



ELSEVIER

Journal of Chromatography A, 973 (2002) 47–59

JOURNAL OF  
CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

## Optimization of artificial neural networks used for retention modelling in ion chromatography

Goran Srečnik<sup>a</sup>, Željko Debeljak<sup>a</sup>, Štefica Cerjan-Stefanović<sup>b</sup>, Milko Novič<sup>c</sup>,  
Tomislav Bolanča<sup>b,\*</sup>

<sup>a</sup>PLIVA, Pharmaceutical Industry, Prilaz baruna Filipovića 25, 10000 Zagreb, Croatia

<sup>b</sup>Laboratory of Analytical Chemistry, Faculty of Chemical Engineering and Technology, University of Zagreb, Marulićev trg 20,  
10000 Zagreb, Croatia

<sup>c</sup>National Institute of Chemistry, Hajdrihova 19, 1001 Ljubljana, Slovenia

Received 20 February 2002; received in revised form 23 July 2002; accepted 24 July 2002

### Abstract

The aim of this work is the development of an artificial neural network model, which can be generalized and used in a variety of applications for retention modelling in ion chromatography. Influences of eluent flow-rate and concentration of eluent anion ( $\text{OH}^-$ ) on separation of seven inorganic anions (fluoride, chloride, nitrite, sulfate, bromide, nitrate, and phosphate) were investigated. Parallel prediction of retention times of seven inorganic anions by using one artificial neural network was applied. MATLAB Neural Networks ToolBox was not adequate for application to retention modelling in this particular case. Therefore the authors adopted it for retention modelling by programming in MATLAB metalanguage. The following routines were written; the division of experimental data set on training and test set; selection of data for training and test set; Dixon's outlier test; retraining procedure routine; calculations of relative error. A three-layer feed forward neural network trained with a Levenberg–Marquardt batch error back propagation algorithm has been used to model ion chromatographic retention mechanisms. The advantage of applied batch training methodology is the significant increase in speed of calculation of algorithms in comparison with delta rule training methodology. The technique of experimental data selection for training set was used allowing improvement of artificial neural network prediction power. Experimental design space was divided into 8–32 subspaces depending on number of experimental data points used for training set. The number of hidden layer nodes, the number of iteration steps and the number of experimental data points used for training set were optimized. This study presents the very fast (300 iteration steps) and very accurate (relative error of 0.88%) retention model, obtained by using a small amount of experimental data (16 experimental data points in training set). This indicates that the method of choice for retention modelling in ion chromatography is the artificial neural network.

© 2002 Elsevier Science B.V. All rights reserved.

**Keywords:** Neural networks, artificial; Retention modelling; Optimization; Fluoride; Chloride; Nitrite; Sulfate; Bromide; Nitrate; Phosphate

\*Corresponding author. Tel.: +385-14-597-205; fax: +385-14-597-250.

E-mail addresses: [zdebelja@pliva.hr](mailto:zdebelja@pliva.hr) (Ž. Debeljak), [milko.novic@ki.si](mailto:milko.novic@ki.si) (M. Novič), [tbolanca@pierre.fkit.hr](mailto:tbolanca@pierre.fkit.hr) (T. Bolanča).

### 1. Introduction

Ion chromatography (IC) is a widely used technique and methods used and applications are de-

scribed in several books [1,2] and reviews [3–6]. IC analysis requires selection of ion chromatographic parameters in order to establish acceptable selectivity and reasonable retention times of late-eluting species as well as to establish better precision and accuracy and lower detection and quantification limits. The retention mechanism is controlled by several variables like stationary phase, eluent and temperature [7–12]. Retention models for ion chromatography were evaluated and only partial agreement was found between theoretical models and experimental data [13–15]. There is still much to be investigated, the problem of the selectivity of mono- and divalent anions having been neglected so far.

The development of computer-assisted retention modelling routines in chromatography can be achieved in a number of ways. In the last decade, neural networks have found widespread popularity amongst analytical chemists for solving chemical problems [16–19]. A lot of different artificial neural networks based on different concepts and aims have been currently explored. For some of these neural networks, a counterpart in statistical methods exists [20]. One of the widely used statistical techniques for retention modelling involves application of experimental design, regression analysis and statistical testing [16,17,19,21]. However, by using the neural networks modelling technique more accurate retention models can be obtained [16,17,19].

Most neural networks are trained by the so-called error backpropagation rule, also known as the generalized Delta rule that can be considered as a gradient descent procedure, which is generalized for multiple layers [22,23]. In optimization theory, it is known that the convergence of gradient descent algorithms can be accelerated using second-order information about the shape of the optimization surface [24].

However, the powerful modelling capability of neural networks causes poor models when the design of the data is not carefully chosen. The generalizing ability of a network model is affected by the data used [25,26]. Especially, when few and noisy data are used, networks are prone to overfitting. It must be stressed that this problem applies for all non-linear methods, and needs to be prevented in practice. Next, the generalizing ability of neural networks strongly depends on the representativity of the

training data with respect to the true but unknown underlying relations. The training sample must contain sufficient and well-distributed training samples in order to generalize for experimental noise. Neural networks are very sensitive to overtraining when little ill dimensioned data are used [27].

In this work the artificial neural network was used for retention modelling of anions in ion chromatography. The aim of this work is to determine the suitable artificial neural network model, which can be generalized and used in a variety of application for retention modelling in ion chromatography. The main advantage of developed ion chromatographic retention model over the others is description of whole ion chromatographic system at the same time by using one artificial neural network for modelling of retention times of all investigated anions (fluoride, chloride, nitrite, sulfate, bromide, nitrate, and phosphate) in relation with all investigated ion chromatographic conditions [eluent flow-rate and concentration of eluent anion ( $\text{OH}^-$ )]. Furthermore the feed forward neural network trained with a Levenberg–Marquardt batch error back propagation algorithm has been used, which ensures significant reduction of calculation time. The huge set of experimental data points used for testing guarantee the reliability of the proposed model. The technique of experimental data sampling strategy for training set was used allowing improvement of neural network prediction power. The number of nodes in the hidden layer, the number of iteration steps and the number of experimental data points used for training set were optimized in terms of obtaining precise and accurate retention model with respect of minimization of unnecessary experimentation and time needed for calculation procedures.

## 2. Methodology

The most common multi-layers feed forward neural network containing one hidden layer with non-linear activation nodes is briefly described. For details on multi-layers feed forward neural network, the reader is referred to comprehensive textbooks [28,29] covering most aspects of neural network design and training which have to be considered in practice. Multi-layer feed forward neural networks

are referred to as multi-purpose universal approximators. This means that given sufficient number of hidden nodes, any function can be approximated to certain degree of accuracy.

