Journal of Chromatography A, 1332 (2014) 14–20

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Prediction of the competitive adsorption isotherms of 2-phenylethanol and 3-phenylpropanol by artificial neural networks

Xiuhong Wu [a,*], Shaoyan Wang [a], Renzhuang Zhang [a], Zhiming Gao [b]

[a] Department of Chemical and Engineering, University of Science and Technology of Liaoning, Anshan, Liaoning 114051, China
[b] School of Material Science and Engineering, Tianjin University, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 11 December 2013
Received in revised form 14 January 2014
Accepted 16 January 2014
Available online 24 January 2014

Keywords:
Adsorption isotherm
Artificial neural network
Binary component
Fitting
Modeling

ABSTRACT

Artificial neural networks (ANNs) were regarded as data-mapping networks with strong nonlinear fitting abilities. A 2-6-2 network was used to determine the competitive adsorption isotherm of 2-phenylethanol (PE) and 3-phenylpropanol (PP). The ANN results were forms of data mapping rather than theoretical mathematical model. The ANN architecture was established after training with a set of experimental data. The established ANN was applied to predict the adsorption isotherms of PE and PP. The selection of parameters for the ANN was discussed. The results indicate that ANN has excellent potential for use in non-linear chromatography for the prediction of adsorption isotherms.

1. Introduction

ANN has played an important role in non-linear mapping for several decades, partially due to its similarity to biological neural networks. ANN is a multi-layer network consisting of input, hidden and output layers. Weights connect the input layer to the hidden layer and the hidden layer to the output layer. Data from one layer can be transferred to another layer using a transform function. In most cases, a back propagation (BP) algorithm is used to extrapolate the experimental data to a target value, after which the ANN structure is established. The ANN network can be used in various areas, such as prediction [1,2], modeling [3,4], optimization [5,6], process control [7] and classification [8,9], especially in analytical chemistry, chemical engineering, environmental science and bioscience. Zupan et al. recently reviewed the wide use of ANN in chemistry [10].

Adsorption isotherms are the basis of non-linear chromatography. Over a long period, adsorption isotherms are expressible using theoretical mathematical models, such as the Langmuir [11], Freundlich [12], bi-Langmuir [13,14], Toth [15,16] and quadratic [17,18] models. To select a suitable model for a set of experimental data, the system used, quantities to determine and a physical sense of the mathematical model must be considered. Thus, the selection of an optimal mathematical model is limited. Langmuir is the model most often applied for fitting due to its simplicity and effectiveness. However, in most cases, larger deviation occurs for higher concentration ranges. Thus, modifications of these models are widely used for the determination of adsorption isotherms. These "hard" models are described exactly by equations and physical–chemical constants. ANNs, on the other hand, are classified as "soft" models because they use an array of simple activation units linked by weighted connections rather than formulae or equations [19].

In the present work, ANN was employed to predict the adsorption isotherms of PE and PP. Competitive adsorption experimental data were trained to establish an ANN architecture for predicting the competitive adsorption isotherms. The results were further compared with the modified competitive Langmuir model. The details of the selection of the ANN parameters were also discussed. A limited number of studies have been reported regarding such applications [20]. One of the objectives of this paper was to verify the nonlinear fitting ability of ANN for nonlinear chromatography.

2. Basic principles of ANN

2.1. Network of ANN

A multi-layer forward network using a back propagation (BP) algorithm featuring back error propagation training is one of the

* Corresponding author. Tel.: +86 412 5929475; fax: +86 412 5929627.
E-mail address: wxh93@163.com (X. Wu).

0021-9673/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.chroma.2014.01.049

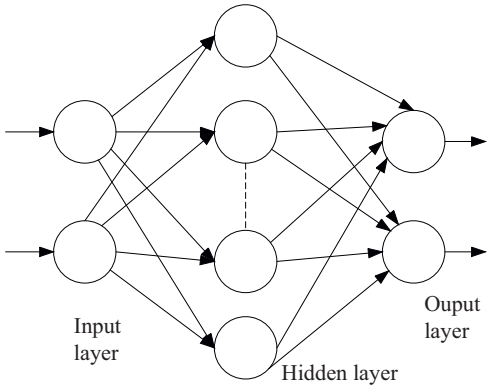

Fig. 1. Schematic of a feed-forward neural network architecture.

most often used networks. BP networks can store large amounts of input and output relationships for mapping and do not need mathematical models to describe mapping relationships. The rule of learning that is generally adopted is the steepest descent method. The weights were adjusted from the back direction to minimize the sum of the squares of errors in the network. A BP network consists of an input layer, hidden layer and output layer, as shown in Fig. 1.

The basic element of the network is a neuron, which imitates the three basic, yet most important functions of a biological neuron: to weight, sum and transfer. The net input of neuron is:

$$S_j = \sum_{i=1}^{n} w_{ji}x_i + b_j = W_jX + b_j \quad (1)$$

The output y_j of neuron j is then given by transferring the net S_j through the transfer function $f(\cdot)$:

$$y_j = f(S_j) = f\left(\sum_{i=0}^{n} w_{ji}x_i\right) = F(W_jX) \quad (2)$$

Here $f(\cdot)$ is a monotonically increasing function and must be bounded. As for biological cells, it is impossible for a transferred signal to increase indefinitely; there must a maximum value.

