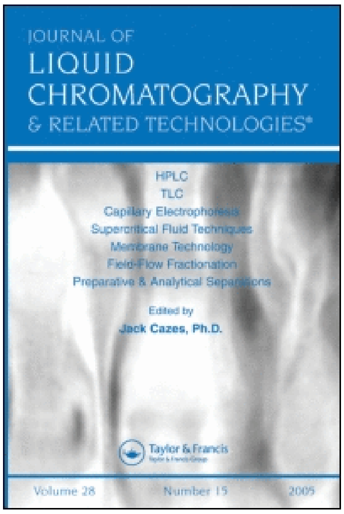

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

Application of Artificial Neural Networks in the Optimization of HPLC Mobile-Phase Parameters

Jogarao V. S. Gobburu[a]; Weilin L. Shelver[b]; William H. Shelver[a]
[a] Department of Pharmaceutical Sciences, North Dakota State University Fargo, North Dakota [b] Minneapolis Medical Research Foundation, Minneapolis, Minnesota

To cite this Article Gobburu, Jogarao V. S. , Shelver, Weilin L. and Shelver, William H.(1995) 'Application of Artificial Neural Networks in the Optimization of HPLC Mobile-Phase Parameters', Journal of Liquid Chromatography & Related Technologies, 18: 10, 1957 — 1972

To link to this Article: DOI: 10.1080/10826079508013953

URL: http://dx.doi.org/10.1080/10826079508013953

JOURNAL OF LIQUID CHROMATOGRAPHY, 18(10), 1957-1972 (1995)

APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN THE OPTIMIZATION OF HPLC MOBILE-PHASE PARAMETERS

JOGARAO V.S. GOBBURU[1], WEILIN L. SHELVER[2], AND WILLIAM H. SHELVER[1]
[1]*Department of Pharmaceutical Sciences*
North Dakota State University
Fargo, North Dakota 58105
[2]*Minneapolis Medical Research Foundation*
914 South 8th Street
Minneapolis, Minnesota 55404

ABSTRACT

The prediction capability of forward feed neural networks was tested c computer generated capacity factors. The capacity factors were simulate from equations reflecting the contribution of mobile phase changes in pH organic modifier concentration, and ion-pair concentration. Simulated dat allows an appropriate experimental design which assures the training of th network does not involve memorization but guarantees the network w generalize. The use of different mathematical forms to calculate th behaviour of capacity factor with changes in pH, methanol concentratior and ion-pair concentration permitted us to explore the capability of neur networks to fit a variety of curves. Each of the independent variables wer studied separately, and then in combination. The effect of variabl transformation played a very important role in effective training of th network. The neural network output equations were used to formulate nonlinear regression problem and the behaviour of this model was compare to the neural network system. When the neural network systems had onl sufficient processing units needed to solve the problem, nonlinear regressio models and neural networks arrived at identical solutions. When the networ

contained excessive neurons, nonlinear regression techniques were unstable, having high intraparameter correlations and showing matrix singularity.

INTRODUCTION

Mobile phase composition, in a reverse-phase ion-pair chromatographic system, plays a very vital role in the resolution of the various components in given sample. Consequently, optimization of the different parameters like - pH, organic modifier concentration, and ion-pair concentration is critical. Usually, the approach toward developing a reliable HPLC method for analysis is more or less intuitive, and is often time-consuming. Some of the crucial aspects in the selection of an appropriate mobile phase include complex nonlinear data manipulations, restrictions on the number of predictor parameters, and necessity for a large volume of data. The complexity of such an approach makes it ideal for the application of versatile data treatment techniques such as neural networks.

The use of artificial neural networks (ANN) to fit complex data is becoming popular in many scientific fields. ANN have been used to predict aqueous solubility of organic compounds [1], analyze quantitative structure-activity relationships [2], analyze NMR data [3], predict protein secondary structure [4], analyze complex pharmacodynamic data [5], and to model nonlinear pharmacokinetic data [6]. The principles of neural networks have been well described in textbooks [7 8], as well as reviewed in journals [9 10]. The process by which neural network is applied to a problem involves selection of input parameters, deciding the desired output, and selection of an appropriate network. The network should be as simple as will adequately reproduce the training data. We have elected to test the applicability of neural networks using the standard backpropagation algorithm for optimizing the composition of the mobile phase. The effect of pH, organic modifier content and ion-pair concentration on capacity factor was simulated using appropriate mathematical relations.