The feed-forward topology of the multi-layered neural network used in this paper is described in Fig. 1. The input of the network consists of  $K$  input variables. The  $L$  hidden nodes inputs are calculated by multiplying the input vector  $\mathbf{x}$ , with the  $K \times L$  weight matrix  $\mathbf{W}$ . The hidden nodes are activated by non-linear transfer function  $\Phi(x_i \cdot \mathbf{W})$  and multiplied with the  $L \times M$  weight matrix  $\mathbf{V}$  in order to obtain the network output as presented by Eq. (1):

$$\hat{y}_i = \Phi(x_i \cdot \mathbf{W}) \cdot \mathbf{V} \tag{1}$$

The transfer operator  $\Phi$  represents a non-linear transfer function, generally the tangents hyperbolic, the sigmoid function or any monotonically increasing function. For  $n$  training patterns  $[(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)]$ , the network is trained by minimizing the squared errors  $(y_i - \hat{y}_i)^2$  for each training sample  $(x_i, y_i)$  for  $i = [1, \dots, n]$ .

Gradient descent methods [28,29] are commonly used as a neural network training method and work by finding a minimum for an error function as given by:

$$F(w) = \sum_i (y_i - x_i \mathbf{W})^2 \tag{2}$$

by means of adapting the weights. The negative gradient [negative first derivative of Eq. (2) denoted as  $-\nabla F(w)$ ] is used as a new search direction at position  $w$  in the weight space, as given by:

$$w_{t+1} = w_t - \eta \nabla F(w)_t \tag{3}$$

The step size  $\eta$  of the new direction  $\nabla F(w)$  is referred to as the learning rate and determines the speed of convergence. The weight adaptations are performed by:

$$w_{t+1} = w_t + \eta \sum_i (y_i - x_i w)_t \cdot x_i \tag{4}$$

The minimization of error function by means of successive weight adaptations can be considered as a gradient descent procedure, using a first-order Taylor expansion:

$$F(w_{t+1}) = F(w_t + \Delta w_t) = F(w_t) + g_t \Delta w_t \tag{5}$$

where the gradient  $g = \nabla F(w_t)$  is evaluated at the previous guess  $w_t$ . In the case of a convex error surface, the Newton method can be applied in order to decrease the number of iteration steps:

$$\begin{aligned} F(w_{t+1}) &= F(w_t + \Delta w_t) \\ &= F(w_t) + \nabla F(w_t) \Delta w_t + \frac{1}{2} \nabla^2 F(w_t) \Delta w_t^2 \end{aligned} \tag{6}$$

It can be shown [30,31] that the  $N$  the Newton's weight update can be computed by:

$$w_{t+1} = w_t - A_t^{-1} g_t \tag{7}$$

where  $A^{-1}$  represents the inverse Hessian  $[\nabla^2 F(w_t)]^{-1}$ . The Hessian represents the matrix of second derivatives, as given in Eq. (8):

$$\nabla^2 F(w_t) = \begin{bmatrix} \frac{\delta^2}{\delta w_1^2} F(w_t) & \frac{\delta^2}{\delta w_1 \delta w_2} F(w_t) & \dots & \frac{\delta^2}{\delta w_1 \delta w_n} F(w_t) \\ \frac{\delta^2}{\delta w_2 \delta w_n} F(w_t) & \frac{\delta^2}{\delta w_2^2} F(w_t) & \dots & \frac{\delta^2}{\delta w_2 \delta w_n} F(w_t) \\ \dots & \dots & \dots & \dots \\ \frac{\delta^2}{\delta w_n \delta w_1} F(w_t) & \frac{\delta^2}{\delta w_n \delta w_2} F(w_t) & \dots & \frac{\delta^2}{\delta w_n^2} F(w_t) \end{bmatrix} \tag{8}$$

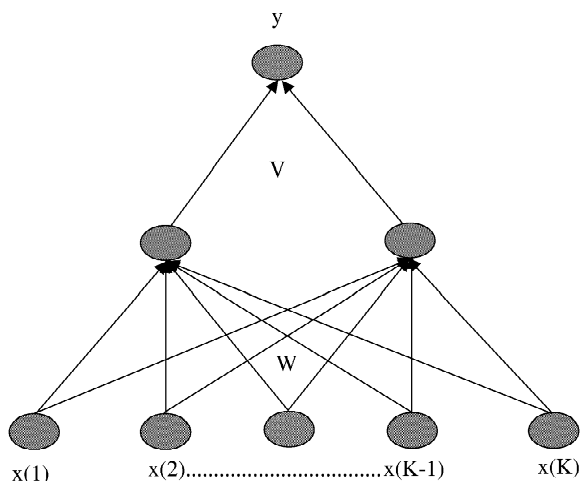


Fig. 1. A three-layer feed-forward neural network. The weight matrix connecting the input variables to the hidden layer is denoted by  $\mathbf{W}$ . The weight matrix containing the output weights is denoted by  $\mathbf{V}$ .

Starting from a favourable position, the global minimum can be found by using significantly lower number of iteration steps, which is major improvement to the iterative approximators of the gradient descent method. Unfortunately, the Newton method requires an inverse and storage of Hessian, which can be quite computational, and memory consuming. Furthermore, an additional problem will arise that multi-layered neural networks consist of more complex error surfaces. The computation of (inverse) Hessian matrix can be avoided by approximation of its properties, known as Gauss–Newton training. The weights are updated as follows:

$$w_{t+1} = w_t - \eta \cdot \frac{1}{2} \cdot \left( \frac{\delta g'_t}{\delta w} \cdot \frac{\delta g}{\delta w} \right)^{-1} g_t \quad (9)$$

Since the Hessian is approximated, the Gauss–Newton update is only accurate near the minimum and, moreover, is sensitive for divergence when training is started far from the minimum. The inverse of approximated Hessian can be stabilized by using the Levenberg–Marquardt [32] update:

$$w_{t+1} = w_t - \eta \cdot \frac{1}{2} \cdot \left( \frac{\delta g'_t}{\delta w} \cdot \frac{\delta g}{\delta w} + \tau I \right)^{-1} g_t \quad (10)$$

which is a combination of the known gradient descent and Gauss–Newton method. The parameters  $\eta$  and  $\tau$  control the behaviour of the weights updates. For a large value of  $\eta$  and  $\tau$ , gradient descent training will be performed, in contrast to Gauss–Newton behaviour for small values for  $\eta$  and  $\tau$ . Levenberg–Marquardt training has been applied successfully because the algorithm is stable and efficient.

### 3. Experimental

#### 3.1. Instrumentation

A Dionex DX500 chromatography system (Sunnyvale, CA, USA) equipped with a quaternary gradient pump (GP50), an eluent generator module (EG40), a chromatography module (LC25) and a detector module (ED40) was used in all experiments. Separation and suppressor columns used were a Dionex IonPac AG15 (50×4 mm) guard column, an IonPac AS15 (250×4 mm) separation column and an

ASRS-ULTRA-4 mm suppressor column, working in recycle mode. The sample-loop volume was 50  $\mu$ l. The void volume of the column was 2.2 ml. The eluent flow-rate was 1.00 to 1.95 ml/min and the concentration of KOH ( $\text{OH}^-$ ) in the eluent was 25.00 to 60.00 mM. The whole system was computer controlled through PeakNet 5.1 software.