The BP algorithm consists of two processes: the forward calculation of data flows (forward propagation) and the back propagation of error signals. Forward propagation refers to propagation from the input layer to hidden layer and then to output layer. The state of the neuron in the present layer only affects the state of the neuron in the next layer. If unexpected results obtained, error signal back propagation activates. These two processes are alternatively carried out. A gradient descent of the error function is performed in weight vector space to find a group of vectors of such a weight as to dynamically minimize the error function of the network. The processes of drawing information and memory are then performed.

The network is valuated by the mean square of errors (MSE): $E = (1/2)\sum_{p=1}^{p}\sum_{j=1}^{m}(t_j - y_j)^2$, which corresponds to the error of the target value t_j compared with the output value y_j of all the sample of P. When the MSE is satisfactory, the ANN architecture is established.

3. Experimental

A Shimazu LC-10A liquid chromatograph was performed with a Shimadzu SPD-10AVP pump equipped with a UV and RI detector and a computer data station were used for all studies. The mobile phase was composed of methanol:water (50:50, v/v). HPLC grade methanol was obtained from Tianjin Yong Da Chemical Ltd. The water was prepared with a Wa Haha device obtained from Hangzhou Wa Haha Purification Water Ltd. (Hangzhou, China). Samples of 3-phenylpropanol and 2-phenylethanol were obtained J&K CHEMICA (Johnson Mattey Company, USA) and Beijing Jinlong Chemical LTD, respectively. All the solvents were filtered using a PTFE membrane of pore size 0.45 μm obtained from Tianjin Yong Da Chemical Ltd., China prior to use. Uracil used as the non-retained compound and acted as marker. A 4.6 mm × 150 mm stainless steel Extend^{TM} 5 micron C_{18} column (Agilent Technologies, USA) was used for all chromatographic studies. The total porosity measured by injecting uracil and was determined as $\varepsilon_t = 0.651$. The column hold-up time (t_m) was 2.077 min, and the efficiency of the column at a flow rate of 0.80 mL min^{-1} was $N = 2880$ theoretical plates. Stock solutions of 2-phenylethanol (PE) and 3-phenylpropanol (PP) were prepared by dissolving 12.2334 g and 13.6283 g of these compounds respectively in 100 mL mobile phase to provide volumetric solutions whose concentrations were 1.001 mol L^{-1} (PE) and 1.003 mol L^{-1} (PP).

4. Results and discussion

4.1. ANN architecture

The number of layers, number of neurons in different layers, transfer function, initial weight and parameters such as the learning rate and momentum constant are necessary to form the ANN architecture. There are some rules for establishing an ANN [21]; however, in most cases, ANNs are formed based on experience and these rules are trivial. In our work, an ANN is used for predicting the competitive adsorption isotherms of PE and PP. The neurons in the input layer contain the concentrations of PE and PP in the mobile phase, whereas the output neurons contain the concentrations of PE and PP adsorbed in the stationary phase. Typically, a three-layer BP network can perform any mapping from dimension n to dimension m [5]. One hidden layer is utilized in this study. The transfer function used in this work is the sigmoid function $f(x) = 1/(1 + \exp(-x))$, which limits the output results to between 0 and 1.

4.2. Number of neurons in the hidden layer

For forward networks, the number of neuron in the hidden layer is a crucial parameter. If the number is too high or too low, the generalization ability may be compromised. Thus, it is important to choose an appropriate number of neurons in the hidden layer. However, the selection of this number is relatively complex. Generally, fewer nodes are preferred to make the network as simple as possible.

The adsorption isotherms of PE and PP have been determined by frontal analysis (FA) and published elsewhere [22]. The experimental data were fitted by competitive Langmiur: $q_1 = 7.26c_1/(1 + 5.08c_1 + 8.34c_2)$, $q_2 = 11.07c_2/(1 + 5.08c_1 + 8.34c_2)$; modified Freundlich equation: $q_1 = 1.43((7.52c_1)^{1/\gamma}/(1 + (7.52c_1)^{1/\gamma} + (11.56c_2)^{1/\gamma}))$, $q_2 = 1.33((11.56c_2)^{1/\gamma}/(1 + (7.52c_1)^{1/\gamma} + (11.56c_2)^{1/\gamma}))$ ($\gamma = 0.95$) and multicomponent Sips: $q_1 = 1.43(7.52c_1/(1 + 7.52c_1 + 11.56c_2))^{1/\gamma}$, $q_2 = 1.33(11.56c_2/(1 + 7.52c_1 + 11.56c_2))^{1/\gamma}$ ($\gamma = 0.91$), respectively. The q_s of PE and PP was determined 1.43 and 1.33 mol L^{-1}, respectively. The experimental data were used for the training and prediction of ANN. The program used was programmed by our group and compiled by Turbo C language.

When the number of hidden nodes was 6 or 9, the average relative error (RE) of each component was below 0.50%, with the lowest errors occurring for 6 nodes. To avoid the problem of over fitting caused by complex network structure, the number of nodes was chosen as 6. Thus, the ANN network was 2-6-2.