ELSEVIER

Journal of Pharmaceutical and Biomedical Analysis
17 (1998) 69–76

JOURNAL OF
PHARMACEUTICAL
AND BIOMEDICAL
ANALYSIS

Application of artificial neural networks in HPLC method development

S. Agatonovic-Kustrin [a,*], M. Zecevic [b], L.j. Zivanovic [b], I.G. Tucker [a]

[a] *School of Pharmacy, University of Otago, Dunedin, New Zealand*
[b] *Institute of Pharmaceutical Chemistry and Drug Analysis, Faculty of Pharmacy, Vojvode Stepe 450, 11000 Belgrade, Serbia*

Received 3 March 1997; received in revised form 29 May 1997

Abstract

The use of artificial neural networks (ANNs) for response surface modelling in HPLC method development for amiloride and methychlothiazide separation is reported. The independent input variables were pH and methanol percentage in mobile phase. The outputs were capacity factors. The results were compared with a statistical method (multiple nonlinear regression analysis). Networks were able to predict the experimental responses more accurately than the regression analysis.

Keywords: Artificial neural network; Backpropagation; HPLC; Amiloride; Methychlothiazide

1. Introduction

An important aspect of method development in liquid chromatography is to achieve adequate separation of all components in a given sample in a reasonable time. Consequently optimization of the different chromatographic factors like pH, mobile phase composition, i.e. organic modifier concentration and temperature is critical for sufficient resolution. Retention mapping techniques are important optimization methods enabling the global optimum to be found [1,2]. The main purpose of this study was to investigate the usefulness of artificial neural networks (ANNs) [3–6] for response surface modelling in HPLC optimisation [7] and to compare the results to those calculated on the basis of the multiple regression method [8]. The combined effect of pH and mobile phase composition on the reverse-phase liquid chromatographic behaviour of amiloride and methychlothiazide was investigated.

The effects of these factors were examined in the range of conditions where they provided acceptable retention and resolution. The effects of methanol (10–50%) and pH (2.5–4.5) were studied. The results show that neural networks offer promising possibilities in HPLC method development. The predictive results obtained by ANNs were better than those obtained with multiple regression models.

* Corresponding author.

0731-7085/98/$19.00
PII S0731-7085(97)00170-2

1.1. Neural networks

ANNs are mathematical systems that simulate biological neural networks. They consists of processing elements (neurons) organized in layers and interconnections between the elements.

ANNs analysis is quite flexible as regards the amount and form of the training (experimental) data which makes it possible to use more informal experimental designs than with statistical approaches. Also, neural network models might generalize better than regression models since regression analyses are dependent on predetermined statistical significance levels. This means that less significant terms are not included in the model. With the ANNs method all data are used potentially making the models more accurate.

Usually a neural network in its basic form is composed of several layers of neurons, there being one input layer, one output layer and at least one hidden layer (Fig. 1). The use of at least one hidden layer enables the ANNs to describe nonlinear systems. A problem in constructing ANNs is to find the optimal number of hidden neurons.

W_{ij} is the weight-connection to neuron j from neuron i, x_i denotes the input values and bias_j is the bias of neuron j. The activation of the jth neuron (Net_j) is defined as the sum of the weighted input signal to that neuron:

$$\text{Net}_j = \sum_i w_{ji} x_i + \text{bias}_j$$

This activation is transformed to the neuron output by a transform function. Different ANN classes use different definitions of the activation function. The most common transform function in back-propagation neural networks (BNN) is a sigmoidal function:

$$y_j = \frac{2}{1 + e^{-\text{Net}_j}} - 1$$

Each neuron in the input layer is connected to each neuron in the hidden layer and each neuron in the hidden layer is connected to each neuron in the output layer to produce the output vector. Information in a BNN is stored as weights, which are connections between neurons in successive

