Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Chromatography A, 1085 (2005) 74–85

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Development of an inorganic cations retention model in ion chromatography by means of artificial neural networks with different two-phase training algorithms

Tomislav Bolanča[a,*], Štefica Cerjan-Stefanović[a], Melita Regelja[a],
Hrvoje Regelja[b], Sven Lončarić[c]

[a] *University of Zagreb, Faculty of Chemical Engineering and Technology, Laboratory of Analytical Chemistry, Marulićev trg 20, 10000 Zagreb, Croatia*

[b] *Helix Ltd., IX Južna Obala 18, 10000 Zagreb, Croatia*

[c] *University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Electronic Systems and Information Processing, Unska 3, 10000 Zagreb, Croatia*

Available online 23 February 2005

Abstract

This paper describes development of artificial neural network (ANN) retention model, which can be used for method development in variety of ion chromatographic applications. By using developed retention model it is possible both to improve performance characteristic of developed method and to speed up new method development by reducing unnecessary experimentation. Multilayered feed forward neural network has been used to model retention behaviour of void peak, lithium, sodium, ammonium, potassium, magnesium, calcium, strontium and barium in relation with the eluent flow rate and concentration of methasulphonic acid (MSA) in eluent. The probability of finding the global minimum and fast convergence at the same time were enhanced by applying a two-phase training procedure. The developed two-phase training procedure consists of both first and second order training. Several training algorithms were applied and compared, namely: back propagation (BP), delta-bar-delta, quick propagation, conjugate gradient, quasi Newton and Levenberg–Marquardt. It is shown that the optimized two-phase training procedure enables fast convergence and avoids problems arisen from the fact that every new weight initialization can be regarded as a new starting position and yield irreproducible neural network if only second order training is applied. Activation function, number of hidden layer neurons and number of experimental data points used for training set were optimized in order to insure good predictive ability with respect to speeding up retention modelling procedure by reducing unnecessary experimental work. The predictive ability of optimized neural networks retention model was tested by using several statistical tests. This study shows that developed artificial neural network are very accurate and fast retention modelling tool applied to model varied inherent non-linear relationship of retention behaviour with respect to mobile phase parameters.

Keywords: Artificial neural networks; Retention modelling; Inorganic cations; Ion chromatography

1. Introduction

Ion chromatography (IC) has become a routine analytical method and extensive literature, including books [1,2] and reviews [3–5], has been published describing methods and applications of ion chromatography in various fields. Selectivity plays the principal role in IC and can be modulated as a function of the nature of analytes, stationary phases, eluents (composition, eluent flow rate) and temperature at which separation occurs. Because the properties of the eluent and temperature on which separation occurs can be varied more easily optimization of separation is usually achieved by changes in those parameters after the selection of an appropriate stationary phase. The most efficient optimization methodologies in

[*] Corresponding author. Tel.: +385 1 4597205; fax: +385 1 4597250.
E-mail addresses: tomislav.bolanca@fkit.hr (T. Bolanča), hregelja@helix.hr (H. Regelja), sven.loncaric@fer.hr (S. Lončarić).

0021-9673/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2005.02.018

ion chromatography are based on the prediction of the elution behaviour through retention models. Several theoretical retention models for anions in ion chromatography were developed [6–15] and compared [16–18]. An effective separation of many metal cations is possible only in the presence of complexing agents forming kinetically labile complexes with the separated ions. The retention and separation of inorganic anions as well as metal complexes may be influenced by employing dissociation/protonation equilibria of analyte species and eluent ions [19]. There is still much to be investigated in the field of alkaline, earth alkaline cations and heavy metals retention modeling.

Artificial intelligence (AI) has emerged as an important empirical modelling tool amongst the analytical chemists for solving ion chromatographic separation problems [20–24]. There are several types of AI tools such as genetic algorithms, expert systems, inductive learning, fuzzy logic and artificial neural networks (ANN). Conventional computers are good at executing algorithms that apply procedural logic to well-defined problems. Unluckily, they are not very clever at emulating human cognitive skills. However computers that are based on ANN are seen as being a promising solution to these kinds of AI problems, since they endeavour to imitate the way in which the human brain works. Their ability to learn and generalize well, fast operation and ease of implementation features have made ANNs method of first choice for retention modelling in ion chromatography [23,24]. Once trained, the neural network can be used during ion chromatographic method development process to provide instant answer to the task it learned.

Of all available types of artificial neural networks, multilayered perceptrons (MLPs) are the most commonly used. There are many algorithms for training MLP networks [25–29]. The popular back propagation (BP) algorithm [30] is simple but reportedly has a problem with slow convergence. Thus, various related algorithms have been introduced to address that problem [25]. Most of them are based on second order information about the shape of the error surface [31]. The need to select the larger number of parameters for second order algorithm increases possibility of incorrect setting their values.

The aim of this work is development of the suitable artificial neural network retention model, which can be used in a variety of applications for method development and retention modelling of inorganic cations in ion chromatography. MLP artificial neural networks were used to model retention behaviour of void peak, lithium, sodium, ammonium, potassium, magnesium, calcium, strontium and barium in relation with eluent flow rate and concentration of methasulphonic acid (MSA) in eluent. MLP training algorithm, activation function, number of hidden layer neurons and 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 the calculation procedures. The advantage of the developed model is application of optimized two-phase training algorithm. The two-phase training algorithm represents a combination of two training algorithms, in this case: back propagation (BP), used for 100 iteration steps followed by BP, delta-bar-delta (DBD), quick propagation (QP), conjugate gradient (CG), quasi Newton (QN) or Levenberg–Marquardt (LM), used until minimum on error surface has been found. Two-phase training enables to use advantages of two training algorithms in one training procedure resulting with better predictive ability obtained within shorter time needed for the calculations.

2. Theory

ANNs are massively parallel, highly connected structures consisting of a number of simple, nonlinear processing elements. Because of their massively parallel structure, they can perform computations at a very high rate if implemented on a dedicated hardware. Their adaptive nature enables them to learn the characteristics of input signals and adapt to changes in the data. Because of their nonlinear nature they can perform functional approximation and data fitting operations which are beyond optimal linear techniques. Feedforward neural networks are a basic type of neural networks capable of approximating generic classes of functions, including continuous and integrable ones. An important class of feedforward neural networks is MLP neural networks. An MLP consists of three types of layers: an input layer, an output layer and one or more hidden layers. An appropriate MLP structure helps to achieve higher model accuracy.