METHODS

A neural network program based on the popular back-propagation algorithm was employed in our study. All the networks consisted of three layers - input, hidden and output layers. Input neurons do not process the data but, feed it to the neurons in the higher layers. The hidden and the output neurons remap the data in a more classifiable form. This is done by transforming the sum of the products of the weights and the corresponding inputs using a flexible function. Our system used the sigmoid function as the squashing function. The connectivity of parallel processing system depended on the complexity of the relationship under investigation. The chosen network was trained with the generated data which was scaled between 0.05 and 1.00.

The effect of each of the mobile phase parameters was studied individually as well as in combination. In the individual cases, either pH, methanol concentration, or ion-pair concentration served as the input and the capacity factor as the target. In the combination case all three parameters were inputs with capacity factor as the output. In almost all the cases variable transformation yielded in lower total sum of squares (tss).

Organic Modifier Concentration

Relevant data was generated from an exponential function which reasonably describes the behaviour of capacity factor with changes in the concentration of methanol. The mathematical relation is depicted in eq(1), where $k_1 = 10$,

$$K' = k_1 e^{k_2 * [methanol]} \qquad (1)$$

$k_2 = 0.202$, and methanol concentration ranged from 50 to 100 per cent.

Ion-Pair Concentration

Data was generated from the hyperbolic equation which mimics the effect of the concentration of the ion-pair reagent on the capacity factor, shown in eq(2),

$$K' = \frac{k_1}{1 + \frac{k_2}{[amine]}} \qquad (2)$$

where $k_1 = 18.94$, $k_2 = 29.86$ and the amine concentration was between 0.01and 40.

pH

The dependence of capacity factor on hydrogen ion concentration was described with eq(3), where Ka = 1.4E-8, k_1 = 0.29, and k_2 = 8.34.

$$K' = k_1 + \frac{k_2 \cdot [H^+]}{1 + \frac{[H^+]}{k_a}} \qquad (3)$$

Combination

The dependence of capacity factor on all of the previous mobile phase components was described by the eq(4). Three amine concentrations (5, 30, 55), twelve pH values from 2 to 8, and ten values of methanol concentrations

$$K' = \frac{\frac{k_1 e^{-k_2 \cdot \ln[methanol]}}{1 + \frac{k_3}{[amine]}} + k_4 e^{-k_4 \cdot \ln[methanol]} \cdot 10^{-pH}}{\frac{10^{-pH}}{k_a}} \qquad (4)$$

were used to generate the training set.

Nonlinear Regression

Nonlinear regression analysis was conducted using SAS[11]. Upon training the data to a particular network, the system output equations were derived to formulate a nonlinear least squares parameter estimation problem. For example, the output equation of system consisting of one input, no hidden, and one output is shown in eq(5). In eq(5), Wij stands for the weight of the connection

$$Output = \frac{1}{1 + e^{(-W_{ij} \cdot I + b_j)}} \qquad (5)$$

between the neurons i and j, I is the input to the neuron j, and bj stands for the bias associated with neuron j.

RESULTS AND DISCUSSION

Variable transformation is one of the important aspects to be considered in the improvement of the training ability of a neural network. Figure 1a shows the changes in the tss along the course of training with the hydrogen ion concentration as the input and capacity factor as the output variables. As shown (Fig 1a), the tss could not reach a minimum value but instead an oscillation is observed. In order to avoid this problem, we transformed the hydrogen ion concentration to pH, and the behaviour of tss is depicted in Figure 1b, which shows a steady tss value of ~0.04.

The behaviour of the capacity factor to the changes in pH was well emulated by a one input (pH), one hidden, and one output (K') network. Figure 2 shows a good correlation between the neural trained output and the generated

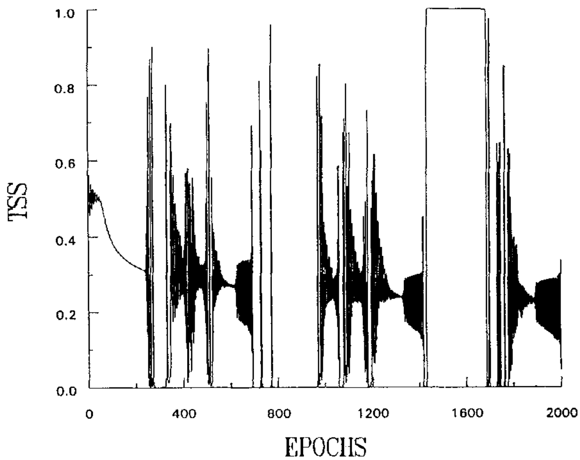

Figure 1a. Plot of the number of epochs and the corresponding tss when hydrogen ion concentration was the input and K' was the output.

