weighed inputs coming from the previous one are summed each other and added to a bias. The result is then transformed by means of a suitable mathematical function to obtain an output called "*activation*" of the neuron. The activation is transferred to neurons in the next layer after another weighing step. In the last layer, output parameter values are estimated by means of a suitable transformation function.

The described process is called "*learning*" and it is repeated iteratively. After each epoch the estimated values (o^*_{ik}) of output parameters are compared with those (o_{ik}) of the corresponding curve in the training matrix and the value of RMS is calculated as:

$$\mathrm{RMS} = \sqrt{\frac{\sum_{i=1}^{m}\sum_{k=1}^{n}(o^*_{ik} - o_{ik})^2}{m \times n}} \tag{6}$$

where m and n are the number of rows of the training matrix and the number of output parameters respectively.

During learning, weight values are changed according to suitable algorithms in order to decrease the value of RMS. In our study the back-propagation algorithm was used. The learning is considered complete when the lowest values of RMS is reached.

For a given system studied, the network architecture must be optimized. While the number of input and output neurons is given by data used, the optimal number of hidden layers and that of their neurons is found using the criteria of the lowest RMS. For each hidden layer, a graph of RMS values vs. the number q of neurons (Fig. 2) shows that at first, increasing the value of q, the RMS decreases rapidly, but after, a poor improvement is obtained. The optimal number of neurons in that hidden layer is given by the point of intersection of the two branches of the graph. For a network with i neurons in the input layer, j in the hidden and k in the output ones, the architecture can be written in annotation as (i, j, k).

After training, the "*verification*" step must be performed using "new" examples. The cross-validation or leave-one-out method consists in the exclusion from the training step of one curve of the training matrix. The excluded curve is then used as

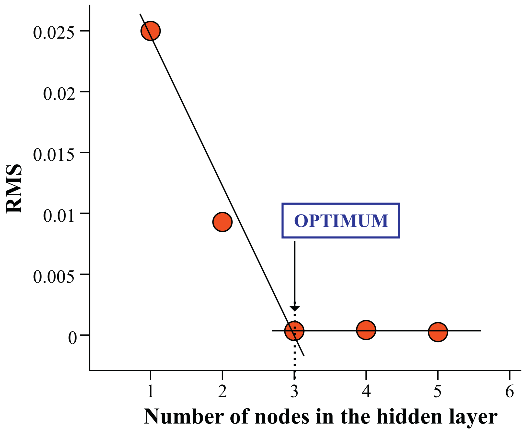

Fig. 2. Optimization of the number of hidden neurons in the first hidden layer.

verification example. Other curves may be supplied by new data, either experimental or calculated. Each curve corresponds to a point on the response surface which is approximated by ANNs. Better definition of the response surface is usually achieved by increasing the number of points. When expensive or time-consuming tasks must be performed to obtain experimental data, another approach should be used.

2.3. Experimental design

Experimental design has been extensively treated in textbooks about chemometrics [22,23]; for this, here we will give only some general remarks.

The area bounded by both lowest and highest values of the selected output parameters (x_j) is called *working space*.

Experimental design is a tool to select points within the working space according to a well-distribution in order to extract the maximum amount of information. The aim of the experimental design is to maximize the ratio:

$$\frac{\text{Information obtained}}{\text{Number of points}} \tag{7}$$

As stated in the Section 2.1, n parameters can be selected as outputs and the optimal n-dimensional experimental design is found. Here we indicate these parameters as "factors". The values assumed by the j-th factor in the experimental design are called "*levels*". The coordinates of the i-th point ($i = 1, 2,\ldots m$) of the experimental design are given by the vector:

$$(\, a_{i1} \quad a_{i2} \quad \ldots \quad a_{ij} \quad \ldots \quad a_{in} \,) \tag{8}$$

in which the element a_{ij} is the level of the x_j factor. According to Eq. (2), from a_{ij} values of x_j, b_{ik} ($k = 1, 2,\ldots p$) values of x_s are measured (or calculated). In this way the $\mathbf{M} = m \times (n + p)$ matrix is obtained:

$$\mathbf{M} = \begin{pmatrix} b_{11} & \ldots & b_{1k} & \ldots & b_{1p} & a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{i1} & \ldots & b_{ik} & \ldots & b_{ip} & a_{i1} & \ldots & a_{ij} & \ldots & a_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{m1} & \ldots & b_{mk} & \ldots & b_{mp} & a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{pmatrix} \tag{9}$$

Each row of the matrix **M** represents a "*curve*".

