An analysis of mixtures using amperometric biosensors and artificial neural networks

R. Baronas*

Faculty of Mathematics and Informatics, Vilnius University, Naugarduko 24, 2600 Vilnius, Lithuania
E-mail: romas.baronas@maf.vu.lt

F. Ivanauskas

Faculty of Mathematics and Informatics, Vilnius University, Naugarduko 24 and Institute of Mathematics and Informatics, Akademijos 4, 2600 Vilnius, Lithuania
E-mail: feliksas.ivanauskas@maf.vu.lt

R. Maslovskis and P. Vaitkus

Faculty of Mathematics and Informatics, Vilnius University, Naugarduko 24, 2600 Vilnius, Lithuania

This paper presents a sensor system based on a combination of an amperometric biosensor acting in batch as well as flow injection analysis with the chemometric data analysis of biosensor outputs. The multivariate calibration of the biosensor signal was performed using artificial neural networks. Large amounts of biosensor calibration as well as test data were synthesized using computer simulation. Mathematical and corresponding numerical models of amperometric biosensors have been built to simulate the biosensor response to mixtures of compounds. The mathematical model is based on diffusion equations containing a non-linear term related to Michaelis–Menten kinetics of the enzymatic reaction. The principal component analysis was applied for an optimization of calibration data. Artificial neural networks were used to discriminate compounds of mixtures and to estimate the concentration of each compound. The proposed approach showed prediction of each component with recoveries greater that 99% in flow injection as well as in batch analysis when the biosensor response is under diffusion control.

KEY WORDS: reaction-diffusion, modelling, biosensor, neural network

1. Introduction

Biosensors are devices that combine the selectivity and specificity of a biologically active compound with a signal transducer and an electronic amplifier

---

* Corresponding author.

[1–3]. The transducer converts the biochemical signal to an electronic one. The biosensor signal is proportional to the concentration of a measured analyte or a group of analytes. The amperometric biosensors measure the current on an indicator electrode due to direct oxidation of the products of the biochemical reaction. In case of the amperometric biosensors, the potential at the electrode is held constant while the current flow is measured. The amperometric biosensors are reliable, relatively cheap and sensitive enough for environment, clinical and industrial purposes [4,5].

Traditionally, changing biological process variables have been treated separately and systematically [6,7]. In recent years, analytical multivariate methods have introduced refined ways to handle complex signal inputs and to interpret their relations to selected observations [8–11]. Methods such as artificial neural networks [12–14] become powerful tools for experimental data analysis to improve sensitivity and selectivity of sensor systems [15–17].

Data collected from complicated samples or in complex processes contains variation from any sources and of several types. Pre-processing methods can be applied in such situations to enhance the relevant information to make resulting models simpler and easier to interpret. The principal component analysis (PCA) allows us to achieve acceptable computing times; the experimental data are therefore efficiently compressed without losing any important information [18–22]. Accordingly, coupling biosensors with artificial neural networks is growing in importance as a tool for multi-component analyses [23–28].

This paper presents a sensor system based on a combination of an amperometric biosensor with the chemometric data analysis of biosensor outputs. Both modes of analyte analysis: batch (BA) and flow injection (FIA) are supported [29]. Artificial neural networks are used for the multivariate calibration of the biosensor signal. Data for biosensor calibration are synthesized using computer simulation. The synthesized data are allocated into two different data sets: a calibration set, with which the PCA is performed, and a test set, which is used for assessing the results of the calibration.

An accurate and reliable calibration of the system as well as the proper test of the methods of chemometrics requires a lot of experimental data. Mathematical and corresponding numerical models of amperometric biosensors were built to generate pseudo experimental biosensor responses to mixtures of compounds. The models are based on non-stationary diffusion equations [30] containing a non-linear term related to Michaelis–Menten kinetics of the enzymatic reaction. The digital simulation was carried out using the finite difference technique [31, 32]. Assuming good enough adequacy of the mathematical model to the physical phenomena, the data synthesized using computer simulation were employed instead of experimental ones. The computer simulation of the physical experiment is usually much cheaper and faster than the physical one. The computer simulation is especially reasonable when the biosensors to be used in practice are in a stage of development. Then the development of smart biosensors to be used

in biochemical systems and the development of effective methods of data analysis may be carried out in parallel.

The proposed system was applied to discriminate compounds of mixtures and to estimate the concentration of each compound. The neural network was calibrated and validated for mixtures of four specific analytes of eight different concentrations. The numerical experiments showed good prediction of each component.

In the case of FIA, the concentrations of each compound were determined with recoveries greater than 99%. Very similar recoveries of concentrations were obtained also in batch analysis when the mass transport by diffusion controls the biosensor response, that is, when the diffusion modulus is greater than unity. In batch analysis, when enzyme kinetics controls the biosensor response, the concentrations were determined with recoveries between 58 and 87%.

2. Mathematical model

We consider an enzyme-catalysed reaction

$$S_k \xrightarrow{\text{E}} P_k, \qquad k = 1, \ldots, K, \tag{1}$$
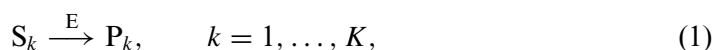
where the mixture of substrates (components) (S_k, $k = 1, \ldots, K$) binds to the enzyme (E) to form enzyme–substrate complex. While it is a part of this complex, the substrate S_k is converted to the product P_k. The rate of the reaction is the rate of appearance of the product. This rate is known to depend upon the concentration of substrate.