The data for further evaluation were obtained by exporting the appropriate chromatograms into ASCII files. ASCII data files were further evaluated using the Microcal Origin (Microcal Software, USA) software package.

#### 3.2. Reagents and solutions

Stock solutions of fluoride (1.0000 g/l), chloride (1.0000 g/l), nitrite (1.0000 g/l), sulfate (1.0000 g/l), bromide (1.0000 g/l), nitrate (1.0000 g/l) and phosphate (1.0000 g/l) were prepared from the air-dried (at 105 °C) salts of individual anions of analytical-reagent grade (Merck, Darmstadt, Germany). An appropriate amount of individual salt was weighed into a volumetric flask (100 ml) and dissolved with Milli-Q water. Working standard solutions of fluoride (2.00 mg/l), chloride (5.00 mg/l), nitrite (10.00 mg/l), sulfate (10.00 mg/l), bromide (20.00 mg/l), nitrate (20.00 mg/l), and phosphate (30.00 mg/l) were prepared by measuring the appropriate volume of stock solution of individual anion into a 100-ml volumetric flask, which was later filled to the mark with Milli-Q water. Working eluent solutions were prepared on-line by appropriate dilution of KOH with Milli-Q water. 18 M $\Omega$  cm $^{-1}$  water (Millipore, Bedford, MA, USA) was used for dilution in all cases.

#### 3.3. Experimental design

The experimental design has been planned in order to describe the chromatographic behaviour in a multi-dimensional space: retention time versus eluent flow-rate and concentration of  $\text{OH}^-$  in eluent. The eluent flow-rate was varied in range from 1.00 to 1.95 ml/min and concentration of  $\text{OH}^-$  in the eluent was varied from 25.00 to 60.00 mM. One hundred and twenty eight experimental data were obtained. The experimental data used for the modelling procedure is presented in Table 1.

Table 1  
Experimental data set used for retention modelling procedure

Eluent flow-rate (ml/min)	$c(\text{OH}^-)$ in eluent (mM)	$\text{F}^-$ (min)	$\text{Cl}^-$ (min)	$\text{NO}_2^-$ (min)	$\text{SO}_4^{2-}$ (min)	$\text{Br}^-$ (min)	$\text{NO}_3^-$ (min)	$\text{PO}_4^{3-}$ (min)
1.05	25.00	5.70	10.17	12.43	22.10	22.85	25.25	70.60
1.15	25.00	5.38	9.45	11.52	20.37	20.78	23.22	64.05
1.25	25.00	5.12	8.85	10.75	18.88	19.15	21.47	58.60
1.35	25.00	4.90	8.33	10.08	17.55	17.93	19.95	54.40
1.45	25.00	4.70	7.90	9.52	16.40	16.97	18.62	50.88
1.55	25.00	4.53	7.52	9.02	15.43	15.93	17.48	47.53
1.65	25.00	4.38	7.17	8.58	14.62	14.92	16.52	44.23
1.75	25.00	4.20	6.78	8.08	13.62	14.15	15.42	41.13
1.85	25.00	4.08	6.50	7.73	13.10	13.10	14.65	38.05
1.95	25.00	3.97	6.27	7.42	12.52	12.52	13.97	36.18
1.10	30.00	5.30	9.07	11.00	16.95	19.32	21.97	45.82
1.20	30.00	5.03	8.47	10.23	15.60	17.85	20.27	41.70
1.30	30.00	4.80	7.97	9.58	14.48	16.60	18.82	38.33
1.40	30.00	4.60	7.53	9.03	13.48	15.55	17.62	35.45
1.50	30.00	4.38	7.00	8.35	12.52	14.07	15.90	31.87
1.60	30.00	4.28	6.83	8.15	12.08	13.82	15.60	31.27
1.70	30.00	4.12	6.48	7.68	11.42	12.88	14.55	29.07
1.80	30.00	3.98	6.22	7.37	10.82	12.27	13.85	27.28
1.90	30.00	3.88	5.98	7.07	10.28	11.72	13.20	25.72
1.05	32.00	5.33	8.85	10.63	15.38	18.27	20.67	37.98
1.15	32.00	5.05	8.25	9.87	14.12	16.85	19.03	34.53
1.25	32.00	4.82	7.75	9.23	13.08	15.65	17.65	31.68
1.35	32.00	4.60	7.32	8.70	12.27	14.60	16.45	29.47
1.45	32.00	4.43	6.95	8.23	11.63	13.67	15.37	27.72
1.55	32.00	4.28	6.63	7.82	11.02	12.87	14.43	26.02
1.65	32.00	4.15	6.35	7.45	10.37	12.22	13.67	24.28
1.75	32.00	3.98	6.02	7.05	9.80	11.43	12.82	22.57
1.85	32.00	3.87	5.78	6.75	9.28	10.83	12.13	21.20
1.95	32.00	3.78	5.58	6.50	8.85	10.43	11.65	20.05
1.00	34.00	5.45	9.13	11.03	15.23	19.20	21.27	37.15
1.10	34.00	5.15	8.48	10.18	13.97	17.58	19.90	33.53
1.20	34.00	4.88	7.93	9.50	12.88	16.27	18.37	30.53
1.30	34.00	4.67	7.47	8.90	12.02	15.12	17.05	28.18
1.40	34.00	4.47	7.07	8.40	11.22	14.20	15.98	26.02
1.50	34.00	4.27	6.58	7.78	10.43	12.87	14.47	23.53
1.60	34.00	4.17	6.43	7.60	10.10	12.63	14.18	23.08
1.70	34.00	4.02	6.10	7.18	9.55	11.80	13.27	21.45
1.80	34.00	3.90	5.87	6.88	9.08	11.25	12.62	20.25
1.90	34.00	3.80	5.67	6.62	8.68	10.73	12.02	19.17
1.05	36.00	5.18	8.32	9.92	12.95	16.73	18.85	28.58
1.15	36.00	4.92	7.77	9.23	11.92	15.47	17.38	26.02
1.25	36.00	4.68	7.32	8.65	11.08	14.37	16.13	23.93
1.35	36.00	4.50	6.92	8.15	10.40	13.43	15.05	22.25
1.45	36.00	4.33	6.58	7.72	9.87	12.60	14.08	20.97
1.55	36.00	4.18	6.28	7.33	9.38	11.85	13.23	19.75
1.65	36.00	4.05	6.02	7.00	8.90	11.25	12.53	18.57
1.75	36.00	3.90	5.72	6.63	8.38	10.57	11.78	17.23
1.85	36.00	3.80	5.50	6.37	8.00	10.03	11.17	16.23
1.95	36.00	3.70	5.32	6.13	7.63	9.65	10.73	15.35
1.00	38.00	5.32	8.62	10.33	12.97	17.67	19.97	28.28