Random selection of other points within the working space allows obtaining new curves to be used for the verification step.

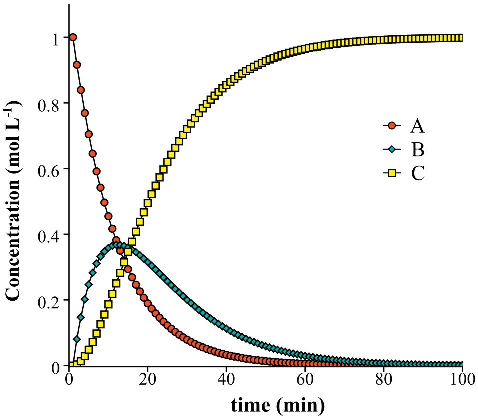

Fig. 3. Concentration profiles for *A*, *B* and *C*.

3. Computational aspects

3.1. Software

ANNs computation was performed using Trajan Neural Network Simulator, Release 3.0 D (Trajan Software Ltd. 1996–1998, UK). Some calculations were also done using STATISTICA V. 6 (StatSoft, Inc., USA). All computations were performed on a standard PC computer with Microsoft Windows Professional XP 2000 as operating system.

4. Results and discussion

The possibility of the use of ANNs in various kinetic systems will be examined while the following items are to be considered:

1. kinetic data modelling;
2. role of experimental design and data properties on modelling and estimation by ANNs.

4.1. Case 1: two consecutive reactions

Thermal decomposition of acetone:

$$C_3H_6O \rightarrow C_2H_2O + CH_4 \rightarrow \frac{1}{2}C_2H_4 + CO + CH_4 \tag{10}$$

is a common industrial process that takes place by two consecutive reactions. In the process some species are also formed as by-products (CH_4, CO). Referring to species that really undergo chemical transformation, the process can be simplified by the scheme:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \tag{11}$$

In Fig. 3 one of the concentration curve profile for the three species *A*, *B*, *C*, is shown. The scheme given by Eq. (11) is mathematically described in matrix notation [24] as:

$$\frac{d}{dt}\begin{pmatrix}[A]\\ [B]\\ [C]\end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{pmatrix}\begin{pmatrix}[A]\\ [B]\\ [C]\end{pmatrix} \tag{12}$$

When $[A]_0 = 1\,\text{mol}\,L^{-1}$ and $[B]_0 = [C]_0 = 0$, the system of differential equations obtained is given by:

$$\begin{aligned} \frac{d}{dt}[A] &= -k_1[A] \\ \frac{d}{dt}[B] &= k_1[A] - k_2[B] \\ \frac{d}{dt}[C] &= k_2[B] \end{aligned} \tag{13}$$

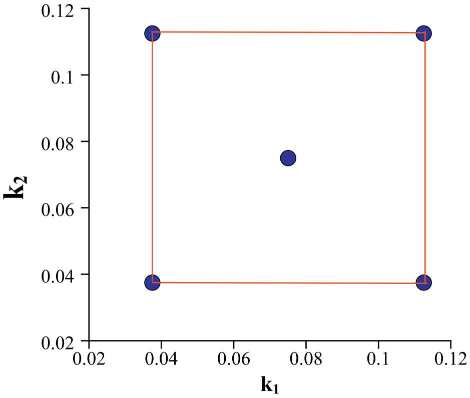

Fig. 4. A 2^2 experimental design with centre point.

and, by integration:

$$\begin{aligned}[A] &= e^{-k_1 t}\\ [B] &= \frac{k_1}{k_2-k_1}(e^{-k_1 t}-e^{-k_2 t})\\ [C] &= [A]_0-[A]-[B]\end{aligned} \qquad (14)$$