An amperometric biosensor can be treated as enzyme electrode, having a layer of enzyme immobilized onto the surface of the probe. Assuming no interaction between analysed substrates (compounds) of the mixture, the symmetrical geometry of the electrode, homogeneous distribution of immobilized enzyme in the enzyme membrane, and considering one-dimensional diffusion, coupling of enzyme reaction with the diffusion described by Fick's law leads to the following equations:

$$\frac{\partial S^{(k)}}{\partial t} = D_S^{(k)} \frac{\partial^2 S^{(k)}}{\partial x^2} - \frac{V_{\max}^{(k)} S^{(k)}}{K_M + S^{(k)}}, \qquad 0 < x < d, \qquad 0 < t \leqslant T, \tag{2}$$

$$\frac{\partial P^{(k)}}{\partial t} = D_P^{(k)} \frac{\partial^2 P^{(k)}}{\partial x^2} + \frac{V_{\max}^{(k)} S^{(k)}}{K_M + S^{(k)}}, \qquad 0 < x < d, \qquad 0 < t \leqslant T,$$
$$k = 1, \ldots, K, \tag{3}$$

where K is the number of compounds, $V_{\max}^{(k)}$ is the maximal enzymatic rate of biosensor attainable with that amount of enzyme, when the enzyme is fully saturated with substrate (component) S_k, K_M is the Michaelis constant, $S^{(k)}$ is the

concentration of substrate S_k, $P^{(k)}$ is the concentration of the reaction product P_k, d is the thickness of the enzyme layer, t is time, T is the duration of biosensor operation, and $D_S^{(k)}$ and $D_P^{(k)}$ are diffusion coefficients of the substrate S_k and product P_k, respectively, $k = 1, \ldots, K$, [3,33,34].

The biosensor operation starts when some substrate appears over the surface of the enzyme layer. This is used in the initial conditions ($t = 0$)

$$S^{(k)}(x, 0) = \begin{cases} 0, & 0 \leqslant x < d, \\ S_0^{(k)}, & x = d, \end{cases} \tag{4}$$

$$P^{(k)}(x, 0) = 0, \quad 0 \leqslant x \leqslant d, \\ k = 1, \ldots, K, \tag{5}$$

where $S_0^{(k)}$ is the concentration of substrate S_k over the biosensor (in bulk solution).

Because of electrode polarization, the concentration of the reaction products at the electrode surface is being permanently reduced to zero. If the analyte is well stirred and in powerful motion, then the diffusion layer ($0 < x < d$) remains at a constant thickness. Consequently, the concentration of substrates as well as products over the enzyme surface (bulk solution/membrane interface) remains constant while the biosensor contact with the substrate. In the FIA, the biosensor contacts the substrate for only a short time [29]. When the analyte disappears, a buffer solution swills the enzyme surface, reducing the substrate as well as product concentration at this surface to zero. Because of substrates remaining in the enzyme membrane, the mass diffusion as well as the reaction still continues for some time even after disconnecting of the biosensor and substrate. This is used in the boundary conditions ($0 < t \leqslant T$) given by

$$\left. \frac{\partial S^{(k)}}{\partial x} \right|_{x=0} = 0, \tag{6}$$

$$S^{(k)}(d, t) = \begin{cases} S_0^{(k)}, & t \leqslant T_F, \\ 0, & t > T_F, \end{cases} \tag{7}$$

$$P^{(k)}(0, t) = P^{(k)}(d, t) = 0, \\ k = 1, \ldots, K, \tag{8}$$

where T_F is the time of injection.

In the BA, the modelled biosensor remains immersed in the substrate all the analysing time. Assuming $T_F = T$, the model expressed by (2)–(8) may be accepted for BA as well. In the BA, the boundary condition (7) reduces to $S^{(k)}(d, t) = S_0^{(k)}$, $t \leqslant T$.

The current is measured as a response of a biosensor in a physical experiment. The biosensor current depends upon the flux of reaction product at the electrode surface, that is, at border $x = 0$. Consequently, the density $I^{(k)}(t)$ of the biosensor current at time t can be obtained explicitly from Faraday's law

$$I^{(k)}(t) = n_e F D_P^{(k)} \left. \frac{\partial P^{(k)}}{\partial x} \right|_{x=0} \qquad k = 1, \ldots, K, \tag{9}$$

where n_e is a number of electrons involved in a charge transfer at the electrode surface, and F is Faraday constant. Assuming the overall biosensor response to a mixture represents the sum total of individual responses to each constituent substrate and having values of the current $I^{(k)}(t)$ for all compounds, $k = 1, \ldots, K$, the common density $I^*(t)$ of the biosensor current can be calculated additively

$$I^*(t) = \sum_{k=1}^{K} I^{(k)}(t). \tag{10}$$

3. Solution of the problem

Analysing the problem (2)–(8), one can notice that there is no direct relationship between pairs of the unknown variables $S^{(k_1)}$, $P^{(k_1)}$ and $S^{(k_2)}$, $P^{(k_2)}$, when $k_1 \neq k_2$, $k_1, k_2 = 1, \ldots, K$. Because of this, the initial and boundary value problem (2)–(8), which consists of $7K$ equations can be split to K problems, containing only seven equations (2)–(8) at given $k, k = 1, \ldots, K$. The problem (2)–(8), formulated for given k_1 (component S_{k_1}), can be solved individually and independently from the problem, formulated for another component S_{k_2}, $k_1, k_2 = 1, \ldots, K$, $k_1 \neq k_2$.

We assume the problem (2)–(8) formulation for a single substrate $S = S_k$ and reaction product $P = P_k, k = K = 1$. Let $V_{\max}$ be the maximal enzymatic rate of the modelled biosensor, S is the concentration of the substrate S and P is the concentration of the reaction product P.

The problem (2)–(8), formulated for a single substrate S and reaction product P, was solved numerically using the finite difference technique [31–36].