Table 1. Continued

Eluent flow-rate (ml/min)	$c(\text{OH}^-)$ in eluent (mM)	$\text{F}^-$ (min)	$\text{Cl}^-$ (min)	$\text{NO}_2^-$ (min)	$\text{SO}_4^{2-}$ (min)	$\text{Br}^-$ (min)	$\text{NO}_3^-$ (min)	$\text{PO}_4^{3-}$ (min)
1.10	38.00	5.02	8.02	9.55	11.92	16.18	18.25	25.62
1.20	38.00	4.77	7.50	8.92	11.03	15.00	16.88	23.37
1.30	38.00	4.57	7.08	8.37	10.32	13.97	15.68	21.65
1.40	38.00	4.38	6.72	7.92	9.67	13.12	14.70	20.07
1.50	38.00	4.18	6.27	7.33	9.02	11.92	13.33	18.22
1.60	38.00	4.08	6.13	7.17	8.72	11.70	13.08	17.80
1.70	38.00	3.93	5.82	6.78	8.27	10.95	12.25	16.65
1.80	38.00	3.82	5.60	6.50	7.88	10.42	11.65	15.70
1.90	38.00	3.80	5.42	6.27	7.43	9.97	11.10	14.70
1.05	40.00	5.07	7.92	9.35	11.22	15.52	17.42	22.32
1.15	40.00	4.82	7.40	8.72	10.37	14.35	16.07	20.42
1.25	40.00	4.60	6.97	8.17	9.68	13.35	14.92	18.85
1.35	40.00	4.40	6.60	7.72	9.08	12.50	13.95	17.52
1.45	40.00	4.25	6.28	7.32	8.63	11.73	13.07	16.52
1.55	40.00	4.12	6.00	6.97	8.23	11.03	12.27	15.63
1.65	40.00	3.98	5.77	6.65	7.83	10.47	11.62	14.73
1.75	40.00	3.83	5.48	6.30	7.38	9.87	10.95	13.67
1.85	40.00	3.73	5.28	6.05	7.07	9.37	10.38	12.93
1.95	40.00	3.63	5.10	5.83	6.75	9.02	9.98	12.23
1.00	42.00	5.20	8.20	9.77	11.35	16.43	18.47	22.43
1.10	42.00	4.92	7.63	9.05	10.45	15.10	16.93	20.33
1.20	42.00	4.68	7.17	8.45	9.70	14.00	15.68	18.62
1.30	42.00	4.48	6.77	7.95	9.08	13.03	14.58	17.23
1.40	42.00	4.30	6.42	7.52	8.57	12.23	13.65	16.10
1.50	42.00	4.10	6.02	6.98	7.98	11.15	12.43	14.62
1.60	42.00	4.02	5.87	6.82	7.73	10.95	12.18	14.27
1.70	42.00	3.87	5.58	6.47	7.35	10.27	11.42	13.40
1.80	42.00	3.77	5.38	6.22	7.03	9.78	10.87	12.70
1.90	42.00	3.67	5.20	5.97	6.77	9.32	10.35	12.10
1.05	44.00	4.98	7.57	8.88	9.95	14.53	16.23	18.00
1.15	44.00	4.73	7.10	8.30	9.25	13.45	15.00	16.55
1.25	44.00	4.52	6.68	7.78	8.67	12.50	13.92	15.37
1.35	44.00	4.33	6.33	7.35	8.13	11.73	13.03	14.27
1.45	44.00	4.18	6.03	6.98	7.72	11.02	12.22	13.47
1.55	44.00	4.05	5.78	6.65	7.40	10.37	11.47	12.82
1.65	44.00	3.92	5.55	6.37	7.05	9.85	10.88	12.10
1.75	44.00	3.77	5.28	6.03	6.65	9.28	10.27	11.22
1.85	44.00	3.68	5.10	5.80	6.38	8.83	9.75	10.67
1.95	44.00	3.58	4.93	5.60	6.13	8.50	9.37	10.15
1.10	46.00	4.83	7.20	8.40	9.08	13.55	15.10	15.60
1.20	46.00	4.60	6.90	8.07	8.72	13.15	14.67	15.18
1.30	46.00	4.40	6.52	7.60	8.18	12.27	13.65	14.10
1.40	46.00	4.23	6.18	7.18	7.73	11.48	12.77	13.25
1.50	46.00	4.05	5.80	6.70	7.23	10.55	11.70	12.10
1.60	46.00	3.97	5.67	6.55	7.00	10.33	11.43	11.70
1.70	46.00	3.82	5.38	6.20	6.65	9.68	10.73	11.03
1.80	46.00	3.72	5.20	5.97	6.40	9.23	10.22	10.47
1.90	46.00	3.63	5.03	5.73	6.15	8.82	9.73	10.03
1.00	50.00	5.03	7.60	8.92	9.27	14.55	15.55	16.27
1.10	50.00	4.77	7.08	8.28	8.52	13.40	14.02	14.98

Table 1. Continued

Eluent flow-rate (ml/min)	c(OH <sup>-</sup> ) in eluent (mM)	F <sup>-</sup> (min)	Cl <sup>-</sup> (min)	NO <sub>2</sub> <sup>-</sup> (min)	SO <sub>4</sub> <sup>2-</sup> (min)	Br <sup>-</sup> (min)	NO <sub>3</sub> <sup>-</sup> (min)	PO <sub>4</sub> <sup>3-</sup> (min)
1.20	50.00	4.53	6.65	7.75	7.95	12.43	12.88	13.87
1.30	50.00	4.35	6.30	7.30	7.48	11.62	11.98	12.93
1.40	50.00	4.18	5.98	6.90	7.10	10.87	11.30	12.07
1.50	50.00	4.00	5.62	6.45	6.62	10.00	10.28	11.10
1.60	50.00	3.92	5.50	6.32	6.43	9.78	10.03	10.83
1.70	50.00	3.77	5.23	5.98	6.13	9.20	9.48	10.20
1.80	50.00	3.68	5.05	5.75	5.88	8.78	9.00	9.70
1.00	55.00	4.62	4.97	7.30	8.47	12.95	13.68	15.23
1.10	55.00	4.70	6.82	7.77	7.88	11.72	12.65	14.03
1.20	55.00	4.48	6.42	7.27	7.40	10.82	11.75	13.02
1.30	55.00	4.28	6.07	6.83	6.97	10.07	10.97	12.13
1.40	55.00	4.13	5.78	6.58	6.58	9.55	10.27	11.33
1.50	55.00	3.95	5.43	6.17	6.17	8.73	9.47	10.43
1.60	55.00	3.87	5.32	5.93	6.03	8.55	9.25	10.18
1.05	58.00	4.77	6.78	7.47	7.80	10.62	12.22	13.50
1.15	58.00	4.55	6.37	6.98	7.30	9.78	11.30	12.47
1.25	58.00	4.33	6.02	6.60	6.87	9.23	10.52	11.57
1.35	58.00	4.17	5.72	6.22	6.50	8.60	9.92	10.88
1.45	58.00	4.02	5.47	5.93	6.18	8.15	9.33	10.23
1.55	58.00	3.90	5.23	5.70	5.90	7.80	8.80	9.63
1.65	58.00	3.75	5.00	5.45	5.62	7.42	8.30	9.08
1.00	60.00	4.90	7.05	7.73	8.17	11.03	12.95	14.33
1.10	60.00	4.63	6.60	7.17	7.62	10.10	11.98	13.27
1.20	60.00	4.42	6.22	6.72	7.13	9.33	11.15	12.30
1.30	60.00	4.25	5.88	6.33	6.73	8.70	10.42	11.47
1.40	60.00	4.08	5.62	6.05	6.38	8.28	9.77	10.73
1.50	60.00	3.90	5.28	5.67	5.98	7.65	9.02	9.90
1.60	60.00	3.83	5.17	5.53	5.85	7.47	8.80	9.63