The kinetic constant's working space was within the interval $0.0375 \le k_j \le 0.1125$. As stated in the Section 2, the training matrix should be obtained using ED on the chosen working space. As first, 2^2 ED with centre point (Fig. 4) was applied. The coordinates of each point of the ED are two values of k_1 and k_2. Using these values the concentrations for each species were calculated by means of expressions in Eq. (14) and were used as input parameters, whence k_1 and k_2 values were used as outputs. In this way the rows of the training matrix **M** (Eq. (9)) are of the form:

$$\begin{pmatrix} [A]_{t_0} & [A]_{t_1} & \cdots & [A]_{t_f} & [B]_{t_0} & [B]_{t_1} & \cdots \\ [B]_{t_f} & [C]_{t_0} & [C]_{t_1} & \cdots & [C]_{t_f} & k_1 & k_2 \end{pmatrix} \qquad (15)$$

The optimal network architecture was found using the criteria of minimum value of RMS (Fig. 2) to minimize the risk of overtraining. Thus, the architecture obtained was (300,3,2). No performance improvements were obtained with architectures with two hidden layers. Under these conditions, the network training process was successful and acceptable results were obtained for both k_1 and k_2.

When a set of kinetic curves randomly chosen within the working space is presented to the network for the verification process no correlation was found between estimated and theoretical k_j values. It means that the network, with the chosen experimental design, is able to model the system that is to "approximate" the unknown relationship between inputs and outputs. However, this is not enough to reach acceptable estimation of k_j values for an unknown set of data. This is a consequence of a too "simple" experimental design. Evidently more points are needed to give to the network sufficient information for a better estimation. The optimal experimental design found is shown in Fig. 5. The external values $k_j = 0.03125$ and $k_j = 0.11875$ were also included.

With this experimental design it was found that the optimal network architecture was (300,8,2). Acceptable results were obtained both for modelling and for verification. The results for the k_1 verification carried out on a set of unknown kinetic curves are shown in Fig. 6 and analogous results were obtained for k_2.

4.1.1. *Sensitivity analysis*

The proposed ED-ANNs approach allows the acceptable estimation of k_1 and k_2 values if the concentration profiles of all the species are used as input data. Sometimes some of the concentration values might not be experimentally accessible. For this reason, sensitivity analysis was performed to follow the effect of parameters on the performance of the model.

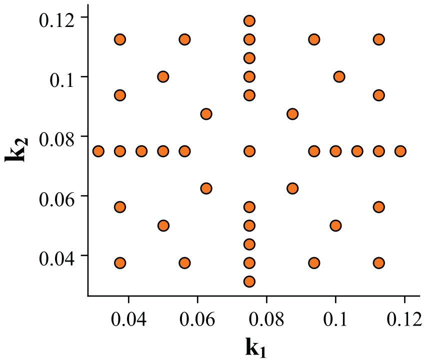

Fig. 5. A 42-points experimental design.

1. Concentration values of the species A only were found not sufficient to build an acceptable model.
2. The concentration profile for the species C is affected by the well known slow-fast ambiguity that consists in the practical undistinguishability of k_1 and k_2 values for time courses data [6]. In agreement with this, we have found that using as input only the concentration values for the species C, the predicted values of k_1 and k_2 for a given curve were equal to the average value of k_1 and k_2 for that curve in the training matrix.
3. Using as input the time courses data only for the intermediate species B, the values of both k_1 and k_2 were estimated with the same accuracy as by using the concentration profiles for all the species.

4.2. *Case 2: cyclic reaction pathways*

As an example of simple cyclic model for the evaluation of ANNs modelling ability, the following one was chosen:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} A \qquad (16)$$

Such kinetic paths are rather common in industrial chemical processes, especially when catalytic steps are involved. The proposed path does not take into account the formation of secondary products.

The kinetic constant's working space was within the interval $0.05 \le k_j \le 0.1$; moreover, the external values $k_j = 0.0375$ and $k_j = 0.1125$ were also included in the experimental design. The ED

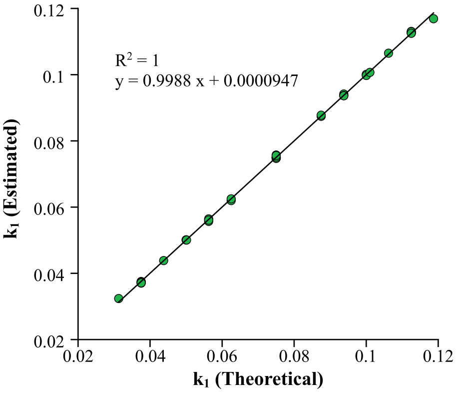

Fig. 6. Case 1: agreement between theoretical and calculated k_1 values for the verification set.