In the common case of K components, having responses of the biosensor to each constituent component, equation (10) allows us to calculate the common biosensor response to the mixture of K components [37]. Consequently, to obtain values $I^*(t_j), j = 0, \ldots, N, t_0 = 0, t_N = T$, of the common biosensor current, it is required to

(1) run computer simulation K times to obtain values $I^{(k)}(t_j)$ of the biosensor current for each component of the mixture, $K = 1, \ldots, K$, $j = 1, \ldots, N$;

(2) calculate the common biosensor current as defined in (10).

In step (1), only values of the parameters $D_S^{(k)}$, $D_P^{(k)}$, $V_{\max}^{(k)}$, and $S_0^{(k)}$ vary when one computer simulation changes the next one. This procedure of computation is valid for both regimes: BA and FIA.

4.   Generation of data sets

The computer simulation software developed was employed to generate data for a calibration of an amperometric biosensor. The biosensor was calibrated for mixtures of four ($K = 4$) components. Each component of eight ($M = 8$) different concentrations was employed in the calibration to have the biosensor response to a wide range of substrate concentrations. Because of this, it was required to solve the problem (2)–(8) for given component S_k numerically $K \times M = 4 \times 8 = 32$ times at four different values of the maximal enzymatic rate $V_{\max}^{(k)}$ and eight values of the substrate concentration $S_0^{(k)}$.

The following values of the model parameters were assumed constant in the all numerical experiments:

$$D_S^{(k)} = D_P^{(k)} = 3 \times 10^{-6}\,\text{cm}^2/\text{s}, \qquad k = 1, \ldots, K,$$
$$K_M = 10^{-7}\,\text{mol/cm}^3, \quad n_e = 2. \tag{11}$$

Each component of the mixture was characterized by the individual maximal enzymatic rate $V_{\max}^{(k)}$:

$$V_{\max}^{(k)} = 10^{-6-k}\,\text{mol/cm}^3\,\text{s}, \qquad k = 1, \ldots, K. \tag{12}$$

To have the biosensor responses for a wide range of the analyte concentration, the following values of the concentration $S_0^{(k)}$ of each of K substrates $S_1, \ldots, S_K$ of the mixture were employed:

$$S_0^{(k)} \in \left\{ S_{0,m} : S_{0,m} = \alpha_m \times S_0, m = 1, \ldots, M \right\}, \qquad k = 1, \ldots, K, \tag{13}$$
$$S_0 = 10^{-8}\,\text{mol/cm}^3, \qquad K = 4, M = 8,$$
$$\alpha_1 = 1, \ \alpha_2 = 2, \ \alpha_3 = 4, \ \alpha_4 = 8, \ \alpha_5 = 12, \ \alpha_6 = 16, \ \alpha_7 = 32, \ \alpha_8 = 64.$$

The biosensor response considerably depends upon being the response under the enzyme kinetics either the mass transport by diffusion control [2,3, 34,38]. The biosensor response is controlled by the diffusion if the enzymatic reaction in the enzyme membrane is faster than the mass transport process, that is, when the dimensionless diffusion modulus is much greater than unity. We express the diffusion modulus for the enzyme reaction with the substrate S_k ($k = 1, \ldots, K$) as a function of the membrane thickness d

$$\sigma_k^2(d) = \frac{V_{\max}^{(k)} d^2}{D_S^{(k)} K_M} \approx 0.33 \times 10^{7-k} d^2. \tag{14}$$

To validate the methods of chemometrics more comprehensively, in calculations we employed two enzyme membranes with different thickness d : 0.02 and 0.05 cm. We notice that $\sigma_k^2(0.02) > 1$ for $k = 1$, 2, 3 and $\sigma_4^2(0.02) < 1$, that is, in the case of $d = 0.02$ cm, the biosensor response is under dual control: by the diffusion (for components S_1, S_2, S_3) and the enzyme kinetics (S_4). The another thickness $d = 0.05$ cm was chosen so that $\sigma_k^2(0.05) > 1$ for $k = 1, \ldots, 4$. In the case of $d = 0.05$ cm, the response is controlled by the diffusion.

Values of two parameters T_F and T depend considerably on the regime of analysis and the membrane thickness d. In FIA, due to the disappearance of the current, time T was considerably less than in BA. In the case of $d = 0.02$ cm, we employed $T = 300$ s, $T_F = T$ in BA, while $T = 150$ s, $T_F = 10$ s in FIA. In the case of thicker membrane ($d = 0.05$ cm), we accepted $T = 500$ s as the response time for BA.

In simulation of the biosensor response for all the values defined in (13), only values of two parameters $V_{\max}^{(k)}$ and $S_0^{(k)}$ vary when one computer simulation changes the next one. In addition, every computer simulation was repeated twice to simulate biosensor response in batch as well as flow injection mode at different values of T_F and T.

Let $I_m^{(k)}(t_j)$ be a value of density $I^{(k)}(t_j)$ of the biosensor current at concentration $S_0^{(k)} = S_{0,m}^{(k)}$ of substrate S_k, $m = 1, \ldots, M$; $j = 1, \ldots, N$; $k = 1, \ldots, K$. Having M numerical solutions (M sets of biosensor response values) $I_m^{(k)}(t_j)$, $j = 1, \ldots, N$ for each $k = 1, \ldots, K$ (in total $K \times M$ solutions), the full factorial $I_m^*(t_j) = I_{m_1, \ldots, m_K}(t_j)$ of $M^K = 8^4 = 4096$ solutions can be produced additively

$$I_m^*(t_j) = \sum_{k=1}^{K} I_{m_k}^{(k)}(t_j), \quad \vec{m} = (m_1, \ldots, m_K), \quad m_k = 1, \ldots, M; \quad j = 1, \ldots, N.$$

(15)

During the computer simulation, values of the biosensor current were stored in a file every second of simulation. Thus, $N = T$ values of $I_m^{(k)}(t_j)$, $t_j = j$ (s), $j = 1, \ldots, N$, for each $K = 1, \ldots, K$ and $m = 1, \ldots, M$ were produced as a result of the computer simulation of the biosensor response (in total $K \times M \times N$ values). Later, using an additional simple utility of summation, a matrix $M^K \times N$ of the biosensor response data were produced following (15) and stored in a file which was passed for chemometrics. This was repeated for batch as well as flow injection regimes.