2.1. Multilayer perceptron neural networks

Basic structure of MLP is shown in Fig. 1. Suppose the total number of hidden layers is L. The input layer is considered as layer 0. Let the number of neurons in hidden layer l be N_l, $l = 1, 2, \ldots, L$. Let w_{ij}^{l} represent the weight of the link between the jth neuron of the $l-1$st hidden layer and ith neuron of the lth hidden layer, and θ_i^l be the bias parameter of ith neuron of the lth hidden layer. Let x_i represent the ith input

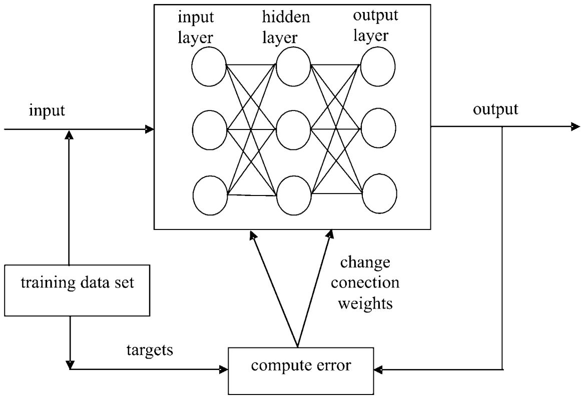


Fig. 1. Basic structure of MLP neural networks.

parameter to the MLP. Let y_{ij}^{-l} be the output of *i*th neuron of the *l*th hidden layer, which can be computed according to the standard MLP formulas as:

$$y_i^{-l} = f\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right) \quad (1)$$

$$i = 1, \ldots, N_l, l = 1, \ldots, L$$

$$\overline{y}_i^0 = x_i, \quad i = 1, \ldots, N_x, \quad N_x = N_0$$

where *f*(·) is the activation function. Let v_{ki} represent the weight of the link between the *i*th neuron of the *L*th hidden layer and the *k*th neuron of the output layer, and β_k be the bias parameter of the *k*th output neuron. The outputs of MLP can be computed as:

$$y_k = \sum_{i=1}^{N_L} v_{ki} y_i^{-L} + \beta_k \quad (2)$$

$$k = 1, \ldots, N_y$$

2.2. Training algorithms

A neural network model can be developed through a process called training. Suppose the training data consists of N_p sample pairs, $\{(x_p, d_p), p = 1, 2, \ldots, N_p\}$, where x_p and d_p are N_x- and N_y-dimensional vectors representing the inputs and the desired outputs of the neural network, respectively. Let w be the weight vector containing all the N_w weights of the neural network. The objective of training is to find w such that the error between the neural network predictions and the desired outputs are minimized,$\min_w E(w)$, where

$$E(w) = \frac{1}{2} \sum_{p=1}^{N_p} \sum_{k=1}^{N_y} (y_{pk}(x_p, w) - d_{pk})^2 = \frac{1}{2} \sum_{p=1}^{N_p} e_p(w) \quad (3)$$

and d_{pk} is the *k*th element of vector d_p, $y_{pk}(x_p, w)$ is the *k*th output of the neural network when the input presented to the network is x_p. The term $e_p(w)$ is the error in the output due to the *p*th sample. Every training algorithm has its own scheme for updating the weights of the neural network.

Training algorithms are an integral part of ANN model development. An appropriate topology may still fail to give a good model, unless trained by a suitable training algorithm. A good training algorithm will shorten the training time, while achieving a better accuracy. Therefore, training process is an important characteristic of the ANNs, whereby representative examples of the knowledge are iteratively presented to the network, so that it can integrate this knowledge within its structure. There are a number of training algorithms used to train a MLP and a frequently used one is called the BP training algorithm [32]. The BP algorithm, which is based on searching an error surface using gradient descent for points with minimum error, is relatively easy to implement. However, the BP algorithm has some problems for many applications. The algorithm is not guaranteed to find the global minimum of the error function since gradient descent may get stuck in local minima, where it may remain indefinitely. In addition to this, long training sessions are often required in order to find an acceptable weight solution because of the well-known difficulties inherent in gradient descent optimization. Therefore, a lot of variations to improve the convergence of the BP were proposed such as DBD, EDBD (extended delta-bar-delta), QP [33,34]. Optimization methods such as second-order methods (CG, QN, LM) have also been used for ANN training in recent years. The LM algorithm combines the best features of the Gauss–Newton technique and the steepest-descent algorithm, but avoids many of their limitations. In particular, it generally does not suffer from the problem of slow convergence [35,36]. The relative performance of algorithms depends on the problem being tackled. Therefore, in this study the MLP were trained with the BP, DBD, QP, CG, QN and LM algorithms.

3. Experimental

3.1. Instrumentation

Dionex DX600 chromatography system (Sunnyvale, CA, USA) equipped with quartenary gradient pump (GS50), chromatography module (LC30) and detector module (ED50A) was used in all experiments. Separation and suppressor columns used were Dionex IonPac CG12A (4 mm × 50 mm) guard column, IonPac CS12A (4 mm × 250 mm) separation column and CAES – 4 mm suppressor column, working in recycle mode, were used, respectively. The sample-loop volume was 25 μL. The eluent flow rate was varied from 0.20 to 2.00 mL/min and concentration of MSA in eluent was varied from 10.00 to 25.00 mmol/L. The whole system was computer controlled through Chromeleon 6.40 + SP1 Build 7.11 software.

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

3.2. Reagents and solutions

Mixed standard solution of lithium (1.0000 g/L), sodium (1.0000 g/L), ammonium (1.0000 g/L), potassium (1.0000 g/L), magnesium (1.0000 g/L), calcium (1.0000 g/L), strontium (1.0000 g/L) and barium (1.0000 g/L) were prepared from the air-dried (at 105 °C) salts of individual anions of p.a. grade (Merck, Darmstadt, Germany). Appropriate amounts of individual salts were weighted into a volumetric flask (100 mL) and dissolved with Milli-Q water. Working standard solutions of lithium (10.00 mg/L), sodium (10.00 mg/L), ammonium (10.00 mg/L), potassium (10.00 mg/L), magnesium (10.00 mg/L), calcium (10.00 mg/L), strontium (10.00 mg/L) and barium (10.00 mg/L) were prepared by measuring the appropriate volume of mixed standard solution onto a 100 mL volumetric flask, which was later filled to the mark with Milli-Q water. 25 mmol/L methansulfonic acid (MSA) standard eluent solution was prepared by dilution of concentrated MSA (Fluka Chemie, Buchs, Switzerland) with degassed Milli-Q water. Working eluent solutions were prepared on-line by appropriate dilution of standard eluent solution with Milli-Q water. A 18 MΩ cm^{-1} water (Millipore, Bedford, MA, USA) was used for dilution in all cases.