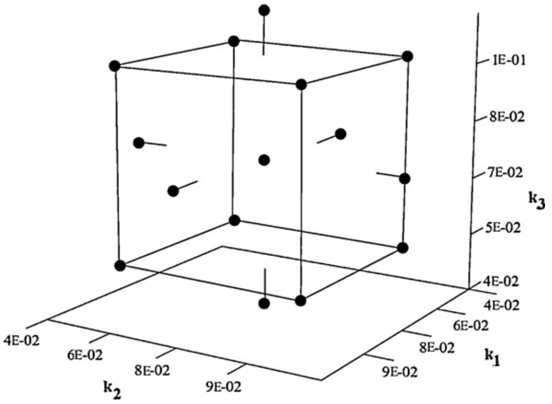

Fig. 7. Central composite experimental design for three factors.

used was a central composite one, as it is shown in Fig. 7. In this case the rows of the training matrix **M** (Eq. (9)) are of the form:

$$([A]_{t_0}[A]_{t_1}\ldots[A]_{t_f}[B]_{t_0}[B]_{t_1}\ldots[B]_{t_f}[C]_{t_0}[C]_{t_1}\ldots[C]_{t_f}k_1k_2k_3) \qquad (17)$$

Using the same criteria as described above, the optimal neural network architecture was found as (150,7,3). The linear regression fit gives the following results:

$$k_{1,(\mathrm{est})} = 0.985 \cdot k_{1,(\mathrm{theor.})} + 1.004 \times 10^{-3}, \quad R = 0.998 \qquad (18)$$

$$k_{2,(\mathrm{est})} = 0.990 \cdot k_{2,(\mathrm{theor.})} + 7.040 \times 10^{-4}, \quad R = 0.999 \qquad (19)$$

$$k_{3,(\mathrm{est})} = 0.999 \cdot k_{3,(\mathrm{theor.})} + 8.867 \times 10^{-4}, \quad R = 0.999 \qquad (20)$$

where the $k_{(\mathrm{est})}$ means *estimate* and $k_{(\mathrm{theor.})}$ means *theoretical*. The results indicate clearly the ability of the chosen network to model the proposed kinetic curves and to determine the values of kinetic parameters. In order to verify the prediction ability of the trained network, cross verification was performed. The agreement found is expressed by the following regression line equation:

$$k_{1,(\mathrm{est})} = 0.987 \cdot k_{1,(\mathrm{theor.})} + 0.0001295, \quad R = 0.998 \qquad (21)$$

For k_2 and k_3, similar agreement was found.

4.3. Case 3: multicomponent kinetic analysis

The use of ANNs as a tool to perform spectrophotometric multicomponent analysis will be evaluated. The system previously described and evaluated using partial least squares method (PLS) [25] will be studied in order to compare the two approaches.

Let us consider the mixture of two analytes M_1 and M_2 that are reacting with an excess of reagent R to form P_1 and P_2 products; the reaction scheme is given in Eq. (22).

$$\begin{aligned} R + M_1 &\xrightarrow{k_1} P_1 \\ R + M_2 &\xrightarrow{k_2} P_2 \end{aligned} \qquad (22)$$

where $k_1 = 1\ \mathrm{mol^{-1}\ L\ min^{-1}}$ and $k_2 = 0.025\ \mathrm{mol^{-1}\ L\ min^{-1}}$. The absorbance is given by:

$$A_T = \varepsilon_R C_R + \sum_i [M_i]_0 \left(\varepsilon_{M_i} e^{-k_{M_i} C_R t} + \varepsilon_{P_i}[1 - e^{-k_{P_i} C_R t}]\right) \qquad (23)$$

For simplicity we consider M_1 and P_2 as the only absorbing species, with $\varepsilon_{\mathrm{M}_1} = \varepsilon_{\mathrm{P}_2} = 100\ \ \mathrm{m^2\ mol^{-1}}$ and then, Eq. (23) can be reduced to:

$$A_T = 100[M_1]_0(e^{-t} + (1 - e^{-0.25t})) + 100[M_2]_0(e^{-t} + (1 - e^{-0.25t})) \qquad (24)$$