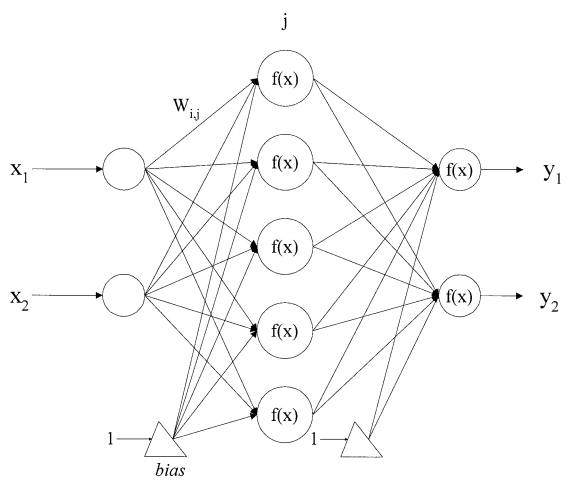

Fig. 1. A simple $2 \times 5 \times 2$ ANN. The lines connecting the neurons represent the weights. Also shown are the bias neurons used to shift the neuron transfer function and to improve the network performance.

Table 1
Fitted K' values for methylclothiazide (M) and amiloride (A) obtained by ANN model with 10 hidden neurones and 15 000 training cycles

Training data		Measured K'		Fitted values	
Methanol	pH	M	A	M	A
10.0	2.8	0.780	2.250	0.939	2.258
10.0	3.5	1.080	2.770	0.997	2.779
10.0	4.2	0.870	4.390	0.867	4.288
30.0	2.8	0.390	0.522	0.352	0.651
30.0	3.5	0.240	0.866	0.320	0.818
30.0	4.2	0.339	1.180	0.325	0.770
50.0	2.8	0.437	0.290	0.519	0.185
50.0	3.5	0.530	0.504	0.272	0.578
50.0	4.2	0.590	0.180	0.845	0.169
30.0	3.5	0.280	0.843	0.320	0.818
			$R^2 =$	0.999	0.999
			SSE =	0.0003	
10	3.2	1.287	2.648	1.011	2.279
15	3.0	0.532	1.450	0.533	1.453
30	3.8	0.283	1.078	0.313	0.800
50	3.0	0.494	0.211	0.488	0.168
			$R^2 =$	0.888	0.899
			SSE =	0.0085	

SSE, Sum of squared errors; R^2, Coefficients of multiple determination.

layers and as bias values (neuron activation threshold). The neural network used in this work is the feed-forward, back-propagation neural network type, most often used in analytical applications. Information from various sets of inputs are fed forward through the BNN to optimize the weight between neurons, or to 'train' it. The error, or bias, in prediction is then propagated through the system and the inter-unit connections are changed to minimize the error in the prediction. This process is continued with multiple training sets until the error is minimized across many sets.

The error of the network is defined as the squared difference (SSE) between the target values t and the outputs y of the output neurons:

$$\mathrm{MSE} = \frac{1}{p \cdot m} \sum_{k=1}^{p} \sum_{l=1}^{m} (y_{kl} - t_{kl})^2$$

where p is the number of training sets, and m is the number of output neurones in the neural network.

During training, neural techniques need to have some way of evaluating their own performance. Since they are learning to associate the inputs with outputs, evaluating the performance of the network on the training data may not produce the best results. If a network is left to train for too long, it will overtrain and will lose the ability to generalize. Thus three types of data sets are used:

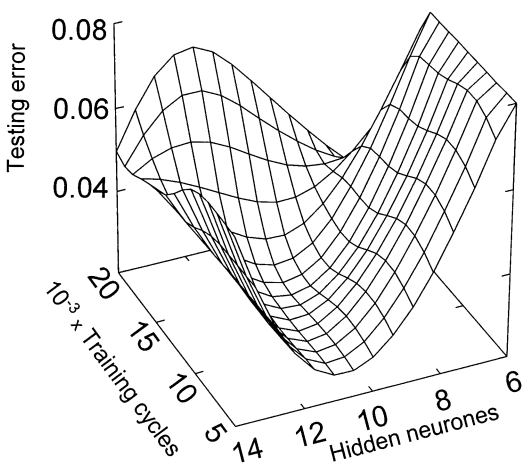

Fig. 2. Effect of the number of hidden neurons and number of cycles during training on the SSE.