3.3. Experimental design

The experimental design has been planed in order to describe the chromatographic behaviour in a multi-dimensional space: retention time versus eluent flow rate and concentration of MSA in eluent. The eluent flow rate was varied in range from 0.20 to 2.00 mL/min and concentration of MSA in eluent was varied from 10.00 to 25.00 mmol/L. The 96 experimental data points were obtained. Fig. 2 presents the experimental design model used for collecting experimental data.

It is preferable that each experimental data point has equal influence on the neural network model, in order for training and testing set to be representative group of data of the whole design area. For this reason the random function was applied for selection of experimental data points used for training, testing and validation set of data. The number of experimental data points used for training set was varied in the following ratios: training: testing: validation = 1:1:1; 2:1:1; 3:1:1. The input experimental data were scaled linearly between values zero and one before modelling. This is necessary as, although most neural networks can accept input values in any range, they are only sensitive to inputs in a far smaller range.

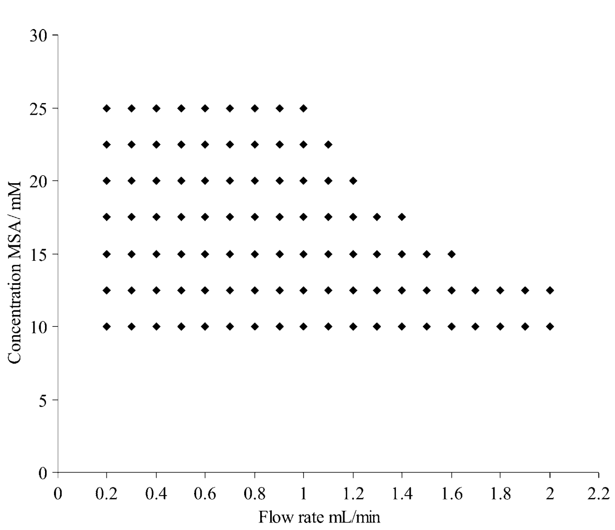


Fig. 2. Design of 96 experimental data points: eluent flow rate vs. concentration of MSA in eluent.

3.4. Neural networks

The neural network used in this paper was the three-layer feed forward neural network. The input layer consists of the two neurons representing eluent flow rate and concentration of MSA in eluent. The output layer consists of one neuron representing the retention of void peak or retention time of one of the particular cations (lithium, sodium, ammonium, potassium, magnesium, calcium, strontium, barium). The training algorithm, activation function, number of neurons in hidden layer and the number of experimental data points used for training calculations need to be optimized. Therefore training algorithm was varied between back propagation (BP), delta-bar-delta (DBD), quick propagation (QP), conjugate gradient (CG), quasi Newton (QN) or Levenberg–Marquardt (LM), the activation function was varied between logistic and hyperbolic, the number of nodes in training set was varied form 5 to 11, and the ratios between sizes of training, testing and validation were varied accordingly 1:1:1; 2:1:1; 3:1:1.

Two-phase training procedure was used for all calculations. The first phase was 100 iteration steps of error back propagation training in order to achieve fast convergence to the region of global minimum on error surface. The second phase was varied between back propagation, delta-bar-delta, quick propagation, conjugate gradient, quasi Newton and Levenberg–Marquardt trainings in order to achieve faster and more accurate convergence. The second phase training procedure was repeated until the global minimum on error surface was found.

The activation function connecting the input and hidden layer of nodes was varied between logistic function, defined as

$$y_i^{-l} = \frac{1}{1 + \mathrm{e}^{-\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right)}} \tag{4}$$

and hyperbolic function defined as

$$y_i^{-l} = \frac{\mathrm{e}^{\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right)} - \mathrm{e}^{-\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right)}}{\mathrm{e}^{\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right)} + \mathrm{e}^{-\left(\sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l\right)}} \tag{5}$$

For computation of output activities linear transfer function was employed:

$$y_i^{-l} = \sum_{j=l}^{N_{l-1}} w_{ij}^l y_j^{-l-1} + \theta_i^l \tag{6}$$

To test the predictive performance of the developed artificial neural network retention model an independent validation set was used. Statistical analysis was used to calculate agreement between measured and predicted values by using standard Pearson-*R* correlation coefficient, confidence intervals for intercept and slope as well as SD ratio. The ratio of the prediction error standard deviation to the original output data standard deviation is called the SD ratio. A lower SD ratio indicates a better prediction. This is equivalent to one minus the explained variance of the model.

All calculations were performed in Statistica 6.1 (StatSoft Inc., USA) environment by using IBM compatible personal computer equipped with 2.66 MHz Pentium IV processor, and 512 Mb RAM.

4. Results and discussion

Tables 1–9 present the results of two-phase training algorithm optimization. The results for the two-phase training algorithm optimization in the case of retention modelling of void volume (Table 1), potassium (Table 5) and magnesium (Table 6) shows that maximal correlation coefficient and minimal SD ratio are obtained by using BP–QN and BP–LM algorithm. On the contrary, results for retention modelling of lithium (Table 2) indicates BP–QN and BP–BP as the two best performing two-phase training algorithms, for sodium (Table 3) BP–QN and BP–BP, for the ammonium (Table 4) BP–QN and DBD, for calcium (Table 7) BP–DBD and BP–LM, for strontium (Table 8) BP–QN and BP–DBD and for the barium (Table 9) BP–LM and BP–DBD, respectively. It can be seen (Tables 1–9) that BP–BP and BP–DBD need significantly more (up to 30 times more for lithium) iteration steps to converge to the global minimum on error surface. It means that significantly more time is needed for BP–BP and BP–DBD training indicating those two methods as a time consuming ones, especially if no significant improvement in prediction ability is obtained (Tables 1–9). On the basis of previous discussion it can be stated that the optimal two-phase training algorithms for retention modelling of inorganic cations are BP–QN and BP–LM. All further calculations were made by using those two training algorithms.