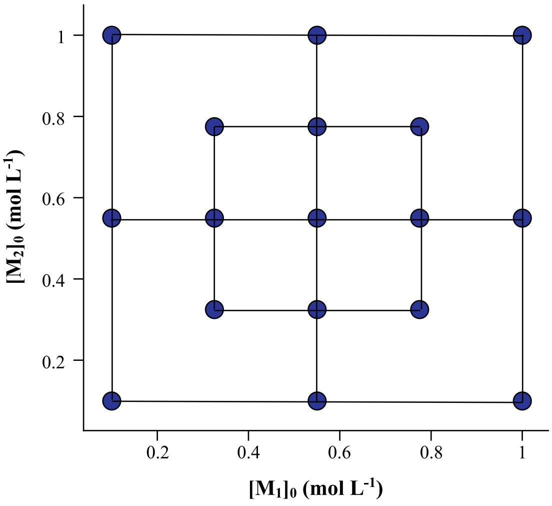

Fig. 8. Case 3: a 5-levels experimental design.

Absorbance values were calculated at time intervals of 0.2 min for a total time of 10 min. The optimized experimental design chosen was formed by 17 points (Fig. 8). The selected range for the initial concentration of analytes was 0.1–1 mM.

The absorbance values (A_T) were used as inputs and initial concentrations of M_1 and M_2 as outputs. The rows of the training matrix **M** (Eq. (9)) are of the form:

$$(\,A_{T,t_0} \quad A_{T,t_1} \quad \ldots A_{T,t_f} \quad [M_1]_0 \quad [M_0]\,) \qquad (25)$$

The optimal neural network architecture found was (51,4,2). The training process gave acceptable results (RMS ≤ 10^{-3}). In order to verify the prediction power of ANNs, cross-validation was performed. The results obtained are reported in Table 1. Moreover, a set of ten new curves was used as external test set; the following regression equation expresses the agreement obtained for theoretical an estimated $[M_1]_0$ values:

$$[M_1]_{0,(\mathrm{est})} = 1.011[M_1]_{0,(\mathrm{theor.})} + 7.9 \times 10^{-3} \qquad (26)$$

Acceptable results were also obtained for $[M_2]_0$. The results show clearly the ability of ANNs to model the chosen chemical system and to estimate the initial concentrations of analytes with an average error of about 0.4%. ED-ANNs approach is in acceptable agreement with PLS treatment of data as previously reported [25].

The model provided by ANNs can be also used to estimate the concentration of only one analyte without any knowledge about the presence of other species or side interactions that may occur. In fact, carrying out the training procedure with the same network, ignoring $[M_2]_0$ values, we obtained acceptable agreement for $[M_1]_0$ values, as expressed by the following regression equation:

$$[M_1]_{0,(\mathrm{est})} = 0.998[M_1]_{0,(\mathrm{theor.})} + 1.24 \times 10^{-3} \qquad (27)$$

The estimation of $[M_1]_0$ values ignoring $[M_2]_0$ ones, was also carried out on the ten curves used as external test; the agreement obtained is shown in Fig. 9.

The results of ED-ANNs approach to multicomponent kinetic analysis indicate that it is possible to take into account the effect of interferences.

4.4. Case 4: optimization of reaction conditions

Here we consider a chemical process studied and nicely commented by Leardi [26] for which no hard model is available. In this case the optimal values for the temperature and for the time of

Table 1
Cross-validation results (multicomponent kinetic analysis).

Correct $[M_1]_0$	Found $[M_1]_0$	Residual	Correct $[M_2]_0$	Found $[M_2]_0$	Residual
0.550	0.5552	0.0520	1.000	0.9966	−0.0034
0.325	0.3185	−0.0065	0.325	0.3274	0.0024
0.100	0.1023	0.0230	0.100	0.1017	0.0017
0.775	0.7840	0.0090	0.775	0.7843	0.0093
0.550	0.5500	0.0000	0.550	0.5485	−0.0015
0.325	0.3223	−0.0027	0.775	0.7845	0.0095
0.325	0.3271	0.0027	0.550	0.5505	0.0005