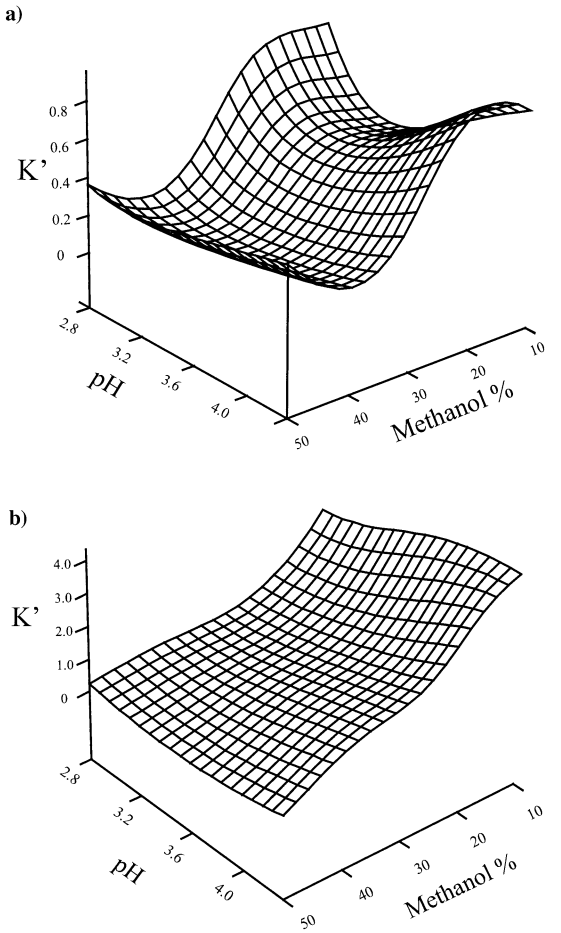

Fig. 3. Response surfaces for multifactor effect of pH and Methanol % on (a) methylclothiazide and (b) amiloride capacity factors generated by ANN with 10 hidden neurons at 15 000 training cycles.

training data: used to train network
test data: used to monitor the neural network performance during training
validation data: used to measure the performance of a trained application, each with corresponding error

1.2. Multifactor regression analysis

A response surface methodology based on multifactor regression analysis [9,10], was used to specify the capacity factors of amiloride and

Table 2
Fitted K' values for methylclothiazide (M) and amiloride (A) obtained by RSM

Methanol	pH	Measured K'		Fitted K'	
		M	A	M	A
10.0	2.8	0.780	2.250	0.863	2.162
10.0	3.5	1.080	2.770	0.887	3.090
10.0	4.2	0.870	4.390	0.893	4.175
30.0	2.8	0.390	0.522	0.276	0.4455
30.0	3.5	0.240	0.866	0.317	0.814
30.0	4.2	0.339	1.180	0.339	1.338
50.0	2.8	0.437	0.290	0.468	0.464
50.0	3.5	0.530	0.504	0.525	0.273
50.0	4.2	0.590	0.180	0.564	0.237
30.0	3.5	0.280	0.843	0.317	0.814
			R^2 =	0.9319	0.9833
			S.E. =	0.1016	0.2626

Standard error of estimation (S.E.).
Coefficients of multiple determination (R^2).

methyclothiazide to all combinations of pH values (2.8–4.2) and all combinations of methanol composition in the mobile phases (10–50%). To construct the model a minimum number of experiments has to be performed and the capacity factor has to be measured at the design points. These measurements are performed simultaneously according to the experimental design and used for modelling the response surface of every solute in the sample. A response surface can simultaneously represent two independent and one dependent variable when the mathematical relationship between the variables is known, or can be assumed. Independent variables were pH and methanol percentage in the mobile phase. The dependent variable was the capacity factor. Ten experimental data were fitted to a polynomial mathematical model by adjusting parameters until calculated values were in close

Table 3
Model fitting results for methylclothiazide (M) and amiloride (A)