Results of the calculations are depicted in figures 1 and 2. Figure 1 shows every 64th of full factorial of M^K simulated biosensor responses for $K = 4$ values of the maximal enzymatic rate and $M = 8$ substrate concentrations in a case of BA. Figure 2 presents generated biosensor responses in FIA. Evolution of biosensor currents are depicted for the first 100 s of biosensor action only because of a petty change of the biosensor current at greater values of time t.

The calculation showed that the maximal biosensor current increases with increase of the maximal enzymatic rate V_{max}. The time of the maximal biosensor current decreases with increase of V_{max}. This property is valid for both regimes of analysis: batch and flow injection. In BA, the maximal biosensor current is the steady-state current. Figure 2 shows, that the current function $I^*(t)$ is not monotonous in FIA. The time of maximal current occurs noticeably later after the time $T_F = 10\,\text{s}$ of analyte removing. The time when the current starts to decrease varies between 19 and 24 s.

5. Determination of component concentrations using an artificial neural network

Let $\vec{c} = (c_1, \ldots, c_K)$ be a vector of concentrations of K components of a mixture and $\vec{z} = \vec{z}(\vec{c}) = (z_1(\vec{c}), \ldots, z_N(\vec{c})) = \left(I_{\vec{c}}^*(t_1), \ldots, I_{\vec{c}}^*(t_N)\right)$ be a vector of the biosensor currents measured at times $t_1, \ldots, t_N$. Thus, $\vec{z}$ defines a response of the biosensor to the mixture of K components of concentrations $\vec{c}$. Our goal is to define a non-linear map $\mathbf{N}$, such as $\mathbf{N}(\vec{z}) = \vec{c}$. The class of feed-forward neural networks (FNN) was chosen to approximate the map $\mathbf{N}$ [13, 39].

While an artificial neural network provides a non-linear approach that needs no *a priori* knowledge of functional dependencies, it requires training. Training is based upon cumulative experimental data. An accurate and reliable calibration of the system as well as the proper test of the methods of
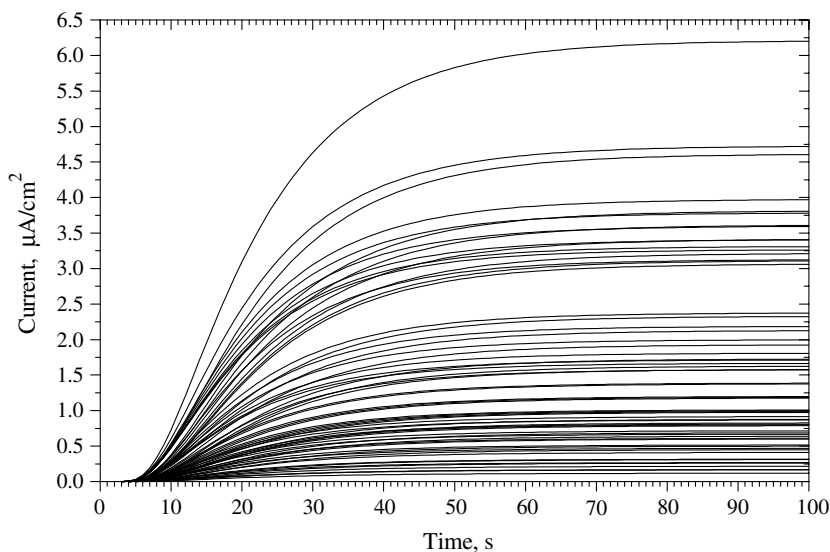

Figure 1. Every 64th biosensor response curve of full factorial of M^K responses at $K = 4$ values of the maximal enzymatic rate and $M = 8$ substrate concentrations in batch analysis, $d = 0.02\,\text{cm}$.
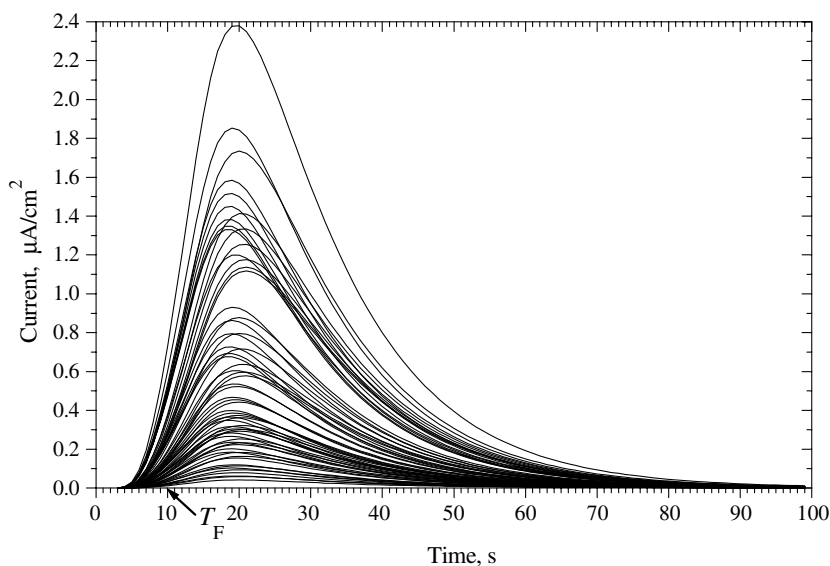
Figure 2. Every 64th biosensor response curve of full factorial of M^K responses at $K = 4$ values of the maximal enzymatic rate and $M = 8$ substrate concentrations in flow injection analysis, $d = 0.02$ cm.

chemometrics requires a lot of experimental data. Assuming good enough adequacy of the mathematical model to the physical phenomena, the data synthesized using computer simulation may be employed instead of experimental one. The computer simulation of experiments is usually much cheaper and faster than the physical one. The computer simulation is especially reasonable when biosensors to be used in practice are in a stage of development. Then the development of smart biosensors to be used in biochemical systems and the development of effective methods of data analysis may be carried out in parallel.