Tables 10–18 present the results of activation function optimization. It can be seen that in the case of void peak (Table 10), lithium (Table 11), potassium (Table 14), strontium (Table 17) and barium (Table 18) retention modelling, maximal correlation coefficient and minimal SD ratio are

Table 1
Results for two-phase training algorithm optimization in the case of void peak retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	68	0.052956	0.99861
Back propagation	100	Quick propagation	404	0.028966	0.99959
Back propagation	100	Back propagation	14997	0.019222	0.99982
Back propagation	100	Delta-bar-delta	14990	0.014482	0.99991
Back propagation	100	Levenberg–Marquardt	490	0.013127	0.99991
Back propagation	100	Quasi Newton	268	0.012636	0.99992

Table 2
Results for two-phase training algorithm optimization in the case of lithium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	159	0.066635	0.99778
Back propagation	100	Quick propagation	533	0.06292	0.99807
Back propagation	100	Delta-bar-delta	14921	0.043393	0.99906
Back propagation	100	Levenberg–Marquardt	328	0.04304	0.99907
Back propagation	100	Back propagation	14433	0.039799	0.99921
Back propagation	100	Quasi Newton	550	0.032361	0.99948

Table 3
Results for two-phase training algorithm optimization in the case of sodium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	139	0.074935	0.99719
Back propagation	100	Quick propagation	1405	0.057557	0.99843
Back propagation	100	Delta-bar-delta	1106	0.049182	0.99887
Back propagation	100	Levenberg–Marquardt	9136	0.048062	0.99897
Back propagation	100	Back propagation	14327	0.043291	0.99906
Back propagation	100	Quasi Newton	342	0.038349	0.99929

Table 4
Results for two-phase training algorithm optimization in the case of ammonium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	33	0.069475	0.99758
Back propagation	100	Quick propagation	1178	0.052241	0.99864
Back propagation	100	Delta-bar-delta	14601	0.049781	0.99877
Back propagation	100	Levenberg–Marquardt	1491	0.047566	0.9989
Back propagation	100	Back propagation	2901	0.04816	0.99893
Back propagation	100	Quasi Newton	68	0.043385	0.99921

Table 5
Results for two-phase training algorithm optimization in the case of potassium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	203	0.07222	0.9974
Back propagation	100	Quick propagation	14998	0.05243	0.9987
Back propagation	100	Delta-bar-delta	14535	0.05165	0.9987
Back propagation	100	Levenberg–Marquardt	1218	0.05125	0.9987
Back propagation	100	Back propagation	1499	0.04687	0.9989
Back propagation	100	Quasi Newton	1480	0.04435	0.999

Table 6
Results for the two-phase training algorithm optimization in the case of magnesium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	6410	0.1128	0.9936
Back propagation	100	Quick propagation	14392	0.08638	0.9963
Back propagation	100	Delta-bar-delta	1409	0.0784	0.997
Back propagation	100	Levenberg–Marquardt	547	0.07534	0.9972
Back propagation	100	Back propagation	1414	0.0684	0.9977
Back propagation	100	Quasi Newton	1365	0.0682	0.9977

Table 7
Results for two-phase training algorithm optimization in the case of calcium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	233	0.1057	0.9949
Back propagation	100	Quick propagation	993	0.08315	0.9966
Back propagation	100	Delta-bar-delta	438	0.0707	0.9975
Back propagation	100	Levenberg–Marquardt	12527	0.07475	0.9976
Back propagation	100	Back propagation	34	0.06326	0.998
Back propagation	100	Quasi Newton	9408	0.05954	0.9982

Table 8
Results for two-phase training algorithm optimization in the case of strontium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	65	0.1112	0.9939
Back propagation	100	Quick propagation	1295	0.09919	0.9951
Back propagation	100	Delta-bar-delta	14947	0.09408	0.9956
Back propagation	100	Levenberg–Marquardt	390	0.09856	0.9958
Back propagation	100	Back propagation	14189	0.07675	0.9971
Back propagation	100	Quasi Newton	447	0.06939	0.9977

Table 9
Results for two-phase training algorithm optimization in the case of barium retention modelling

Training, step one		Training, step two		Validation error	
Method	Number of iteration steps	Method	Number of iteration steps	SD ratio	Correlation
Back propagation	100	Conjugate gradient descent	127	0.1152	0.9935
Back propagation	100	Quick propagation	14196	0.08685	0.9964
Back propagation	100	Delta-bar-delta	867	0.08392	0.9965
Back propagation	100	Levenberg–Marquardt	499	0.07481	0.9972
Back propagation	100	Back propagation	13578	0.07398	0.9973
Back propagation	100	Quasi Newton	1488	0.06335	0.998

Table 10
Results for activation function optimization in the case of void peak retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	435	0.052956	0.99861
Hyperbolic	Back propagation	100	Quasi Newton	481	0.038976	0.99926
Logistic	Back propagation	100	Levenberg–Marquardt	499	0.092538	0.99572
Hyperbolic	Back propagation	100	Levenberg–Marquardt	113	0.41026	0.91344

Table 11
Results for activation function optimization in the case of lithium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	1432	0.040694	0.9992
Hyperbolic	Back propagation	100	Quasi Newton	167	0.04672	0.99893
Logistic	Back propagation	100	Levenberg–Marquardt	1488	0.035414	0.99939
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1085	0.034124	0.99942

Table 12
Results for activation function optimization in the case of sodium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	1447	0.031122	0.99952
Hyperbolic	Back propagation	100	Quasi Newton	707	0.034706	0.9994
Logistic	Back propagation	100	Levenberg–Marquardt	1447	0.031122	0.99952
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1207	0.036478	0.99934

Table 13
Results for activation function optimization in the case of ammonium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	959	0.039588	0.99926
Hyperbolic	Back propagation	100	Quasi Newton	157	0.044346	0.99902
Logistic	Back propagation	100	Levenberg–Marquardt	1485	0.060147	0.99828
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1497	0.054459	0.99865

Table 14
Results for activation function optimization in the case of potassium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	140	0.04451	0.999
Hyperbolic	Back propagation	100	Quasi Newton	658	0.04142	0.9991
Logistic	Back propagation	100	Levenberg–Marquardt	732	0.04442	0.999
Hyperbolic	Back propagation	100	Levenberg–Marquardt	204	0.05078	0.9987

Table 15
Results for activation function optimization in the case of magnesium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	1443	0.0603	0.9982
Hyperbolic	Back propagation	100	Quasi Newton	136	0.07901	0.9969
Logistic	Back propagation	100	Levenberg–Marquardt	530	0.06813	0.9979
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1449	0.08581	0.9965