Dependence of retention times of analysed ions (fluoride, chloride, nitrite, sulfate, bromide, nitrate, and phosphate) on ion chromatographic analysis conditions (eluent flow-rate and concentration of KOH in eluent).

It is preferable for the each experimental data point to have equal influence on the neural network model, if one wants the training and testing set to be a representative group of data of the whole design area. For that purpose the design space was divided into 8–32 equidistant subspaces depending on number of experimental data points used for the training set. It means that an equal number of experimental data points lie in each subspace of the whole design area. From each subspace one experimental data point was chosen for the training set using a random function (total 8–32 experimental data points). The rest of the experimental data points (total 96–120 experimental data points) were used for validation procedures. The input data of retention times are logarithmically ( $\log_{10}$ ) transformed before modelling, because of the heteroscedastic nature of the

retention time variance. This provides the homogeneous variance in the output of the network. Fig. 2 presents the experimental design model where 16 experimental data points used for the training set of optimized model are marked.

### 3.4. Neural networks

The neural network used in this paper was the three-layer feed forward backpropagation neural network. The input layer consists of the two nodes representing eluent flow-rate and concentration of OH<sup>-</sup> in eluent. The output layer consists of seven nodes representing the retention times of seven inorganic anions (fluoride, chloride, nitrite, sulfate, bromide, nitrate and phosphate). That means that parallel prediction of the retention times of the seven

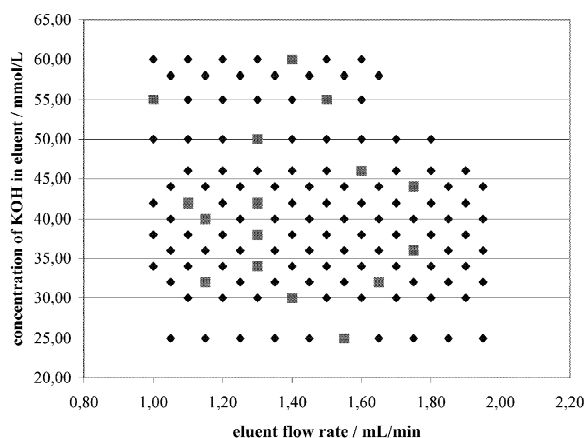


Fig. 2. Design of 128 experimental data points: concentration of  $\text{OH}^-$  in eluent vs. eluent flow-rate. Sixteen experimental data points used for the final ANN model are marked.

inorganic anions by using one artificial neural network was applied, one of the advantages of the developed retention model. The number of nodes in the hidden layer, the number of iteration steps and the number of experimental data points used for the training set need to be optimized. Therefore the number of nodes in the hidden layer was varied from 1 to 6, the number of iteration steps was varied from 100 to 1100 and the number of experimental data points used for training set was varied from 8 to 32. The optimal neural network model was determined by employing the calculation of prediction quality with an external test set consisting of 96–120 experimental data during training. The optimization of ANN shows that the small number of experimental data points used for the training set and the small number of iteration steps, will provide good accuracy of retention model, with average relative errors below 1%.

Two different transfer functions were used for the retention modelling procedure with neural networks. The hyperbolic tangent transfer function was used for computation of hidden layer nodes activities:

$$\theta(\mathbf{x}_i^* \mathbf{W}) = \frac{1 - e^{-\mathbf{x}_i^* \mathbf{W}}}{1 + e^{-\mathbf{x}_i^* \mathbf{W}}} \quad (11)$$

where  $\theta$  represents the transfer function,  $\mathbf{x}_i$  represents the input vector, and  $\mathbf{W}$  represents the weight

vector. For computation of output activities linear transfer function was employed:

$$\hat{y}_i = \Phi(\mathbf{x}_i, \mathbf{W}) \cdot \mathbf{V} \quad (12)$$

The Levenberg–Marquardt batch learning procedures using momentum were applied, and the number of iteration steps used for training procedures was optimized leading to the adequate description of retention behaviour. The advantage of applied batch training methodology is the significant increase in speed of calculation algorithm in comparison with the delta rule training methodology described in other studies [16–19].

The testing procedures involve the application of Dixon's outlier test. There are certain problems regarding Levenberg–Marquardt training optimization procedures. Some values of randomly chosen weights could cause some problems connected with matrix inversion resulting in earlier termination of the training procedure. In that case the training procedure was repeated. Dixon's outlier test was used for exclusion of outliers, during the optimization process, followed by a retraining procedure using different neural network parameters (initial weights vector, momentum factor). The purpose of Dixon's test is to obtain better generalization of the retention mechanism. The results of the optimization procedures of the neural network are shown in Figs. 3–5.

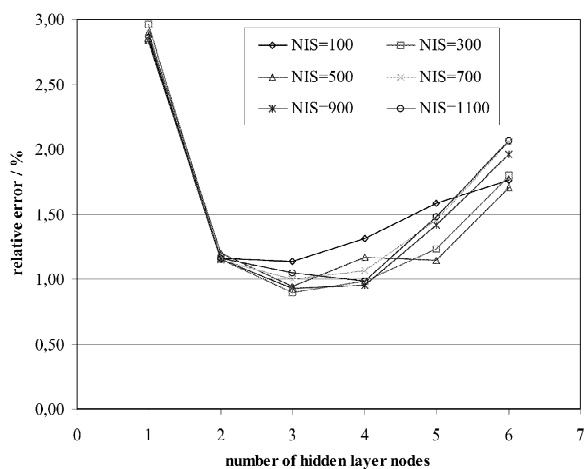
To test the predictive performance of the developed artificial neural network retention model an independent test set (96–112 experimental data points) was used to calculate relative error. The following equations were used for calculations:

$$\text{relative error} = \left( \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \frac{|y_{ij} - \hat{y}_{ij}|}{\hat{y}_{ij}} \right) \times 100\%$$

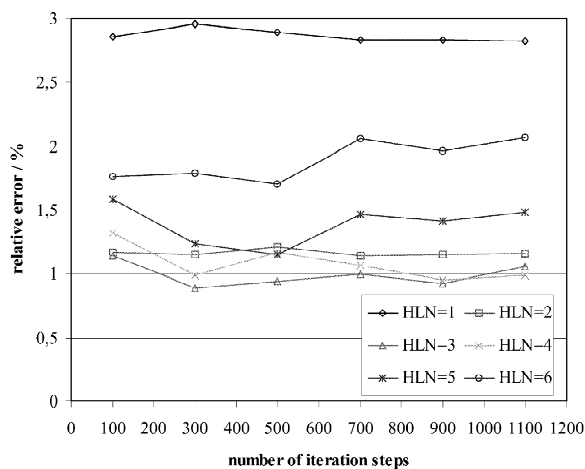
where  $y_{ij}$  represents the measured value of the neural network output (retention times of particular ion),  $\hat{y}_{ij}$  represents predicted value of neural network output,  $N$  represents the number of experimental data points used for the test set and  $M$  represents the number of retraining steps. All relative errors were calculated on non-logarithmically transformed data. The results of optimization procedures of neural network are shown in Figs. 3–5.