5.1. Input data compression

Data collected in complicated processes contains a lot of redundant information, since the variables are collinear. Pre-processing methods can be applied in such situations to enhance the relevant information to make resulting models simpler and easier to interpret. Two different approaches were applied to reduce the dimensionality of the vector $\vec{z}$ of input data: the correlation coefficients analysis (CCA) and the PCA.

With CCA, the correlation coefficients are calculated for every point of the input vector $\vec{z}$ and each component of the mixture. Based on the calculated correlation coefficients, only a few points of $\vec{z}$ with the highest values of the coefficients are accepted as inputs to a neural network. Therefore, every vector $\vec{z}$ of

the original data is replaced with a resulting vector $\vec{x}$, having the dimensionality J, $J \leqslant N$. Each element of the resulting vector $\vec{x}$ belongs also to $\vec{z}$. However, a lot of elements of $\vec{z}$ are usually missing in $\vec{x}$. The CCA resulting vector $\vec{x}$ is passed to a neural network.

With the more advanced PCA, the so-called principal components are extracted, which are statistically independent from each other and which are therefore orthogonal relative to one another, yet are still capable of adequately reconstructing the original data [18,19]. Using PCA, the input data $\vec{z}$ are expressed as a linear combination of a few basic vectors, were the basic vectors capture as much variation of original data as possible. The purpose of PCA is to find basic vectors $\vec{w}_1, \ldots, \vec{w}_J$, $J \leqslant N$, for $i = 1, \ldots, J$, satisfying the following conditions:

(1) $\mathbf{E}\{\vec{w}_i^{\mathrm{T}}\vec{z}\}^2$ are maximized under the constraints,

(2) $\vec{w}_i^{\mathrm{T}}\vec{w}_j = \delta_{ij}$, for $j = 1, \ldots, i$,

where $\mathbf{E}$ stands for an expectation operator, N is the number of measurements of biosensor current during the biosensor operation and the superscript T indicates transposition. The vectors $\vec{w}_1, \ldots, \vec{w}_J$ satisfying that conditions can be found as dominant eigenvectors of the data covariance matrix $\Sigma = \mathbf{E}\{\vec{z}\vec{z}^{\mathrm{T}}\}$. Therefore, each vector $\vec{z}$ of original data can be represented by its principal component vector $\vec{x}$ having dimensionality J, $J \leqslant N$. In the cases of large dimensionality of input vector $\vec{z}$, the dimensionality of the resulting vector $\vec{x}$ is usually significantly less than the dimensionality of $\vec{z}$, $J \ll N$. There exist some rules of thumb on how many dimensions to use, such as keeping all dimensions whose contribution to the total variation exceed 80%. The PCA resulting vector $\vec{x}$ is passed to a neural network.

5.2. Artificial neural network set-up

Several network topologies were applied in preliminary experiments, whereby networks with two hidden layers were accepted in the cases when CCA is used to reduce the dimensionality of the input vector, while a single hidden layer was accepted in the cases when input data are corrected with the PCA.

Figure 3 shows a three-layered multi-input and multi-output FNN. The first layer is the input layer with J nodes. Each input neuron represents a response of a biosensor to a mixture. The second layer is the hidden layer consisting of p nodes with sigmoid functions as their activation (transfer) functions. The third layer is the output layer comprises K nodes with linear functions. Each output neuron represents a concentration of the mixture component. The non-linear mapping from $\vec{x}$ to $\vec{c}$, $\vec{c} = (c_1, \ldots, c_K)$, can be expressed as follows
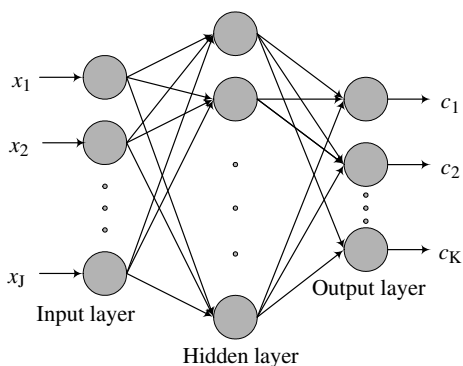
Figure 3. Schematic diagram of three-layer artificial neural network, where $x_1, \ldots, x_J$ are values of the biosensor current and $c_1, \ldots, c_K$ are the determined concentrations of mixture components.

$$c_k = \sum_{s=1}^{p} \alpha_{s,k} \varphi \left(\left\langle \vec{\beta}_s^{\mathrm{T}}, \vec{x} \right\rangle + \gamma_s \right) + \varepsilon_k, \qquad k = 1, \ldots, K, \qquad (16)$$

where c_k is the output of kth output node expressing the concentration of kth component of the mixture, p is number of nodes in the hidden layer; $\alpha_{sk}, \vec{\beta}_s, \gamma_s, \varepsilon_k$ are the weights and φ is the activation function. The sigmoid (logistic) function was employed as the activation function φ, that is, $\varphi((u) = 1/(1 + \exp(-u))$. The number p of nodes in the hidden layer was chosen basing on the thumb rules. Having L observations (elements) in the learning set, the degree of freedom in the neural network should not exceed $0.1 \times L$, that is, $(p + 1) \times K + p(J + 1) < 0.1 \times L$.