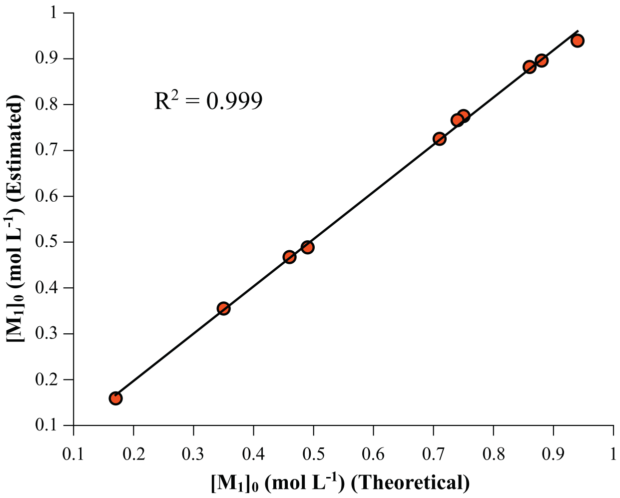

Fig. 9. Case 3: agreement between "*predicted" and* "theoretical" $[M_1]_0$ values ignoring $[M_2]_0$ values for ten test curves.

reaction are searched for in order to achieve the highest yield for the final product.

Data reported in Table 2 have been collected according to a 3^2 factorial design. A model for this system has been obtained [26] by multiple regression approach using the function:

$$f(x, y) = b_1 + b_2x + b_3y + b_4xy + b_5x^2 + b_6y^2 \tag{28}$$

However, ED-ANNs combined approach may be used as well. The temperature (T) and the time of reaction (t) were selected as input whence the yield (Y%) was selected as output parameter. The rows of the training matrix **M** (Eq. (9)) are of the form:

$$(\,T \quad t \quad Y\%\,) \tag{29}$$

To train the network only 5 points of the former 3^2 factorial design were used. The optimal neural network architecture found was (2,5,1).

In the parameter's working space, 153 points were selected according to a suitable grid and, for each point; the value of the yield was estimated using both approaches. The two response surfaces obtained, together with their difference, are shown in Fig. 10. Residual analysis of the results obtained with both ANNs and multiple regression models was performed. The mean value is equal to 0.30 (it is close to zero as it should be). Also, the standard deviation (s = 3.3) is comparable with the accuracy of the yield values. In addition, ANNs approach allows to reduce the number of experimental points needed to build the model.

Table 2
Experimental data for optimization of reaction conditions.

Temperature (°C)	Time of reaction (min)	Yield (%)
40	20	20.6
40	40	44.9
40	60	51.0
60	20	39.9
60	40	55.1
60	60	52.1
80	20	43.0
80	40	49.1
80	60	37.0

From the data obtained with the multiple regression approach, the maximum value of the yield is achieved for T = 60 °C and t = 45 min. ANNs estimate T = 60 °C and t = 40 min as optimal reaction conditions.

Using less experimental points than that required for the multiple regression, the ED-ANNs approach can model the system giving also an acceptable estimate of the optimal conditions for the reaction. This result is relevant especially for the optimization of industrial processes for which the cost of each experiment must also be considered.

Using ED-ANNs combined approach the knowledge of the hard model of the system in terms of reaction mechanism and rate constant values can be completely avoided. The optimal reaction conditions to achieve the highest yield may be found without solving the system of differential equations. Moreover, soft-modelling has another advantage with respect to hard-modelling approach. Frequently, a change of the reaction conditions causes a change of mechanism. However, using ANNs and experimental design, this possibility is implicit in the model. The change of mechanism may then be viewed as a supplementary factor not explicitly considered in the experimental design, but whose effects are taken into account implicitly.