M						
Parameter	1.2001	−0.0716	0.1413	0.0010	−0.0188	0.0012
S.E.	1.6672	0.0163	0.9582	0.0002	0.1357	0.0036
A						
Parameter	0.4029	−0.0606	0.72027	0.0022	0.1596	−0.0400
S.E.	4.3085	0.0421	2.4763	0.0004	0.3508	0.0094

ANOVA for RSM model fitting

M	Sum of Squares	df	F-Ratio	P-value
Model	0.56465	5	10.9*	0.02
Error	0.04129	4		
Lack of fit	0.00647	3	0.062**	
Pure error	0.03483	1		
Total	0.60593	9		
A				
Model	16.2346	5	47.1*	0.001
Error	0.27575	4		
Lack of fit	0.00334	3	0.004**	
Pure error	0.27 241	1		
Total	16.5103	9		

* $F_{5,4} = 6.25$.
** $F_{3,1} = 215.7$.

agreement with the experimental values [11]. The capacity factors can than be predicted at every pH and methanol percentage in mobile phase composition.

2. Experimental

2.1. Equipment

Separations were made on a Waters 5 μm μBondapak C-18 column (300 × 3.9 mm i.d. Waters Milford, MA). The injection volume was 10 μl, elution was performed at a flow rate of 1.5 ml min^{-1} and the column was maintained at ambient temperature. The absorbance was monitored at 286 nm. The mobile phase was 0.05 M aqueous solution of KH_2PO_4-methanol (pH adjusted with phosphoric acid).

2.2. Solvents and chemicals

Standards of amiloride and methylclothiazide and Lometazid® tablets were supplied by ICN Galenika (Belgrade, Serbia). The chromatographic internal standard was phenacetin. All the solvents used for the preparations of the mobile phase were HPLC grade and the mixtures were filtered and degassed before use.

2.3. Solutions

2.3.1. Internal standard solution

A 800 μg ml^{-1} solution of phenacetin in methanol was prepared.

2.4. Stock solution

About 10 mg of amiloride reference material and 5 mg of methyclothiazide reference material was precisely weighed, dissolved in internal standard solution and diluted to 100 ml with the same solvent to form a stock solution.

2.5. Standard solutions

Working standard solutions were prepared by dilution of a 4 ml volume of this solution to 10 ml with the internal standard solution.

2.6. Sample preparation

A finely powdered tablet was accurately transferred to a 100 ml calibrated flask and diluted to volume with internal standard solution. The mixture was sonicated for 5 min at room temperature and than centrifuged at 2500 × *g* for 5 min. The supernatant liquid was filtered through a 1.5 μm membrane filter. A 4 ml volume of this solution was diluted to 10 ml with the internal standard solution.

2.7. Data analysis

2.7.1. ANN simulator software

MS-Windows based artificial neural network simulator software, NNMODEL Version 1.404 (Neural Fusion) was used. Calculations were performed on a 486 personal computer.

2.8. Training data

The properties of the training data determine the number of inputs and output neurons. The behaviour of the capacity factors (K') of amiloride and methyclothiazide to the changes in pH and mobile phase composition were emulated using a network of two inputs (pH, and methanol (%)), one hidden layer and two outputs (K' for amiloride and methylclothiazide) (Table 1).

Neural networks were trained using different numbers of hidden neurons (2–20) and training cycles (5000–35 000). At the start of a training run, both weights and biases were initialized with random values. During training, modifications of the network weights and biases were made by back-propagation of the error. While the network was being optimized, the testing data (Table 1), were fed into the network to evaluate the trained network.

A problem in constructing the ANN was to find the optimal number of hidden neurons. Another problem was overfitting which occurs when the training data contain noise and the ANN is modelling the noise instead of the underlying features.