The neural network was trained by the supervised batch learning procedure that requires a set of examples for which the desired network response is known. The learning of the network was carried out comparing the calculated target values and the desired outputs by means of calculation of the sum square error. In the learning, values of the weights $\alpha_{sk}, \vec{\beta}_s, \gamma_s, \varepsilon_k, s = 1, \ldots, p, \ k = 1, \ldots, K$, are updated after the whole training data set has been passed for the network.

An advanced variant of back-propagation (BP) algorithm called Levenberg–Marquardt (LM) was used to optimize the process of learning [40]. BP is one of the most commonly used training algorithms and LM is one of the fastest variants of BP for networks of moderate size [41]. It is important for learning, that values of $\vec{b}_s, s = 1, \ldots, p$, are the same for all $k = 1, \ldots, K$. The resulting network is able to generalize (give a good response) when presented with cases not found in the set of examples.

The forecasting quality of concentration $y_m = S_{0,m}^{(k)}$ of each component $S_k, k = 1, \ldots, K, m = 1, \ldots, M$, was estimated by the mean parameter

(percentage of true interval predictions)

$$Q_k = \frac{1}{L} \sum_{i=1}^{L} \text{Ind}\left(O_{i,k} \in \Delta y\right) \cdot \text{Ind}\left(C_{i,k} = y\right) \cdot 100\%, \tag{17}$$

where the indicator function $\text{Ind}\left(O_{i,k} \in \Delta y\right)$ equals unity when kth output of the network belongs to the interval $(y - \delta_{1,y}, y + \delta_{2,y})$ of concentrations, and zero otherwise, L is a number of observations in calibration or test set.

6.   Results and discussion

To achieve acceptable computing times, the generated pseudo-experimental data were therefore efficiently compressed with the aid of CCA as well as PCA. During calibration, the input data are allocated to two different data sets:

(1) calibration set (C-set), with which the compression is performed;

(2) test set (T-set), which is used for assessing the results of the calibration.

The data sets, which were generated by measuring mixtures of all the components, are based on the full factorial designs, ensuring that no constraint of a constant total concentration of the analytes is presented for the calibration. It is a crucial point when using neural networks for data analysis to validate the calibration by independent test data. On the other hand, the chosen samples for the calibration set must be independent from the test set. Because of this, C-set was constructed in random order.

The total set of full factorial of $M^K = 8^4 = 4096$ responses was split randomly into C and T sets having approximately the same number of responses. A total of 2000 response curves were chose independently as a T-set. The remaining 2096 samples were accepted as a learning (C) set.

6.1. Correlation coefficients analysis

The CCA was applied to C-set of responses calculated at $d = 0.02 \, \text{cm}$ for both regimes of analysis: batch and flow injection. For every point of the input vector $\vec{z} \in$ C-set and each component we calculated the correlations. The points with the highest correlation coefficients were used as inputs to the neural networks. Let $\vec{z} = (z_1, z_2, \ldots, z_N)$ be a vector of the points on the response and $\rho(c_k, z_i)$ the correlation coefficient for the compounds c_k and the point $z_i, 1 \leqslant i \leqslant N, 1 \leqslant k \leqslant K$.

In the case of BA, when $d = 0.02 \, \text{cm}$ and $N = 300$, the calculated highest values are: $\rho(c_1, z_2) = 0.9886$, $\rho(c_1, z_3) = 0.9794$, $\rho(c_1, z_{75}) = 0.6683$, etc. Vectors having only six points $\vec{x} = (x_1, \ldots, x_J) = (z_2, z_3, z_{75}, z_{77}, z_{295}, z_{297})$, $J = 6$, were

accepted as inputs to a neural network. The network with two hidden layers was chosen. In the first hidden layer there were 16 neurons and in the second—eight neurons. Because of six input points, the input layer consists of six nodes. Due to mixtures of four components used in the analysis, the output layer was of four nodes.

Using (17), the forecasting quality Q_k of concentration $y = S_{0,m}^{(k)}$ of the mixture component $S_k, k = 1, \ldots, K, m = 1, \ldots, M$, was estimated by the mean parameter accepting the intervals $\Delta = \Delta_1$ of the prediction accuracy. The intervals Δ_1 are defined in Table 1. Corresponding values Q_k of the forecasting quality are presented in Table 2.

On can see in Table 1, the designed neural network well enough estimates the concentrations of three components only. The concentration of the component S_4 is estimated with forecasting quality less than 40%, $Q_4 = 38.8\%$ at learning set and 36.9% at T-set. In the case of the other three components, the forecasting quality is grater than 82%, $Q_k > 82\%$, at C as well as T-set.

Variation of the recognition quality of different components well correlates with variation of the diffusion modulus (14). The diffusion modulus is less than unity for one component (substrate) S_4 only, $k = 4$, $\sigma_4^2(0.02) < 1$. For all other components the diffusion modulus is greater than unity. To approve the influence of the diffusion modulus on the quality of estimation of concentrations, we applied that procedure of estimation to a system having 2.5 times thicker enzyme membrane, $d = 0.05$ cm.

Table 1
The accuracy intervals Δ_1 and Δ_2 for prediction of the analyte concentration y using forecasting quality, defined in (17).

y, nmol/cm^3	1	2	4	8	12	16	32	64
Δ_1 y, nmol/cm^3	<1.5	[1.5,3)	[3,6)	[6,10)	[10,14)	[14,24)	[24,48)	$\geqslant 48$
Δ_2 y, nmol/cm^3	[0,1.5)	[1.5,2.9)	[3.1,5)	[7,9)	[11,13)	[15,17)	[31,33)	[63,65)

Table 2
The forecasting quality (17) of concentrations prediction at calibration (C) and test (T) sets using both accuracies: Δ_1 and Δ_2 in batch (BA) and flow injection (FIA) modes, when input data were processed with CCA.