The program for neural network was made by the





(a)

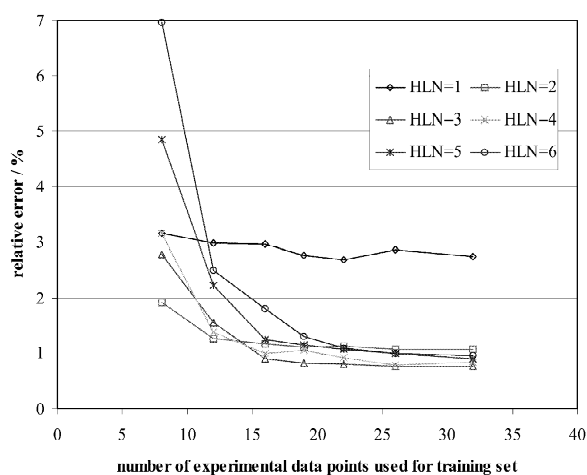


(b)

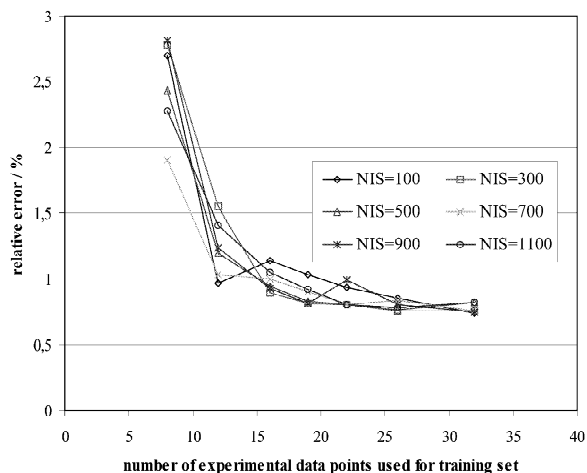
Fig. 3. (A) Relative error against number of hidden layer nodes obtained by using different number of iteration steps (NIS). (B) Relative error against number of iteration steps obtained by using different number of hidden layer nodes (HLNs).

authors in the MATLAB environment (MATLAB 6.0, MathWorks, Sherborn, MA, USA) by using the MATLAB Neural Networks ToolBox. The MATLAB Neural Networks ToolBox was not adequate for application to retention modelling in this particular case. Therefore the authors adopted it for retention modelling by programming in MATLAB metalanguage. The following routines were written:

- The division of experimental data set on training and test set.



(a)

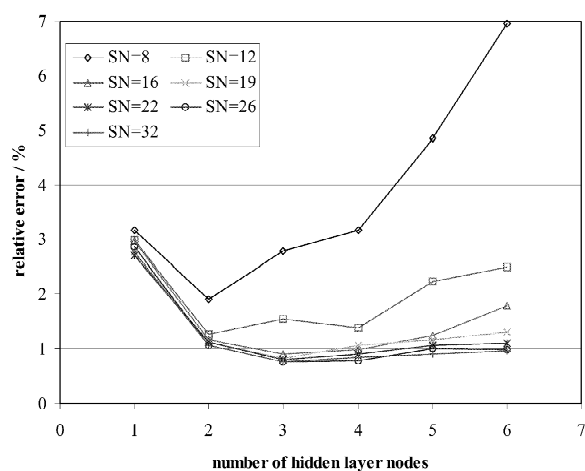


(b)

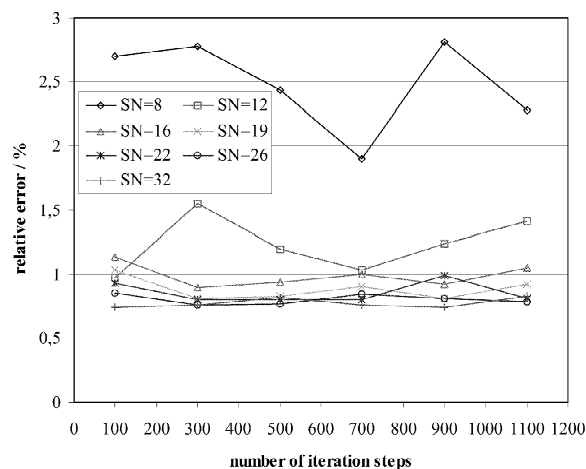
Fig. 4. (A) Relative error against number of number of experimental data points used for training set obtained by using different number of hidden layer nodes (HLNs). (B) Relative error against number of number of experimental data points used for training set obtained by using different number of iteration steps (NIS).

- Selection of data for training and test set.
- Dixon's outlier test.
- Calculations of relative error.

All other MATLAB Neural Network ToolBox options used in the described ANN retention model were adopted for the needs of retention modelling. The calculations were performed on an IBM compat-



(a)



(b)

Fig. 5. (A) Relative error against number of hidden layer nodes obtained by using different experimental data points used for training set (SN). (B) Relative error against number of iteration steps obtained by using different number of experimental data points used for training set (SN).

ible personal computer equipped with an 800 MHz Pentium III processor, and 512 Mb RAM.

#### 4. Results and discussion

The results of optimization procedures, using 128 experimental data points for the modelling procedure

(Table 1) are shown in Figs. 3–5 and are discussed as follows.

From Fig. 3A it can be seen that minimal relative error of proposed neural network retention model is obtained by using three hidden layer nodes and 300 iteration steps. Those results can be proved by examining Fig. 3B. Fig. 3B clearly indicates that the optimal number of hidden layer nodes is three in the whole domain of investigated numbers of iteration steps and the smallest relative error is obtained by using 300 iteration steps. It is also shown that the number of iteration steps has a low influence on relative error, particularly in comparison with the influence of the number of hidden layer nodes on relative error. However, one of the important factors which has to be considered is the reduction of computation time. If the number of iteration steps is higher the computation time is longer. That fact only confirms the previous conclusion that the optimal number of iteration steps is 300.