4.5. *Effect of random errors*

The kinetic data used in Sections 4.1–4.3 were affected only by the truncation error. The evaluation of the behaviour of ANNs in the elaboration of data affected by random errors will be given here. For

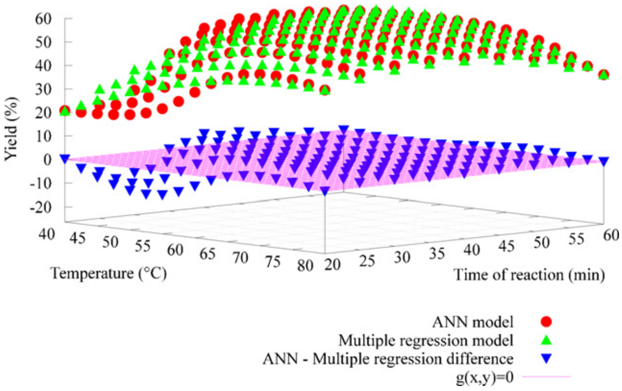

Fig. 10. Case 4: response surfaces obtained with ANN and multiple regression approach.

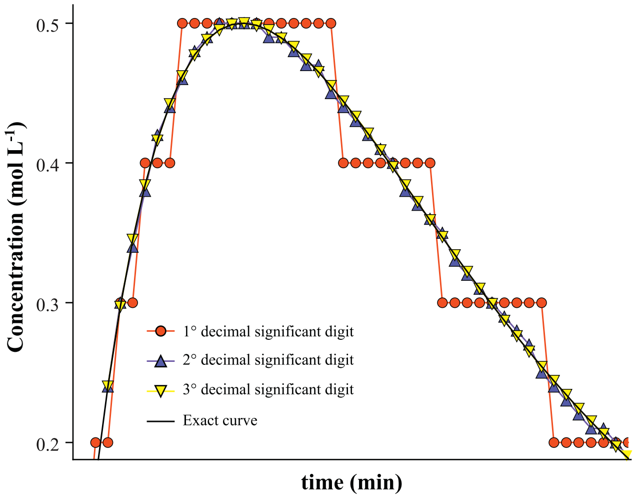

Fig. 11. Effect of roundoff errors on the concentration profile curve.

this purpose random errors were introduced in the error-free data of *Case 1* (Section 4.1). It is well known that it is not possible to generate true random numbers [27]; therefore pseudo “random” errors were introduced by rounding the significant digits of the independent variable values (concentrations) to the third, second and first decimal. The change in concentration profile is shown in Fig. 11.

The values of outputs (k_j) were kept with 4 decimal significant digits. Under this condition we reproduced the situation in which some parameters are measured with a different level of accuracy. For the three levels of rounding examined, the neural network was able to give acceptable modelling of the curves in the training matrix as for the error-free ones. It is evident that the neural network is able to compensate for the presence of different levels of roundoff errors.

The prediction ability of the trained network so obtained was evaluated using 20 new curves. The data in this verification set were rounded as before. The results obtained are expressed by the following regression lines:

Rounding level: ±0.001

$$k_{1,(\mathrm{est.})} = 0.981 \times k_{1,(\mathrm{theor.})} + 3.7 \times 10^{-4} \tag{30}$$

$$k_{2,(\mathrm{est.})} = 0.961 \times k_{2,(\mathrm{theor.})} - 9.1 \times 10^{-5} \tag{31}$$

Rounding level: ±0.01

$$k_{1,(\mathrm{est.})} = 0.954 \times k_{1,(\mathrm{theor.})} + 2.62 \times 10^{-3} \tag{32}$$

$$k_{2,(\mathrm{est.})} = 0.910 \times k_{2,(\mathrm{theor.})} + 2.94 \times 10^{-3} \tag{33}$$

Rounding level: ±0.1

$$k_{1,(\mathrm{est.})} = 0.60 \times k_{1,(\mathrm{theor.})} + 2.11 \times 10^{-2} \tag{34}$$

$$k_{2,(\mathrm{est.})} = 0.62 \times k_{2,(\mathrm{theor.})} + 1.93 \times 10^{-3} \tag{35}$$

It follows from the results that ANNs are able to give acceptable predictions for both rounding levels ±0.001 and ±0.01. These calculations clearly indicate that ANNs are able to predict the values of output parameters even from experimental data affected by rather high standard deviations.

5. Conclusions

(i) ANNs are able to model with sufficient precision any kind of kinetic curves.
(ii) Even if the performance of combined ED-ANNs approach was demonstrated for consecutive and cyclic reaction paths, this approach can be used for data concerning whichever reaction path.
(iii) ED-ANNs approach is also applicable to multicomponent kinetic analysis without any knowledge about the chemical reactions involved.
(iv) Optimization of chemical processes can be achieved by ANNs without knowing the kinetic behaviour of the system. Moreover, the use of ED-ANNs combined approach allows reducing the number of experiments needed.