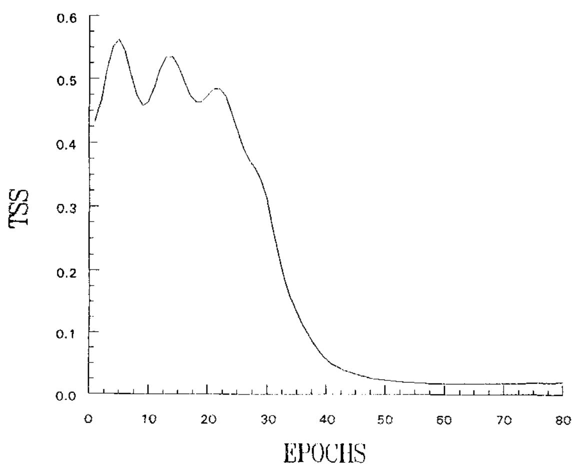

Figure 1b. Plot of the number of epochs and the corresponding tss when pH was the input and K' was the target.

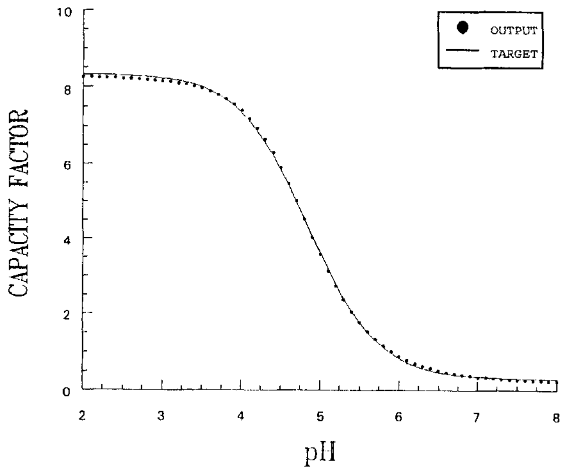

Figure 2. Plot of neural network trained output and calculated capacity factors reflecting the changes in pH of the mobile phase (r=0.9999).

Table 1
SAS Estimates and the Neural Network Derived Parameters Associated with Effect of pH of the Mobile Phase.

PARAMETER	NEURAL NETWORK	SAS
w1	-6.6352	-6.6082
w2	9.1653	9.2044
b1	2.8344	2.8170
b2	-3.3941	-3.400

K' (r = 0.9999). As seen in Figure 2, the magnitude of change due to alterations in pH is maximum in the pH 4-6 range, the capacity factor plateaus in the either extremes. Upon successful training, the system output equations were derived and nonlinear regression was applied to estimate the parameters which include the various weights and biases. Table 1 shows the similarity between SAS estimates and the neural network parameters ('wi' denotes the ith weight and 'bi' denotes ith bias).

The exponential curve describing the effect of change in methanol concentration in the mobile phase was emulated with a network consisting of no hidden neurons. The logarithmic transformation of capacity factor resulted in lower tss. The neural network trained output and the calculated capacity factors with methanol concentrations ranging from 50 to 100% are depicted in Figure 3, (r = 0.9935). The SAS estimates from the output equations and the neural network parameters are presented in Table 2.

The curvature governing the effect of the ion-pair reagent in the mobile phase on the capacity factor was well reproduced with a neural network system consisting of one input (amine concentration), no hidden neuron and one output (K'). Figure 4 shows the correlation between the neural network output and the calculated K' (r = 0.9925). The weights and biases of the neural network and the SAS nonlinear estimates are presented in Table 3.

The complex nature of the combined effects of all three mobile phase variables could be emulated by a three hidden neuron network. Figure 5 depicts

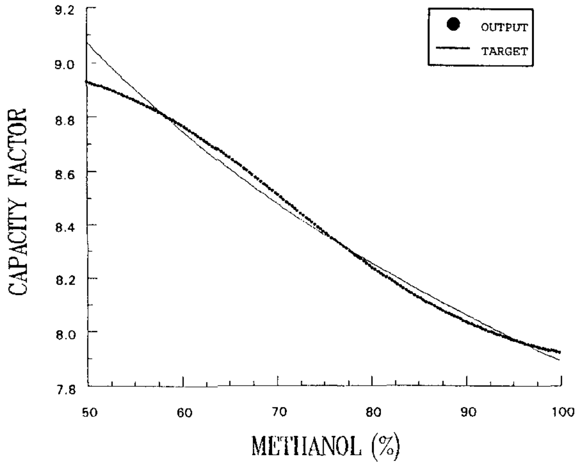

Figure 3. Plot of neural network trained output and calculated capacity factors reflecting the changes in methanol concentration in the mobile phase. (r=0.9935)

Table 2

SAS Estimates and Neural Network Derived Parameters Associated with the Effect of Methanol Concentration in the Mobile Phase.