Component, k	BA, Δ_1, $d = 0.02$		BA, Δ_1, $d = 0.05$		FIA, Δ_2, $d = 0.02$	
	C-set	T-set	C-set	T-set	C-set	T-set
1	89.88	90.5	98.52	98	100	100
2	92.89	93	98.18	98.9	100	100
3	82.6	82.69	93.46	93.95	100	100
4	38.8	36.9	78.67	78.95	99.95	100

Values Q_k of the forecasting quality at the membrane thickness $d = 0.05$ cm are presented also in Table 2. In that case, $N = 500$ points $(z_1, \ldots, z_N)$ of each biosensor response were employed in the analysis. Similarly to the previous case, after CCA analysis, $J = 6$ input points $(z_2, z_3, z_{49}, z_{50}, z_{499}, z_{500})$, were selected as inputs for the neural network. The network architecture (6–16–8–4) was the same as in the case of thinner enzyme membrane.

One can see in Table 2, for the same batch mode of analysis, the forecasting values are significantly higher than in the case of thinner enzyme layer, $d = 0.02$ cm. In the case of component S_4, $k = 4$, the quality is even more than two times higher, $Q_4 \approx 79\%$. In all the other cases the forecasting quality is also better, $Q_k > 93\%$, $k = 1, 2, 3$. Consequently, in batch analysis, the quality of estimation of the component concentrations can be increased by increasing the thickness d of the enzyme membrane or, more accurately, by adjusting conditions of the analysis so, that the diffusion modulus increases.

In the analysis of the data obtained in the flow injection regime of the biosensor operation, distinctly better results of the estimation of concentrations were achieved. The results remained very good even if the intervals Δ of the prediction accuracy get narrow. The employed intervals $\Delta = \Delta_2$ of the prediction accuracy are defined in Table 1.

For the processing of FIA data, a neural network with only one hidden layer containing seven neurons was constructed. In that case $N = 150$ points $(z_1, \ldots, z_N)$ of the biosensor response was employed in the analysis. After CCA analysis, $J = 4$ input points $(z_2, z_{37}, z_{113}, z_{115})$, were enough for the input of the neural network. Results of the calibration and test of the network of architecture (4–7–4) are presented in Table 2.

One can see in Table 2, the results of concentrations prediction, obtained on data produced in FIA, are considerably better than the results obtained in BA. The concentrations of all the components were predicted with recoveries greater that 99.9%. Let us notice, that such level of recognition quality was obtained at the membrane thickness $d = 0.02$ cm, when the diffusion modulus calculated for component S_4 is less than unity, that is the biosensor response is under enzyme kinetics control.

6.2. Principal component analysis

Application of PCA to the C-sets resulted with six principal components in all three cases of biosensor operation: (1) BA at membrane thickness $d = 0.02$ cm, (2) BA at $d = 0.05$ cm and (3) FIA at $d = 0.02$ cm. Due to the PCA, neural networks having six nodes in input layer were employed. Since mixtures to be analysed consist of four components, the networks have four nodes in output layer. Additional analysis showed that one hidden layer is enough to achieve sufficiently good results of concentrations estimation. Twelve nodes in a single

hidden layer were used in the case of BA when membrane thickness d is $0.02\,cm$. In two other cases of BA at $d = 0.05\,cm$ and FIA at $d = 0.02\,cm$, eight nodes in a single hidden layer were enough. Values Q_k, $k = 1, 2, 3, 4$, of the forecasting quality, calculated for narrower intervals Δ_2 (see Table 1) of prediction accuracy are presented in Table 3. For the BA applied at $d = 0.02\,cm$, Table 3 presents also results of concentrations recognition at wider intervals Δ_1. The calculated recognition qualities at the wider intervals Δ_1 in two other cases (BA at $d = 0.05\,cm$ and FIA) were practically the same as at the intervals Δ_2. Because of this, these results are missing in Table 3.

Comparing results presented in Table 3 with the results from Table 2, one can see that in the case of BA, input data pre-processing with the PCA is considerably preferable in comparison with CCA. However, the recognition quality of biosensor data obtained in FIA is practically the same for both types of input data pre-processing: CCA and PCA.

7. Conclusions

In the cases when no interaction between components of a mixture the mathematical model (2)–(8) describes an operation of amperometric biosensors in BA and FIA. The problem (2)–(8) can be solved numerically for each component independently. The common biosensor current is calculated additively from the individual biosensor responses to each constituent component.

Computer simulation of the biosensor response can be used to generate pseudo-experimental biosensor responses to mixtures of compounds. Assuming K is a number of mixture components and M is a number of different concentrations of each component, the biosensor response data for full factorial of mixtures (M^K samples) can be synthesized by a simple routine of summation from the results of $K \times M$ computer simulations of the response. The generated data was employed to calibrate and validate a sensor system based on an

Table 3

The forecasting quality (17) of concentrations prediction at calibration (C) and test (T) sets using both accuracies: Δ_1 and Δ_2 in batch (BA) and flow injection (FIA) modes, when input data were processed with PCA.

Component, k	BA, Δ_1, $d = 0.02$		BA, Δ_2, $d = 0.02$		BA, Δ_2, $d = 0.05$		FIA, Δ_2, $d = 0.02$	
	C-set	T-set	C-set	T-set	C-set	T-set	C-set	T-set
1	100	100	100	100	99.9	100	100	100
2	100	100	100	100	99.8	99.8	100	100
3	99.76	99.6	92.56	96	100	100	100	100
4	87.97	86.95	61.16	57.4	99.95	99.9	99.85	99.75

amperometric biosensor and an artificial neural network. The CCA as well as the PCA analysis were applied for an optimization of the calibration data.