From Fig. 4A and B it can be seen that the number of experimental data points used for the training set has a significant influence on the selection of optimal parameters (number of hidden layer nodes and number of iteration steps) for the artificial neural network retention model. Reducing the number of experimental data points used for the training set is crucial for the development of the retention model without losing time on unnecessary experimentation. It is also important that the small number of experimental points in training set do not decrease the predictive ability of the retention model. For that reason it is important to determine the optimal number of experimental data points used for the training set. From Fig. 4A and B it can be seen that relative error has higher values if a small number of experimental data points are used for the training set (8, 12, 16 experimental data points). When using more experimental data points for the training set (16, 19, 22, 26, 32 experimental data points) the relative error is smaller but nearly the same value. It can be concluded that the optimal number of experimental data points used for the training set is 16.

From Fig. 5A it can be seen that the optimal number of hidden layer nodes is three when using 16, 19, 22, 26 or 32 experimental data points for the training set. If 8 or 12 experimental data points are used for the training set the optimal number of

hidden layer nodes is two, but the relative error in that case is significantly higher than in the case of using 16 to 32 experimental data points.

From Fig. 5B it can be seen that the optimal number of iteration steps is 300 when using 16, 19, 22, 26 or 32 experimental data points for the training set. If 8 or 12 experimental data points are used for the training set the optimal number of iteration steps is 700. Relative error in that case is significantly higher than in the case of using 16 to 32 experimental data points. By increasing the number of iteration steps over 300, relative error is slightly higher and the computation time is longer.

The trained neural network retention model was used to predict retention times for all seven anions (fluoride, chloride, nitrite, sulfate, bromide, nitrate, and phosphate) at all 128 ion chromatography conditions (eluent flow-rate, concentration of  $\text{OH}^-$  in eluent). Simulated retention times were plotted against experimentally measured retention times (Fig. 6). It can be seen that correlation coefficient is in range of 0.9830 to 0.9977. From those results it can be concluded that the proposed neural network retention model generalizes data well and that it can be used for retention modelling.

Results shown in Table 1 indicate that the selectivity of ion chromatographic methods strongly depends on the applied ion chromatographic conditions [eluent flow-rate, concentration of  $\text{KOH}$  ( $\text{OH}^-$ ) in eluent]. By adjusting the eluent flow-rate and concentration of ( $\text{OH}^-$ ) in eluent it is possible to increase selectivity. That is a crucial factor for numerous different applications of ion chromatography analysis, particularly for wastewater analysis and the analysis of samples with great differences in concentration of analyte components. The possibility of adjustment of the retention times of fluoride and chloride is crucial for the determination of organic acids (acetate, formate and propionate) and the possibility of adjustment of the retention times of nitrite and sulfate is crucial for the determination of butyrate and carbonate. By adjusting the retention times of late eluting anions (phosphate and nitrate) it is possible to obtain shorter ion chromatographic run and speed up analysis, without decreasing the selectivity of fast eluting anions (fluoride and chloride). The developed artificial neural network retention model enables one both to improve performance

characteristic of applied method and to speed up the new method development by reducing unnecessary experimentation.

## 5. Conclusions

In this work, artificial neural networks were used for the retention modelling of anion separation in ion chromatography. Parallel prediction of retention times of seven inorganic anions by using one artificial neural network was applied. The MATLAB Neural Networks ToolBox was not adequate for application to retention modelling in this particular case. Therefore the authors adopted it for retention modelling by programming in MATLAB metalanguage. The following routines were written; the division of experimental data set on training and test set; selection of data for training and test set; Dixon's outlier test; calculations of relative error.

Problems arise when validation of the predictive ability with an external test set during training is applied to neural network models trained with the Levenberg–Marquardt batch error back propagation algorithm, due to the fact that weight initializations yield irreproducible neural network retention models. Every new initialisation can be regarded as a new start position for the Levenberg–Marquardt batch error back propagation algorithm search for the global minimum. Although special learning parameters (e.g., momentum factor) can help to avoid local minima, no guarantee of finding the global minimum can be given. The probability of finding the global minimum was enhanced by selecting various random start positions for the Levenberg–Marquardt batch error back propagation algorithm search. Consequently, there is a bigger chance of avoiding the local minima. Obviously, the chance of finding the global minimum directly depends on the smoothness of the error hyperplane (Figs. 3–5) and the number of local minima. Validation of the predictive ability with an external test set during training by means of reinitializations and retraining the networks using Dixon's filter as a criterion, provides establishment of a better neural network retention model.

In this work the number of hidden layer nodes, the number of iteration steps and the number of experimental data points used for training set are