Table 16
Results for activation function optimization in the case of calcium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	460	0.08159	0.9967
Hyperbolic	Back propagation	100	Quasi Newton	226	0.08784	0.9961
Logistic	Back propagation	100	Levenberg–Marquardt	460	0.08159	0.9967
Hyperbolic	Back propagation	100	Levenberg–Marquardt	187	0.08305	0.9966

obtained by using hyperbolic function, while in the case of sodium (Table 12), ammonium (Table 13), magnesium (Table 15) and calcium (Table 16) retention modelling maximal correlation coefficient and minimal SD ratio are obtained by using logistic function. In the case of void peak, lithium, sodium, ammonium, magnesium, strontium and barium retention modelling (Tables 10–13, 15, 17, 18), faster convergence to the global minimum on error surface is obtained by using hyperbolic activation function. Logistic activation function provides faster convergence in the case of potassium and calcium retention modelling (Tables 14 and 16). Results for potassium show that correlation coefficient in the case of using either logistic or hyperbolic function in combination with BP–QN and BP–LM training algorithm does not differ significantly. Results for calcium retention modelling shows that significantly lower prediction is obtained in the case of using logistic function and BP–LM training algorithm then by using hyperbolic activation function. This indicates hyperbolic transfer function as the optimal choice for retention modelling of inorganic cations. All further calculations were made by using hyperbolic activation function.

Table 19 presents the results of optimization procedures of the number of hidden layer neurons and the number of experimental data points used for training calculations. The combination of neural networks parameters, number of hidden layer neurons, number of experimental data points used for training set and training algorithm, which provides maximal correlation coefficient and minimal SD ratio presents optimized neural network retention model. From Table 19 it can be seen that the optimal number of hidden layer neurons is 11 for void peak, lithium, ammonium, potassium and magnesium, 9 for sodium and barium, 7 for strontium and 5 for calcium. The optimal number of experimental data points used for training set (training:testing:validation) is 3:1:1 for lithium, sodium, ammonium, potassium, magnesium, calcium, strontium and barium and 2:1:1 for void peak. BP–LM training algorithm provides maximal correlation coefficient and minimal SD ratio for all cations. Predictive ability of optimized artificial

Table 17
Results for activation function optimization in the case of strontium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	810	0.06726	0.9979
Hyperbolic	Back propagation	100	Quasi Newton	120	0.06723	0.9979
Logistic	Back propagation	100	Levenberg–Marquardt	1636	0.06671	0.9978
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1319	0.1139	0.9936

Table 18
Results for activation function optimization in the case of barium retention modelling

Activation function	Training, step one		Training, step two		Validation error	
	Method	Iteration steps	Method	Iteration steps	SD ratio	Correlation
Logistic	Back propagation	100	Quasi Newton	694	0.0688	0.9976
Hyperbolic	Back propagation	100	Quasi Newton	274	0.09577	0.9954
Logistic	Back propagation	100	Levenberg–Marquardt	697	0.09453	0.9957
Hyperbolic	Back propagation	100	Levenberg–Marquardt	1489	0.06641	0.9978

Table 19
Optimization of number of hidden layer neurons and number of experimental data points used for training calculations

Cation	Phase one, iteration steps	Phase two, iteration steps	Number of hidden layers neurons	Sampling, training:testing:validating	Correlation	SD ratio
Void	BP, 100	QN, 470	7	3:1:1	0.9998	0.0120
	BP, 100	LM, 491	11	2:1:1	0.9999	0.0119
Lithium	BP, 100	QN, 745	7	2:1:1	0.9994	0.0308
	BP, 100	LM, 1486	11	3:1:1	0.9995	0.0299
Sodium	BP, 100	QN, 239	11	3:1:1	0.9988	0.0312
	BP, 100	LM, 1490	9	3:1:1	0.9990	0.0295
Ammonium	BP, 100	QN, 1305	7	3:1:1	0.9992	0.0348
	BP, 100	LM, 678	11	3:1:1	0.9994	0.0344
Potassium	BP, 100	QN, 697	9	3:1:1	0.9995	0.0315
	BP, 100	LM, 1490	11	3:1:1	0.9995	0.0327
Magnesium	BP, 100	QN, 60	11	2:1:1	0.9982	0.06051
	BP, 100	LM, 320	11	3:1:1	0.9984	0.0566
Calcium	BP, 100	QN, 458	9	3:1:1	0.9984	0.0567
	BP, 100	LM, 885	5	3:1:1	0.9983	0.0399
Strontium	BP, 100	QN, 796	11	3:1:1	0.9982	0.0519
	BP, 100	LM, 486	7	3:1:1	0.9985	0.0553
Barium	BP, 100	QN, 2637	9	3:1:1	0.9990	0.0444
	BP, 100	LM, 3188	9	3:1:1	0.9995	0.0305

neural networks retention models was tested by employing statistical calculations.

The prediction powers of developed artificial neural networks retention models were tested and results are shown in Figs. 3–11. It is shown that relationships between simulated retention times (y) against measured retention times (x) were investigated. If there were no modeling errors and no measurement random errors were made and if there were no bias, this would yield the relationship $y = x$. Because at least random errors were made the coefficients of linear relationship were different (intercept was different from zero and/or slope was different from one). There are four possibilities, namely:

- if intercept is equal to zero and slope is equal to one, there is no systematic error,
- if intercept is equal to zero and slope is different of one, there is proportional systematic error,
- if intercept is different from zero and slope is equal to one, there is absolute systematic error,

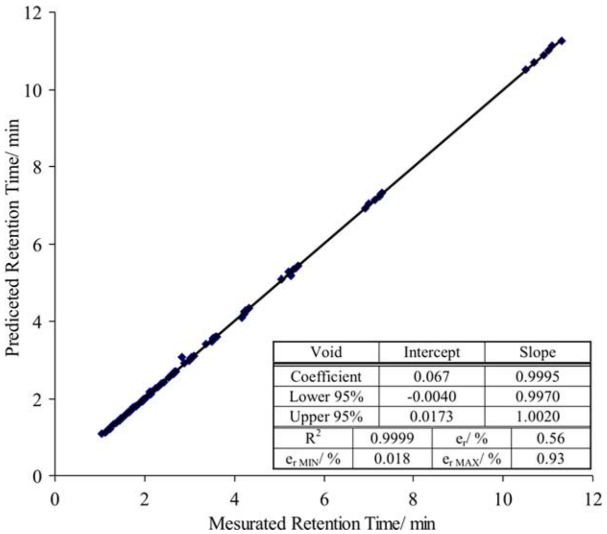


Fig. 3. Predictive ability calculations for optimized void peak neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\,AVER}}$ denotes average relative error, $e_{\mathrm{r\,MIN}}$ denotes minimal relative error and $e_{\mathrm{r\,MAX}}$ denotes maximal relative error.