PARAMETERS	NEURAL NETWORK	SAS
w1	-4.862	-4.862
b1	2.682	2.682

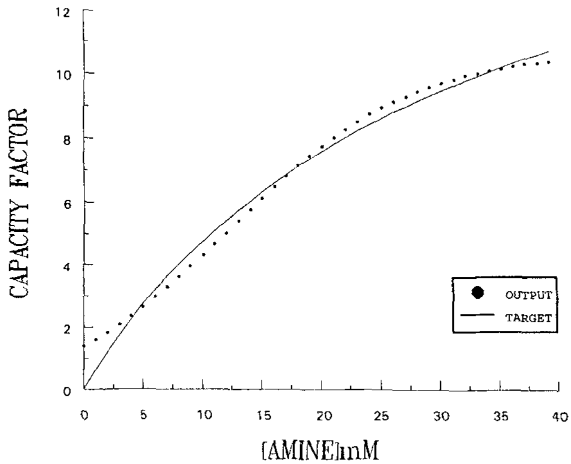

Figure 4. Plot of neural network trained output and calculated capacity factors reflecting the changes in the ion-pair reagent in the mobile phase. (r=0.9925)

the neural network output and the calculated K' (r = 0.9999). Table 4 presents a comparison between the SAS estimates and neural network parameters.

An important aspect that needs to be focussed is the feasibility of developing an ANN system that has potential practical significance. The attributes of such a system would be to generalize the solution with sparse data and robustness toward any noise in the data set. We explored this issue by training the neural network with only few data points and test the system for the other untrained points. The case of combined effects on the capacity factor was considered, which would be a rigorous test for the ability of ANN to recognize the pattern governing the relation between the input and target

Table 3
SAS and Neural Network Derived Parameters Associated with the Effect of Ion-pair Reagent Concentration in the Mobile Phase.

PARAMETER	NEURAL NETWORK	SAS
w1	5.3330	5.3333
b1	-1.8441	-1.8441

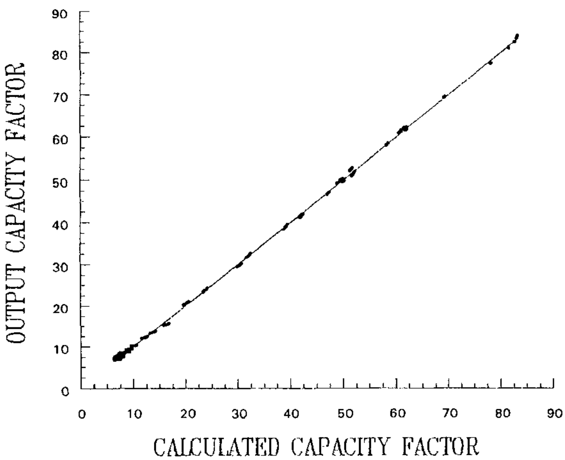

Figure 5. Plot of neural network trained output and calculated capacity factors reflecting the effects of pH, methanol and amine concentrations in the mobile phase (r=0.9999).

Table 4

SAS and Neural Network Derived Parameters Associated with the Combined Effects of pH, Methanol, and Ion-pair Reagent in the Mobile Phase

PARAMETER	NEURAL NETWORK	SAS
w1	-0.6826	-0.6727
w2	-8.7640	-8.7600
w3	-0.0370	-0.0357
w4	10.2680	10.0138
w5	-5.3810	-5.4455
w6	0.0507	0.0465
w7	2.7207	2.7164
w8	1.8429	1.8874
w9	-0.0386	-0.0337
w10	6.4646	6.4559
w11	2.0858	2.0791
w12	2.3476	2.3368
b1	3.89433	3.8900
b2	3.3506	3.4054
b3	-2.8092	-2.8263
b4	-4.2470	-4.2408

variables with minimum information. Neural networks were trained with as low as 8, 12, and 16 points in three different experiments and using the optimized system parameters we tested for 429 points. In order to examine the robustness of the ANN, we also trained the network with data adulterated with 30% error. Neural networks were successful in predicting the capacity factors