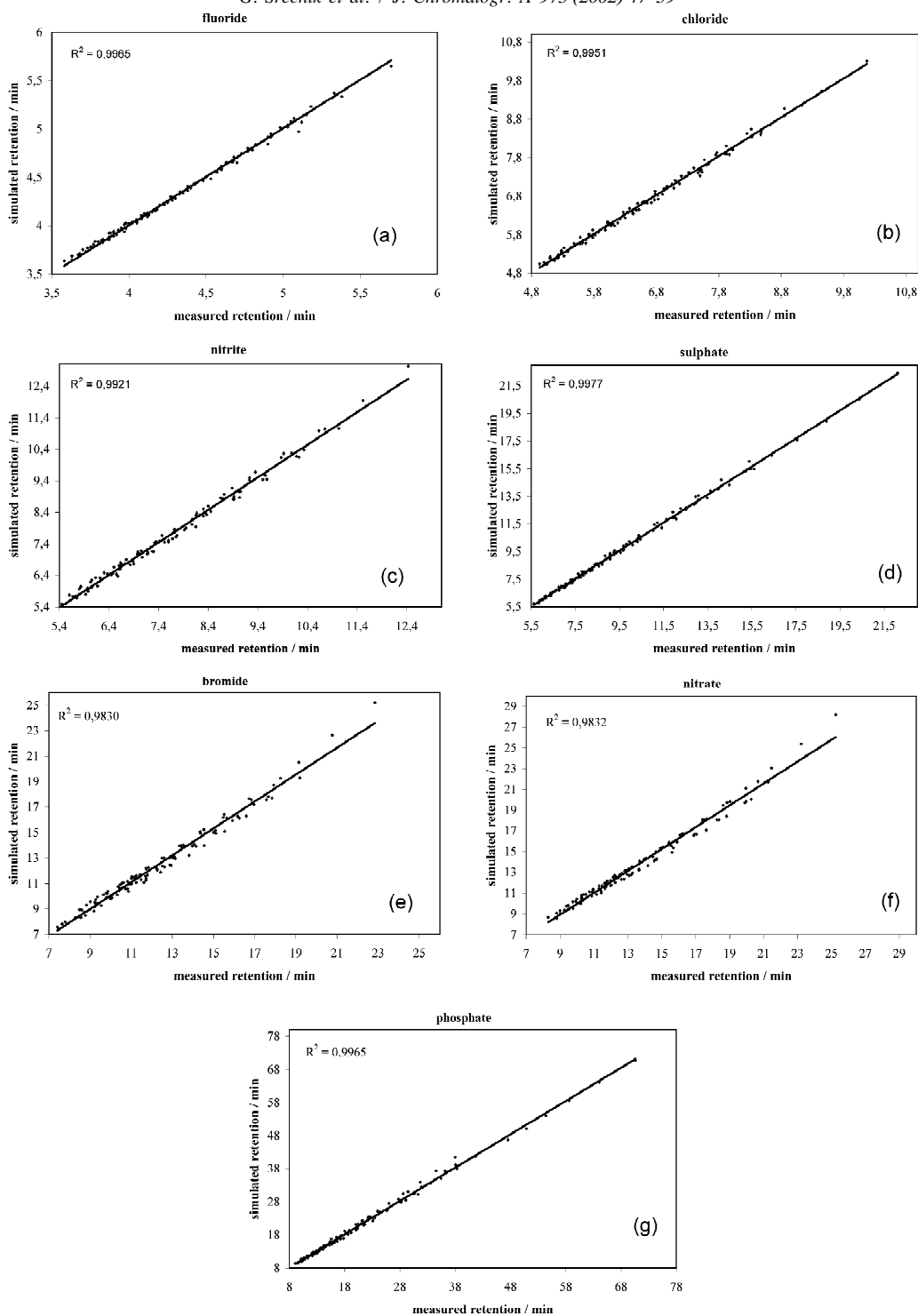


Fig. 6. Plot of experimental measured retention times versus simulated retention times using neural network model with three hidden layer nodes trained with 300 iteration steps and 16 experimental data points: (A) fluoride, (B) chloride, (C) nitrite, (D) sulfate, (E) bromide, (F) nitrate, (G) phosphate.

optimized in order to obtain a neural network model with good predictability with respect to speeding up retention modelling procedure by reducing the unnecessary experimental work and computation procedure. The optimal number of hidden layer nodes of the developed neural networks model retention models is three, the optimal number of iteration steps is 300 and the optimal number of experimental data points used for training set is 16.

The trained neural network retention model was used to predict the retention times for all seven anions (fluoride, chloride, nitrite, sulfate, bromide, nitrate, and phosphate) at all 128 ion chromatography conditions (eluent flow-rate, concentration of  $\text{OH}^-$  in eluent). It can be seen that correlation coefficient is in range of 0.9830 to 0.9977. From that results it can be concluded that the proposed neural network retention model generalizes data well and that it can be used for retention modelling.

It is shown that selectivity of ion chromatographic methods strongly depends on applied ion chromatographic conditions (eluent flow-rate, concentration of  $\text{OH}^-$  in eluent). The developed retention model allows manipulation of the appearance of a particular peak on the chromatogram and allows improvement of selectivity between particular anions. By using this retention model it is possible both to improve the performance characteristics of the applied method and to speed up the new method development by reducing unnecessary experimentation.

Finally, this study shows that optimized artificial neural networks are a very accurate and fast retention modelling tool with a small amount of experimental data needed to model varied inherent non-linear relationships of retention behaviour with respect to the mobile phase parameters. This indicates that the method of choice for retention modelling in ion chromatography is the artificial neural network.

## References

[1] D.T. Gjerde, J.S. Fritz, *Ion Chromatography*, 2nd ed., Hüthig, New York, 1987.

- [2] J. Weiss, *Ion Chromatography*, 2nd ed., VCH, Weinheim, New York, Basel, Cambridge, Tokyo, 1995.
- [3] G.D. Franklin, *Int. Lab.* 15 (1985) 56.
- [4] P.R. Haddad, A.L. Heckenberg, *J. Chromatogr.* 300 (1984) 357.
- [5] H. Small, *J. Chromatogr.* 546 (1990) 3.
- [6] Š. Cerjan-Stefanović, T. Bolanča, L. Čurković, *J. Chromatogr. A* 918 (2001) 325.
- [7] R.E. Barron, J.S. Fritz, *J. Chromatogr.* 284 (1984) 13.
- [8] F. Vlačil, I. Vinš, *J. Chromatogr.* 26 (1987) 133.
- [9] R.W. Slingsby, Ch.A. Pohl, *J. Chromatogr.* 485 (1988) 241.
- [10] L. Dwarth, J.S. Fritz, *J. Chromatogr. Sci.* 29 (1988) 241.
- [11] D.L. DuVal, J.S. Fritz, *J. Chromatogr.* 295 (1984) 89.
- [12] Š. Cerjan-Stefanović, T. Bolanča, L. Čurković, *J. Liq. Chromatogr. Rel. Technol.* 23 (2000) 2169.
- [13] J.E. Madden, P.R. Haddad, *J. Chromatogr. A* 829 (1998) 65.
- [14] J.E. Madden, P.R. Haddad, *J. Chromatogr. A* 850 (1999) 29.
- [15] J.E. Madden, N. Avdalović, P.E. Jackson, P.R. Haddad, *J. Chromatogr. A* 855 (1999) 65.
- [16] G. Sacchero, M.C. Buzzoniti, C. Sarzanini, E. Mentasti, H.J. Metting, P.M.J. Coenegracht, *J. Chromatogr. A* 799 (1998) 35.
- [17] J. Havel, J.E. Madden, P.R. Haddad, *Chromatographia* 49 (1999) 481.
- [18] J.E. Madden, N. Avdalović, P.R. Haddad, J. Havel, *J. Chromatogr. A* 910 (2001) 173.
- [19] A. Yan, G. Jiao, Z. Hu, B.T. Fan, *Comput. Chem.* 24 (2000) 171.
- [20] B. Cheng, D.M. Titterington, *Statistical Sci.* 9 (1) (1994) 2.
- [21] M.N. Hasan, P.C. Jurs, *Anal. Chem.* 60 (1988) 978.
- [22] J. Smits, W.J. Melssen, L.M.C. Buydens, G. Kateman, *Chemom. Intell. Lab. Syst.* 22 (1994) 165.
- [23] B.J. Wyhoff, *Chemom. Intell. Lab. Syst.* 18 (1993) 115.
- [24] P. Devika, L. Achenie, *J. Intell. Fuzzy Syst.* 3 (1995) 287.
- [25] D.M. Allen, *Technometrics* 16 (1974) 125.
- [26] S. Geisser, *J. Am. Stat. Assoc.* 70 (1975) 320.
- [27] R. Tibshirani, *Neural Comput.* 8 (1996) 152.
- [28] M. Smith, *Neural Networks for Statistical Modeling*, Van Nostrand Reinhold, New York, 1993.
- [29] J. Zupan, J. Gasteiger, *Neural Networks in Chemistry and Drug Design*, 2nd ed., Wiley-VCH, Weinheim, 1999.
- [30] J. Tetteh, E. Metcalfe, S. Howells, *Chem. Intell. Lab. Syst.* 32 (1996) 177.
- [31] R. Tibshirani, *Neural Comput.* 8 (1996) 152.
- [32] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in Pascal. The Art Scientific Computing*, Cambridge University Press, New York, 1992.