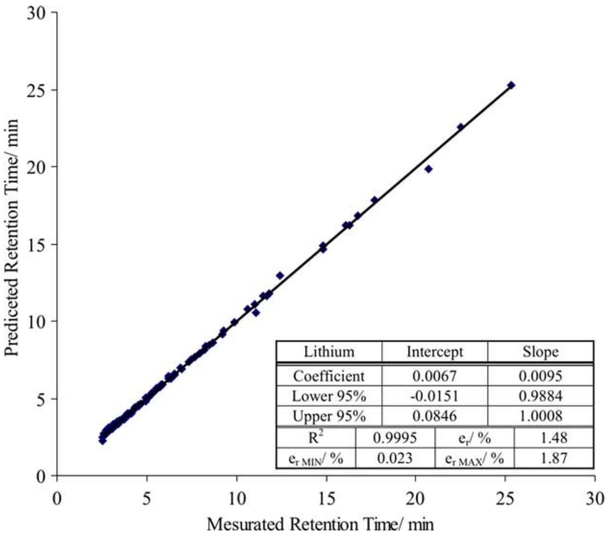


Fig. 4. Predictive ability calculations for optimized lithium neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\,AVER}}$ denotes average relative error, $e_{\mathrm{r\,MIN}}$ denotes minimal relative error and $e_{\mathrm{r\,MAX}}$ denotes maximal relative error.

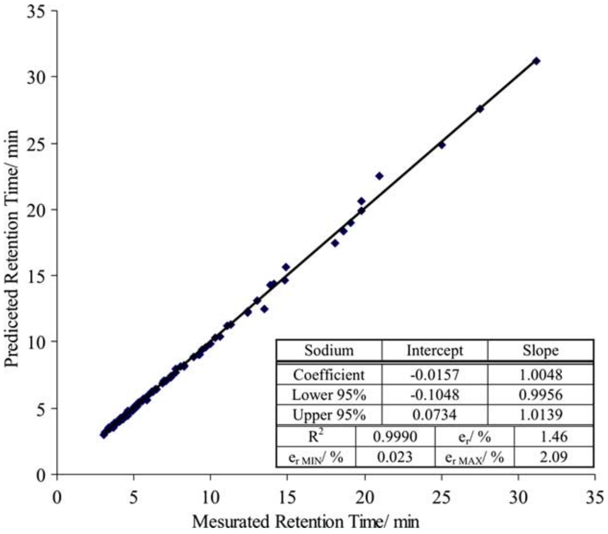


Fig. 5. Predictive ability calculations for optimized sodium neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\ AVER}}$ denotes average relative error, $e_{\mathrm{r\ MIN}}$ denotes minimal relative error and $e_{\mathrm{r\ MAX}}$ denotes maximal relative error.

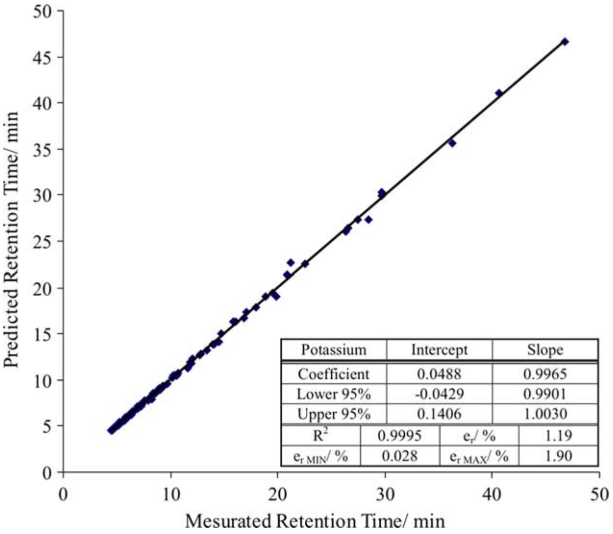


Fig. 7. Predictive ability calculations for optimized potassium neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\ AVER}}$ denotes average relative error, $e_{\mathrm{r\ MIN}}$ denotes minimal relative error and $e_{\mathrm{r\ MAX}}$ denotes maximal relative error.

- if there is no linearity, it is necessary to carry out the modeling over a shorter ion chromatographic parameter range (if this is still compatible with original aim).

From Figs. 3–11, it can be seen that lower and upper 95% confidence interval limits boundaries for intercept include value zero. This proves that intercept of calibration curve is not significantly different from zero whit respect of confidence 95%. Furthermore, lower and upper 95% confidence interval limits boundaries for slope include value one, what proves that slope of calibration curve is not significantly different from one with respect of confidence 95%. On the basis of previous discussion, it can be stated that there is no systematic error present in optimized artificial neural network retention models for all cations.

Since correlation coefficient is the measure of the joint variation between two variables, it presents the strength of the proposed linear relationship between a predicted and measured retention times. It can be seen (Figs. 3–11) that correlation coefficient (R^2) has satisfactory values in the range between 0.9983 and 0.9999, and strong linear relationship between predicted and measured retention times exist for all cations. The average of relative errors are ranged between 0.56 and 1.52%, minimal relative errors are ranged between

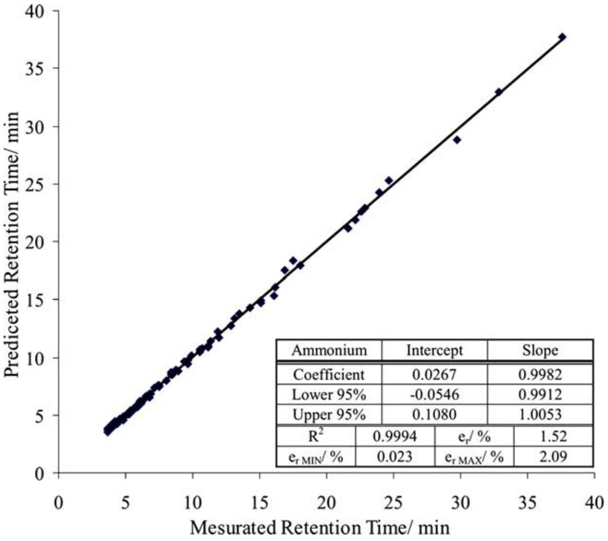


Fig. 6. Predictive ability calculations for optimized ammonium neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\ AVER}}$ denotes average relative error, $e_{\mathrm{r\ MIN}}$ denotes minimal relative error and $e_{\mathrm{r\ MAX}}$ denotes maximal relative error.