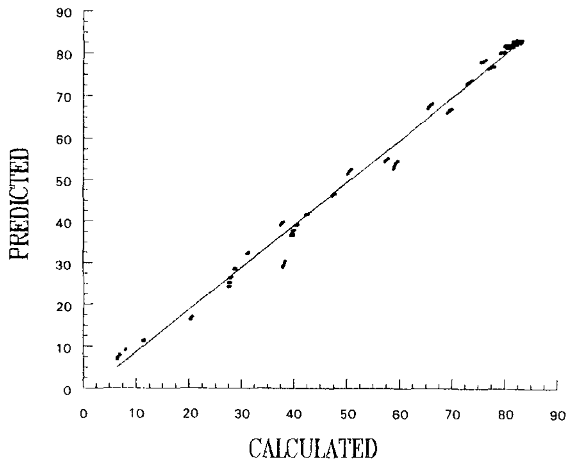

Figure 6. Plot of the neural network predicted and the actual calculated capacity factors when the network was trained with only 16 points and tested for 429 points (r=0.9969).

Table 5
Correlation Coefficients of the Neural Network's Predictions Trained with Sparse Data

NUMBER OF DATA POINTS TRAINED	CORRELATION COEFFICIENT (r)
8	0.9037
8[a]	0.9077
12	0.9354
16	0.9969
16[a]	0.9658

[a] 30% error was included in the training set

accurately. Figure 6 depicts the actual and the predicted capacity factors reflecting the changes in all three mobile phase parameters, when the network was trained with 16 points. This proves the power of ANN, and their potential for practical applications. The summary of results from these tests is shown in Table 5.

CONCLUSIONS

Neural networks prove to be very powerful in elucidating the individual as well as combined effects of the various mobile phase variables considered on the capacity factor of a chromatographic method. It is also shown that neural networks yield similar results as nonlinear regression technique. At the same time neural networks offer greater flexibility and potential than nonlinear regression techniques in that ANN can generalize the pattern even with few data points. Data points fewer than the parameters to be estimated makes the application of the standard nonlinear least squares awkward.

As investigated in this article, parallel distributed processing systems offer a great advantage over the traditional approaches in that model specification is not necessary in neural networks, which is otherwise quite cumbersome. Other important considerations include emulating patterns with sparse data, which economizes the number of experiments and robustness of the technique.

REFERENCES

1 N. Bodor, M. Huang and A. Harget. Neural Network Studies. 4. An Extended Study of the Aqueous Solubility of Organic Compounds. *Int. J. Quant. Chem.* : Quantum Chem. Symposium. 26:853-867 (1992).

2 T. Aoyoma, Y. Suzuki and H. Ichikawa. Neural Networks Applied to Quantitative Structure-Activity Relationship Analysis. *J. Med. Chem.* 33:2583-2590 (1990).

3 B. Meyer, T. Hansen, D. Nute, P. Albersheim, A. Darvill, W. York and J. Sellers. Identification of the ^{1}H-NMR Spectra of Complex Oligosaccharides with Artificial Neural Networks. *Science*. 251:542-544 (1991).

4 N. Qian and T.J. Sejnowski. Predicting the Secondary Structure of Globular Proteins Using Neural Network Models. *J. Mol. Biol.* **202**:865-884 (1988).

5 P. Veng-Pedersen and N.B. Modi. Application of Neural Networks to Pharmacodynamics. *J. Pharm. Sci.* **82**:918-926 (1993).

6 Jogarao V.S. Gobburu and W.H. Shelver. Neuronets: A Solution to Nonlinear Pharmacokinetic Models. *Presented at IX-AAPS, San Diego, CA.* (1994).

7 J.L. McClelland and D.E. Rumelhart. *Explorations in Parallel Distributed Processing*. MIT Press, Cambridge, MA (1988).

8 P.D. Wasserman. *Neural Computing Theory and Practice.* Van Nostrand Reinhold, NY (1989).

9 R.J. Erb. Introduction to Backpropagation Neural Network Computing. *Pharm. Res.* **10**:165-170 (1993)

10 P.A. Jansson. Neural Networks: An Overview. *Anal. Chem.* **63**:357A-362A (1991).

11. SAS (r) *Proprietary software release 6.04.* SAS Institute Inc., Cary, NC, USA.

Received: December 20, 1994
Accepted: January 5, 1995