Artificial neural networks can be successfully used to discriminate components of mixtures and to estimate the concentration of each component from the biosensor response data. In the cases when biosensor operates in BA, the forecasting quality of concentrations is considerably higher when the calibration data is processed using PCA rather than CCA. In FIA, the quality of concentrations recognition is practically the same for both types of input data pre-processing: CCA and PCA.

In the case of BA, the prediction quality significantly depends on the biosensor response being under either diffusion or enzyme kinetics control. The concentration of components is predicted more accurately when the diffusion modulus is greater than unity, that is, the response is under diffusion control. Because of this, the enzyme membrane thickness as one of factors determining the diffusion modulus is of crucial importance for the detection limit of the biosensor system.

Acknowledgement

References

[1] L.C. Clarc and C. Loys, Ann. N.Y. Acad. Sci. 102 (1962) 29.
[2] A.P.F. Turner, I. Karube and G.S. Wilson, *Biosensors: Fundamentals and Applications* (Oxford University Press, Oxford, 1987).
[3] F. Scheller and F. Schubert, *Biosensors* (Elsevier, Amsterdam, 1992).
[4] U. Wollenberger, F. Lisdat and F.W. Scheller, *Frontiers in Biosensorics 2: Practical Applications* (Birkhauser Verlag, Basel, 1997).
[5] A. Chaubey and B.D. Malhotra, Biosens. Biolectron. 17 (2002) 441.
[6] C.R. Rao, *Linear Statistical Inference and its Application* (Wiley, New York, 1973).
[7] P.H. Hopke, Anal. Chim. Acta 500 (2003) 365.
[8] B. Lavine, Anal. Chem. 70 (1998) 209R.
[9] T. Artursson, T. Eklöv, I. Lundström, P. Mårtensson, M. Sjöström and M. Holmberg, J. Chemometrics 14 (2000) 711.
[10] P. Malkavaara, R. Alén and E. Kolehmainen, J. Chem. Inf. Comput. Sci. 40 (2000) 438.
[11] R. Bro, Anal. Chim. Acta 500 (2003) 185.
[12] J.A. Hertz, A.S. Krogh and R.G. Palmer, *Introduction to the Theory of Neural Computation* (Addison Wesley, Reading, MA, 1991).
[13] S. Haykin, *Neural Networks: A Comprehensive Foundation,* 2nd ed. (Prentice Hall, New York 1999).
[14] D. Patterson, *Artificial Neural Networks, Theory and Applications* (Prentice Hall, Upper Saddle River, 1996).
[15] J.R. Long, V.G. Gregoriou and P.J. Gemperline, Anal. Chem. 62 (1990) 1791.

[16] S. Hanaki, T. Nakamoto and T. Moriizuni, Sens. Actuat. B 57 (1996) 65.

[17] C. Ziegel et al., Biosens. Bioelectron. 13 (1998) 539.

[18] J.R. Llinas and J.M. Ruiz, in: *Computer Aids to Chemistry,* ed. G. Vemin and M. Chanon (John Wiley, New York, 1986).

[19] H. Martens and T. Næs, *Multivariate Calibration* (Wiley, Chichester, 1989).

[20] J.W. Gardner, Sens. Actuat. B 4 (1991) 109.

[21] C.-F. Mandenius, A. Hagman, F. Dunås, H. Sundgren and I. Lundström, Biosens. Bioelectron. 13 (1998) 193.

[22] D. Brodnjank-Vončina, D. Dobčnik, M. Novič and J. Zupan, Anal. Chim. Acta 462 (2002) 87.

[23] Y.G. Vlasov, A.V. Legin, A.M. Rudnitskaya, A.D. D'Aminco and K. Natale, Zh. Analit. Khim. 52 (1997) 1199.

[24] R.W. Harrison, J. Math. Chem. 26 (1999) 125.

[25] J. Zupan and J. Gasteiger, *Neural Networks in Chemistry and Drug Design,* 2nd ed. (Wiley-VCH, Weinheim, 1999).

[26] T.T. Bachmann et al., Biosens. Bioelectron. 15 (2000) 193.

[27] A.V. Lobanov et al., Biosens. Bioelectron. 16 (2001) 1001.

[28] S. Reder, F. Dieterle, H. Jansen, S. Alcock and G. Gauglitz, Biosens. Bioelectron. 19 (2003) 447.

[29] J. Ruzicka and E.H. Hansen, *Flow Injection Analysis* (John Wiley and Sons, New York, 1988).

[30] J. Crank, *The Mathematics of Diffusion,* 2nd ed. (Clarendon Press, Oxford, 1975).

[31] W.F. Ames, *Numerical Methods for Partial Differential Equations,* 2nd ed. (Academic Press, New York, 1977).

[32] D. Britz, *Digital Simulation in Electrochemistry,* 2nd ed. (Springer-Verlag, Berlin, 1988).

[33] T. Schulmeister, Selective Electrode Rev. 12 (1990) 260.

[34] R. Baronas, F. Ivanauskas and J. Kulys, J. Math. Chem. 32 (2002) 225.

[35] P.N. Bartlett and K.F.E. Pratt, Biosens. Bioelectron. 8 (1993) 451.

[36] K. Yokoyama and Y. Kayanuma, Anal. Chem. 70 (1998) 3368.

[37] R. Baronas, J. Christensen, F. Ivanauskas and J. Kulys, Nonlinear Anal. Modell. Control 7 (2002) 3.

[38] R. Baronas, F. Ivanauskas and J. Kulys, Sensors 3 (2003) 248.

[39] J. Principe, N. Euliano and W. Lefebvre. *Neural and Adaptive Systems: Fundamentals through Simulations* (Wiley, New York, 2000).

[40] L.W. Chan and C.C. Szeto, in: *Proceedings of the IEEE IJCNN'99,* Washington, D.C., (1999).

[41] A.F. Atiya and A.G. Parlos, IEEE Trans. Neural Networks 11 (2000) 697.