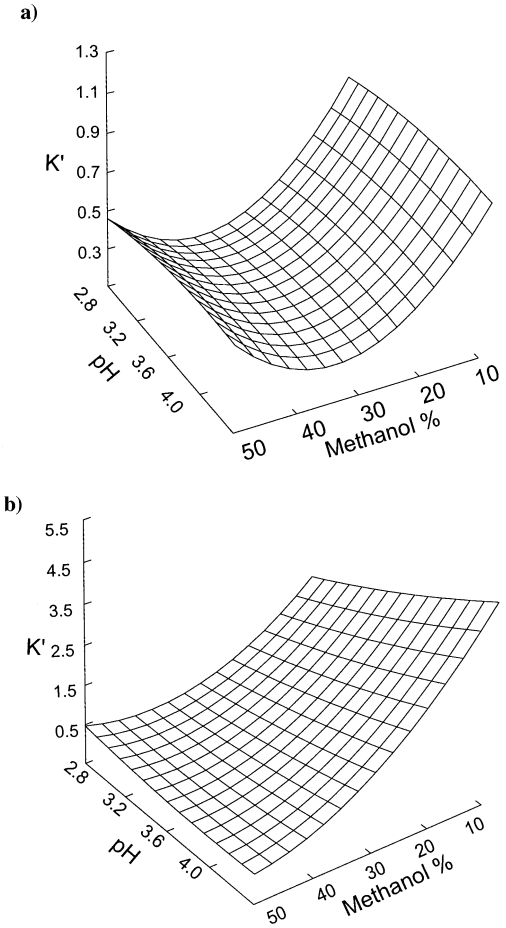

Fig. 4. Response surfaces for multifactor effect of pH and methanol % on a) methylclothiazide and (b) amiloride capacity factors generated by RSM.

3. Results and discussion

3.1. Network topologies

The number of connections in the network is dependent upon the number of neurons in the hidden layer. In the training phase the information of the training data is transformed to weight values of the connections. Therefore, the number of connections might have a significant effect on the network performance. Since there are no theoretical principles for choosing the proper network topology several structures were tested. The number of neurons in the hidden layer ranged from 2 to 20. Neurons were added to the hidden layer two at a time. The networks were trained and tested after each addition. After addition of the 20th hidden neurons, it became evident that more hidden neurons did not improve the generalization ability of the network (Fig. 2).

3.2. Training of the networks

To compare the predictive power of the neural network model, sum of squares errors were calculated and compared after each training cycle. The performance of the network on the testing set gives a reasonable estimate of the network prediction ability (Table 1).

The lowest testing SSE was obtained with 10 hidden neurones and at 15 000 training cycles. After 20 000 cycles extra training made the prediction ability worse and the test error began to increase. This effect is called overtraining or overfitting (Fig. 2).

3.3. Data validation

In order to test the predictions of the ANN and RSM five additional experiments were performed. The factor levels of the input variables were chosen so that they were within the range of the training experimental data. This operation is called interrogation of the model. The average error percentage [12] for each drug (ERR%) was used to examine the best generalization ability of the models,

$$\mathrm{ERR}\% = \sum_{i=1}^{n} \mathrm{abs}\left(1 - \frac{y_i}{t_i}\right) \times 100\%/n$$

where n is the number of validation sets for a drug, t_i is the measured capacity factor value and y_i denotes the predicted capacity factors by the model for a drug (Table 4).

The combined effect of pH and methanol percentage on the capacity factors generated by the best ANN model is presented in Fig. 3.

3.4. Multifactor regression model

Ten experiments were performed (pH 2.8, 3.5, 4.2 and methanol percentage of 10, 30, 50%, with

Table 4
Predicted K' values for methylclothiazide and amiloride

Methanol (%)	pH	Measured K'	Predicted K'		
			RSM	ANN*	ANN**
Predicted K' for methylclothiazide					
30	3	0.382	0.288	0.374	0.285
20	3	0.581	0.483	0.370	0.424
50	3	0.494	0.480	0.487	0.424
40	3.9	0.229	0.340	0.237	0.217
15	3.2	0.547	0.662	0.581	0.734
		ERR(%) =	0.2208	0.0995	0.1436
Predicted K' for amiloride					
30	3	0.819	0.562	0.657	0.564
20	3	0.891	1.268	1.034	0.953
50	3	0.198	0.470	0.168	0.132
40	3.9	1.037	0.495	0.812	0.932
15	3.2	1.564	2.008	1.677	1.678
		ERR (%) =	0.5707	0.1595	0.4035