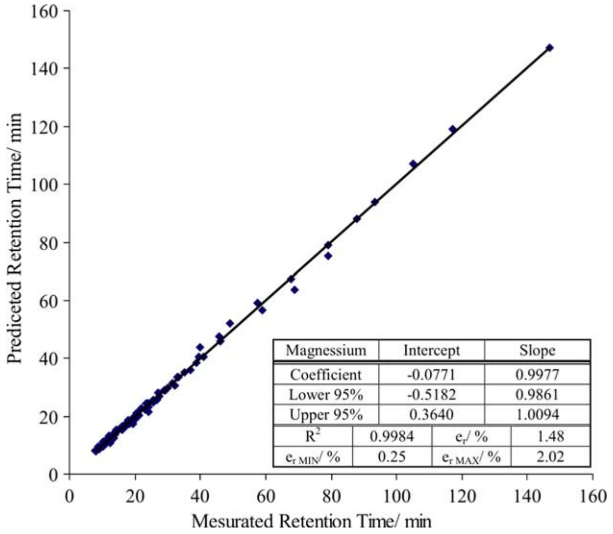
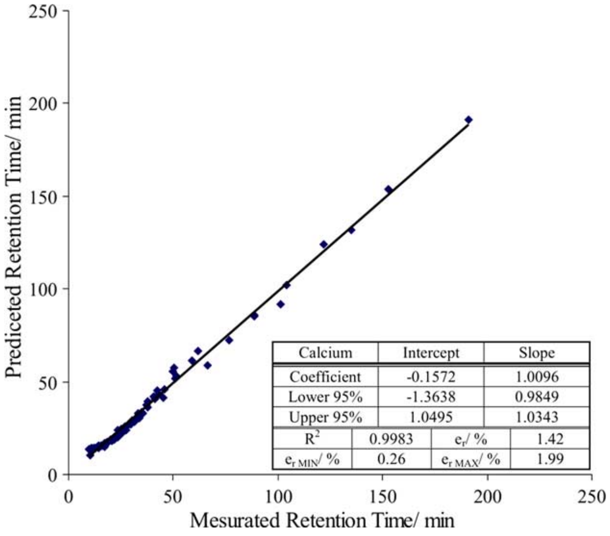


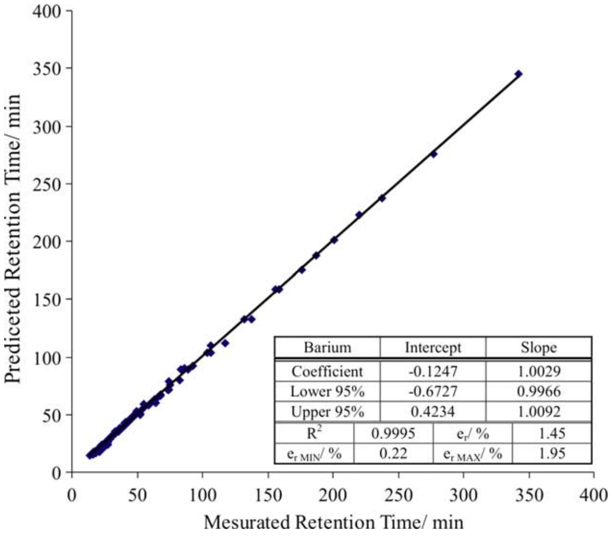
Fig. 8. Predictive ability calculations for optimized magnesium neural network retention model. R^2 denotes standard Pearson-R correlation, $e_{\mathrm{r\ AVER}}$ denotes average relative error, $e_{\mathrm{r\ MIN}}$ denotes minimal relative error and $e_{\mathrm{r\ MAX}}$ denotes maximal relative error.



Calcium	Intercept	Slope
Coefficient	-0.1572	1.0096
Lower 95%	-1.3638	0.9849
Upper 95%	1.0495	1.0343

R^2	0.9983	e_r/ %	1.42
$e_{r\ MIN}$/ %	0.26	$e_{r\ MAX}$/ %	1.99

Fig. 9. Predictive ability calculations for optimized calcium neural network retention model. R^2 denotes standard Pearson-*R* correlation, $e_{r\ AVER}$ denotes average relative error, $e_{r\ MIN}$ denotes minimal relative error and $e_{r\ MAX}$ denotes maximal relative error.



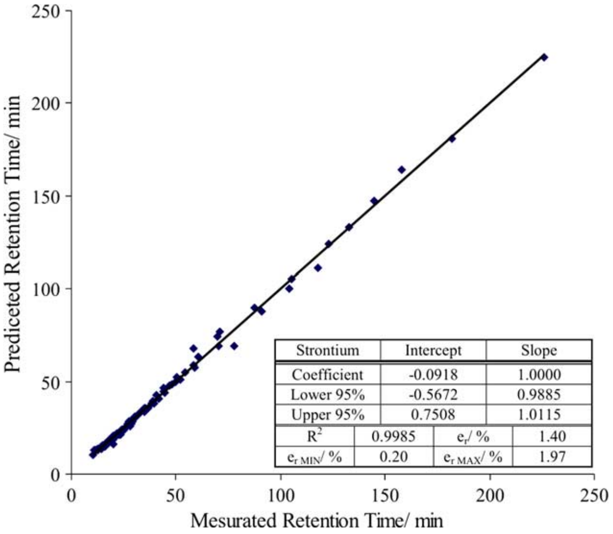
Barium	Intercept	Slope
Coefficient	-0.1247	1.0029
Lower 95%	-0.6727	0.9966
Upper 95%	0.4234	1.0092

R^2	0.9995	e_r/ %	1.45
$e_{r\ MIN}$/ %	0.22	$e_{r\ MAX}$/ %	1.95

Fig. 11. Predictive ability calculations for optimized barium neural network retention model. R^2 denotes standard Pearson-*R* correlation, $e_{r\ AVER}$ denotes average relative error, $e_{r\ MIN}$ denotes minimal relative error and $e_{r\ MAX}$ denotes maximal relative error.

0.018 and 0.028% and maximal relative errors are ranged between 1.87 and 2.02%. On the basis of conducted statistical test, it can be concluded that developed ANN retention models have good generalization ability.