Concluding, the results are showing that ANNs are a suitable tool to evaluate even complex and highly non-linear kinetic data.

Acknowledgements

Support from Ministry of Education, Youth and Sports of the Czech Republic (Projects MSM, 0021622411, 0021627501, and CZ.1.05/2.100/03.0086) and Czech Science Foundation (Projects No. 104/08/0229, 202/07/1669) are greatly acknowledged. This research has been also supported by the project R&D Center for Low-cost Plasma and Nanotechnology Surface Modifications CZ.1.05/2.1.00/03.0086 funding by European Regional Development Fund.

References

[1] N. Ingri, L.G. Sillén, Arkiv Kemi 23 (1964) 97.
[2] P. Gans, A. Sabatini, A. Vacca, Talanta 43 (1996) 1739–1753.
[3] L.G. Sillén, A Mimeographed Sheet, Department of Inorganic Chemistry, KTH, S 10044, Stockholm, 1967 (July).
[4] H. Wennerholm, R. Mellström, O. Rehnberg, N. Ingri, Chem. Scripta 7 (1975) 67.
[5] H. Gampp, M. Maeder, C.J. Meyer, A.D. Zuberbühler, Talanta 32 (1985) 251–264.
[6] J.L. González-Hernández, J. Dušek, J. Havel, Kinet. Catal. L. 38 (1989) 385–393.
[7] J.L. González-Hernández, M.M. Canedo, C. Grande, Chemometr. Intell. Lab. 39 (1997) 77–84.
[8] S. Vajda, H. Rabitz, J. Phys. Chem. 98 (1994) 5265–5271.
[9] K.A. Connors, Chemical Kinetics. The Study of Reaction Rates in Solution, VCH, Publishers, Inc., 1990.
[10] D. Pérez-Bendito, M. Silva, Trends Anal. Chem. 15 (1996) 232–240.
[11] E.J. Molga, B.A.A. vanWoezik, K.R. Westerterp, Chem. Eng. Process. 39 (2000) 323–334.
[12] D. Baş, F.C. Dudak, H. Boyaci, J. Food Eng. 79 (2007) 622–628.
[13] P. Valeh-e-Sheida, F. Yaripour, G. Moradi, M. Saber, Ind. Eng. Res. 49 (2010) 4620–4626.
[14] D. Baş, F.C. Dudak, H. Boyaci, J. Food Eng. 79 (2007) 1152–1158.
[15] N.H.T. Lemes, E. Borges, J.P. Braga, Chemometr. Intell. Lab. 96 (2009) 84–87.
[16] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, M. Redon, Anal. Chem. 67 (1995) 4477–4483.
[17] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, M. Redón, N. Villegas, Analyst 121 (1996) 395–400.
[18] S. Ventura, M. Silva, D. Pérez-Bendito, Anal. Chem. 67 (1995) 4458–4461.
[19] Y. Li, A.M. Rauth, X. Yu Wu, Eur. J. Pharm. Sci. 24 (2005) 401–410.
[20] J. Havel, E.M. Peña-Méndez, A. Rojas-Hernández, J.P. Doucet, A. Panaye, J. Chromatogr. A 793 (1998) 317–329.
[21] M. Leshno, V.Y. Lin, A. Pinkus, S. Schocken, Neural Netw. 6 (1993) 861–867.
[22] D.L. Massart, B.G.M. Vandeginste, L. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A. Data Handling in Science and Technology, vol. 20A, Elsevier Science, 1997.
[23] R.G. Brereton, Chemometrics Data Analysis for the Laboratory and Chemical Plant, vol. 8, Wiley, 2003.
[24] M.N. Berberan-Santos, J.M.G. Martinho, J. Chem. Educ. 67 (1990) 375–379.
[25] J. Havel, F. Jiménez, R.D. Bautista, J.J. León, Analyst 118 (1993) 1355–1360.
[26] R. Leardi, Anal. Chim. Acta 652 (2009) 161–172.
[27] P. Hellekalek, Math. Comput. Simulat. 46 (1998) 485–505.