* ANN with 10 hidden neurons at 15 000 training cycles.
** ANN with 10 hidden neurons at 20 000 training cycles.

a replicate at the mid point) and according to the experimental data (Table 2) model fitting methods gave the equation for the relationship between the methyclothiazide capacity factor and pH and mobile phase composition:

$$K' = 1.2001 - 0.0716x_1 + 0.1423x_2 + 0.0010x_1^2 - 0.0188x_2^2 - 0.0012x_1x_2$$

and for the amiloride capacity factor:

$$K' + 0.4029 - 0.0606x_1 + 0.7202x_2 + 0.0022x_1^2 + 0.1596x_2^2 - 0.0400x_1x_2$$

$x_1 = \text{methanol}(\%)$

$x_2 = \text{pH}$

The standard errors of the regression parameters and F-test for the significance of the regression and for the lack of fit of the model are given in Table 3. Predicted response surfaces drawn from the fitted equations are shown in Fig. 4.

3.5. Comparison of the best network and the regression model

To examine the predictive power of the regression model with the neural network model we compared experimental and predicted capacity factor values and ERR% for each drug (Table 4). These results show that, although the predictive power of the polynomial regression model is very good, it is possible to predict capacity factors more accurately using the neural networks model. Perhaps alternative regression models using various transforms would have given better prediction, but it would be time consuming to select the best from the infinite number of possible models. The network, however, constructed one model for all capacity factor solutes at all design points used for training. In this way the information is obtained more completely as the peak sequence in the different chromatograms can contribute to the model.

4. Conclusions

The neural network approach is a very powerful tool in HPLC method development. Capacity factors can be estimated with better or equal results to those obtained by the multiple regres-

sion technique. At the same time neural networks offer greater flexibility and potential than a non-linear multiple regression. We have selected to test the applicability of ANNs using the standard backpropagation algorithm for optimizing the composition of mobile phase, but there are several different training algorithms available, and their suitability should be studied in the future.

Acknowledgements

This work was supported by a 1996 FIP Fellowship (Foundation for education and research), awarded to S.A. Kustrin.

References

[1] P.J. Schoenmzkers, S. Vzn Molle, C.M.G. Hayes, L.G.M. Uunk, Effect of pH in reversed-phase liquid chromatography, Anal. Chim. Acta 250 (1991) 1–19.
[2] J.L. Glajch, J.J. Kirkland, Method development in high-performance liquid chromatography using retention mapping and experimental design techniques, J. Chromatogr. 485 (1989) 51–63.
[3] J. Zupan, J. Gasteiger, Neural network: a new method for solving chemical problems or just a passing phase?, Anal. Chim. Acta 248 (1991) 1–30.
[4] J.V.S. Gobburi, W.L. Shelver, W.H. Shelver, Application of artificial neural networks in the optimization of HPLC mobile-phase parameters, J. Liq. Chromatogr. 18 (10) (1995) 1957–1972.
[5] H.J. Metting, P.M.J. Coenegracht, Neural networks in high-performance liquid chromatography optimization: response surface modelling, J. Chromatogr. 728 (1–2) (1996) 47–53.
[6] R.J. Erb, Introduction to backpropagation neural network computing, Pharm. Res. 10 (1993) 165–170.
[7] D.L. Massart, B.G.M. Vanderginste, S.N. Deming, Y. Michotte, L. Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.
[8] L.J. Zivanovic, M. Vasiljevic, S. Agatonovic-Kustrin, M. Maksimovic, J. Pharm. Biamed. Analysis 14 (8–10) (1996) 1245–1250.
[9] S.N. Deming, S.L. Morgan, Experimental Design: A Chemometric Approach, vol. 3, Elsevier, Amsterdam, 1987.
[10] P.J. Schoenmakers, Optimization of Chromatographic Selectivity, A Guide to Method Development, Elsevier, Amsterdam, 1986.
[11] Statistical Graphic System Statgraphics 3.0, by Statistical Graphic Corporation.
[12] E. Murtoniemi, J. Yliruusi, P. Kinnunen, P. Merkku, K. Leiviska, Int. J. Pharmaceut. 108 (1994) 155–164.