By applying developed artificial neural networks for adjustment of eluent flow rate and concentration of MSA in eluent it is possible to increase selectivity. That is crucial factor for numerous different applications of ion chromatography analysis, particular for wastewater analysis and analysis of samples with great differences in concentration of analyte components. By adjusting the retention times of sodium and ammonium it is possible to determine mentioned cations in ratios even above 1000:1. The possibility of retention times adjustment of potassium, calcium and magnesium is crucial for determination of rubidium and cesium as well as for determination of ethyilamines and morpholine. By adjusting retention times of late eluting cations especially strontium and barium it is possible to obtain shorter ion chromatographic run and speed up analysis, without decreasing selectivity of fast eluting anions (lithium, sodium, ammonium).



Strontium	Intercept	Slope
Coefficient	-0.0918	1.0000
Lower 95%	-0.5672	0.9885
Upper 95%	0.7508	1.0115

R^2	0.9985	e_r/ %	1.40
$e_{r\ MIN}$/ %	0.20	$e_{r\ MAX}$/ %	1.97

Fig. 10. Predictive ability calculations for optimized strontium neural network retention model. R^2 denotes standard Pearson-*R* correlation, $e_{r\ AVER}$ denotes average relative error, $e_{r\ MIN}$ denotes minimal relative error and $e_{r\ MAX}$ denotes maximal relative error.

5. Conclusion

In this work artificial neural networks were used for retention modelling of cations in ion chromatography. First order training algorithms have shown not to be optimal due to a fact that training procedures were very slow and time consuming. When second order training algorithms were applied weight initialization could yield irreproducible neural network retention model. Every new initialisation can be regarded as a new start position for the second order algorithm search for the global minimum. Obviously, the chance of finding the global minimum directly depends on the smoothness of the error hyperplane and the number of local minima. 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 applying two-phase training procedure consisting of both first and second order training algorithms. This study shows that the optimized two-phase training consisting of first and second order algorithms insures faster training with a higher probability of avoiding local minima.

In this work the activation function, the number of hidden layer neurones, and the number of experimental data points used for training set are optimized in order to obtain a neural network model with good predictive ability. The optimized neural network retention models were used to predict retention times for void peak and all eight cations (lithium, sodium, ammonium, potassium, magnesium, calcium, strontium, barium). It can be seen that there is no systematic error present in optimized artificial neural network retention models for void peak and all cations. Correlation coefficients are ranged between 0.9983 and 0.9999. From these results it can be concluded that developed 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 MSA in eluent). The developed retention model allows manipulating with appearance of the particular peak on chromatogram and allows improvement of separation between particular cations. By using this retention model it is possible both to improve performance characteristic of applied method and to speed up the new method development by reducing unnecessary experimentation.

References

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

[2] J. Weiss, Ion Chromatography, second ed., VCH Verlagsgwsellshaft mbH, Weiheim, 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] M.J. Van Os, J. Slanina, C.L. De Ligny, W.E. Hammers, J. Agterdenbos, Anal. Chim. Acta 144 (1982) 73.

[7] P.R. Haddad, C.E. Cowie, J. Chromatogr. 303 (1984) 321.

[8] A. Jardy, M. Caude, A. Diop, C. Curvale, R. Roset, J. Chromatogr. 439 (1988) 137.

[9] T.B. Hoover, Sep. Sci. Technol. 17 (1982) 295.

[10] D.R. Jenke, G.K. Pagenkopf, Anal. Chem. 56 (1984) 85.

[11] D.R. Jenke, G.K. Pagenkopf, Anal. Chem. 56 (1984) 88.

[12] M. Maruo, N. Hirayama, T. Kuwamoto, J. Chromatogr. 481 (1989) 315.

[13] N. Hirayama, T. Kuwamoto, J. Chromatogr. 508 (1990) 51.

[14] A. Yamamoto, K. Hayakawa, A. Matsunaga, E. Mizukami, M. Miyazaki, J. Chromatogr. 627 (1992).

[15] D.R. Jenke, Anal. Chem. 66 (1994) 4466.

[16] J.E. Madden, P.R. Haddad, J. Chroamtogr. A 829 (1998) 65.

[17] J.E. Madden, P.R. Haddad, J. Chroamtogr. A 850 (1–2) (1999) 29.

[18] J.E. Madden, N. Avdalović, P.E. Jackson, P.R. Haddad, J. Chromatogr. A 855 (1) (1999) 65.

[19] P. Janoš, J. Chromatogr. A 789 (1–2) (1997) 3.

[20] G. Sacchero, M.C. Buzzoniti, C. Sarzanini, E. Mentasti, H.J. Metting, P.M.J. Coenegracht, J. Chromatogr. A 799 (1998) 35.

[21] J. Havel, J.E. Madden, P.R. Haddad, Chromatographia 49 (1999) 481.

[22] J.E. Madden, N. Avdalovic, P.R. Haddad, J. Havel, J. Chromatogr. A 910 (2001) 173.

[23] G. Srečnik, Ž Debeljak, Š. Cerjan-Stefanović, T. Bolanča, M. Novič, K. Lazarić, Ž GumhalterLulić, Croat. Chem. Acta 75 (3) (2002) 713.

[24] G. Srečnik, Ž Debeljak, Š. Cerjan-Stefanović, M. Novič, T. Bolanča, J. Chromatogr. A 973 (2002) 47.

[25] A. Maren, C. Harston, R. Pap, Handbook of Neural Computing Applications, Academic Press, London, 1990.

[26] D.T. Pham, S. Sagiroglu, Proc IMechE, Part B, J. Eng. Manufact. 210 (1996) 69.

[27] D. Alpsan, M. Towsey, O. Ozdamar, A.C. Tsoi, D.N. Ghista, Neural Networks 8 (6) (1995) 945.

[28] K.W. Tang, G. Pingle, G. Srikant, J. Intell. Syst. 7 (1997) 307.

[29] F. Stager, M. Agarwal, Neural Networks 10 (1997) 1435.

[30] R. Drake, M.S. Packianather, Int. J. Adv. Manufact. Technol. 14 (1998) 280.

[31] P. Devika, L. Achenie, J. Intell. Fuzzy Syst. 3 (1995) 287.

[32] B.B. Chaudhuri, U. Bhattacharya, Neurocomputing 34 (2000) 11.

[33] R.A. Jacobs, Neural Networks 1 (1988) 295.

[34] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Nature 323 (1986) 533.

[35] M.T. Hagan, M.B. Menhaj, IEEE Trans. Neural Networks 5 (6) (1994) 989.

[36] F. Wang, V.K. Devabhaktuni, C. Xi, Q. Zhang, Int. J. RF Microwave Comput. Aided Eng. 9 (3) (1999) 216.