The European Journal of Mineral Processing and Environmental Protection
Vol.3, No.1, 1303-0868, 2003, pp. 119-127

ejmp&ep

Prediction of two-metal biosorption equilibria using a neural network

K.H. Chu*

Department of Chemical and Process Engineering, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

Received 30 June2002; accepted 15 January 2003

ABSTRACT

A feedforward neural network model with a single hidden layer was used to correlate and predict biosorption equilibrium data in a binary metal system. Experimental data on the biosorption of Fe(III) and Cr(VI) by the microalga *Chlorella vulgaris* reported in the literature was used to assess the performance of the neural network. It was demonstrated that the neural network approach was significantly more accurate than the traditional modeling approach based on Langmuir-type models. To assess the predictive capability of the neural network model, the network was trained using a subset of available data. The suitably trained neural network was found to be capable of predicting fresh data not belonging to the training set. However, training data should be selected carefully if the best results are to be achieved.

Keywords: Artificial neural network; Biosorption; Equilibrium isotherm; Modeling

1. INTRODUCTION

Metal contaminants at low concentrations are difficult to remove from aqueous solutions. Chemical precipitation, reverse osmosis, and many other methods become inefficient when metals are present in trace quantities. Adsorption is one of the few alternatives available for such situations. Recent research has focused on the use of biomass as metal biosorbent. Various microbial biomass including bacteria, fungi, and algae as well as plant biomass are typical biological solids that exhibit surface specificity toward heavy metal ions. Most of the studies have focused on the biosorption behavior of biomass in single metal systems. In most practical cases contaminated water will contain a number of metal ions, and competitive uptake, which may have significant influences on the biosorption performance, is likely. Hence, data obtained from single metal systems is of limited practical use and reliable data on multimetal systems is essential for process design, scale-up, and optimization.

Due to the numerous possible combinations of metal ions and biomass species in multimetal biosorption systems, the experimental measurement of multimetal equilibrium data is tedious and time consuming. As a result, only a limited number of multimetal studies have been reported in the literature (Chong and Volesky, 1995; Aksu *et al.*, 1997; Chang and Chen, 1998; Kaewsarn and Yu, 1999; Puranik and Paknikar, 1999; Sanchez *et al.*, 1999; Sag *et al.*, 2000; Klimmer *et al.*, 2001; Lee *et al.*, 2002).

Mathematical models with predictive capability offer an alternative way of generating multicomponent equilibrium data within a limited amount of time and with a limited amount of experiments. The existing modeling techniques may be divided into two categories. The first involves the use of empirical or semi-empirical models such as Langmuir and Freundlich isotherms, while the second is based on the use of mechanistic models such as ion exchange.

* Corresponding author. E-mail: khim.chu@canterbury.ac.nz

Models based on the adaptations of the single component Langmuir and Freundlich isotherms have been used by various researchers to describe multimetal biosorption data (Chong and Volesky, 1995; Aksu *et al.*, 1997; Chang and Chen, 1998; Puranik and Paknikar, 1999; Sanchez *et al.*, 1999; Sag *et al.*, 2000). Models providing mechanistic characterization of the ion exchange phenomenon have also been used to quantify multimetal uptake (Schiewer and Volesky, 1996). However, most of these studies have treated the models as mathematical functions for *correlating* the multimetal equilibrium data where parameter estimation was based on all available experimental data. The ability of these models to *predict* unseen data (data that was not used for curve fitting/parameter estimation) has not been adequately established.

This contribution describes a predictive modeling approach based on neural networks which offers a useful alternative to the popular correlative modeling methodology based on Langmuir-type models. A neural network model is appropriate for this kind of application due to its inherent capability to capture nonlinear relationships effectively. A number of previous results have demonstrated the potential applicability of neural networks to the field of multicomponent adsorption. For example, Syu *et al.* (1993) reported the application of neural networks to predict the adsorption from mixtures containing organic acids and alcohols. In another application, Yang *et al.* (1996) developed a neural network to predict the binary adsorption of various organic solutes. More recently, neural network models were used to predict the adsorption equilibria of binary vapor mixtures (Carsky and Do, 1999). By comparison, this work utilizes a neural network model to predict the biosorption equilibrium data of two heavy metal ions. Experimental data on the binary uptake of Fe(III) and Cr(VI) by the microalga *Chlorella vulgaris* reported by Aksu *et al.* (1997) was used to assess the predictive capability of the neural network model.

2. NEURAL NETWORKS

Neural networks are purely data driven models and have been proven to be universal approximators (Hornik *et al.*, 1989). In recent years there has been considerable interest in employing neural networks to model chemical and biochemical processes due to their ability to identify complex input-output relationships (Baughman and Liu, 1995; Bulsari, 1995; Sato *et al.*, 1999). A typical neural network has an input layer, an output layer, and hidden layers. A number of interconnected processing elements or neurons are logically organized in these layers. The neurons in the hidden layer which are linked to the neurons in the preceding layer and the succeeding layer by adjustable weights enable the network to compute complex associations between inputs and outputs. The process of determining these weights is known as training. Basically, the network is taught to model an input-output relationship during a supervised training procedure by using series of input and associated output data.

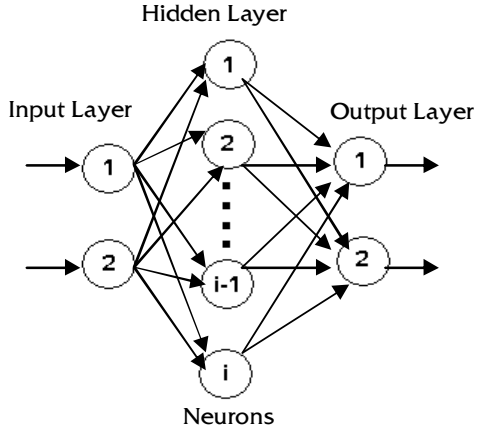


Figure 1. Neural network structure

Figure 1 shows the general layout of the three-layer network used in this study. In this structure there are two neurons in the input layer, *i* neurons in the hidden layer, and two neurons in the output layer. Also, the neurons in the hidden and output layers are associated with a bias neuron (not shown). Inside each hidden and output neuron, a weighted sum of the inputs is calculated, a bias weight is added, and this value, called *z*, is then processed using an activation function prior to being fed as an input to the next layer. The sigmoid function defined by Eq. (1) is used as the activation function:

$$F(z) = \frac{1}{1+\exp(-z)} \tag{1}$$

Training constitutes the first stage in the implementation of a neural network designed to identify the relationship between the input variables and the output variables of a given process. The training of a network with the above topology is achieved by adjusting the weights of the neurons through an iterative algorithm that minimizes the root mean square (RMS) error between the network-predicted outputs and actual data defined by Eq. (2):

$$\text{RMS error} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(u_{\mathrm{k}} - v_{\mathrm{k}}\right)^2} \tag{2}$$

where *u* represents experimental data, *v* refers to model prediction, and *n* is the number of data points. The popular back-propagation algorithm was used in this study to minimize the RMS error for a given set of input-output data. The neural network model was developed using the standard C/C + + computer language under the Windows NT environment.

3. RESULTS AND DISCUSSION

Aksu *et al.* (1997) reported binary equilibrium data for Fe(III) and Cr(VI) biosorption on the microalga *C. vulgaris*. The results obtained from batch equilibrium experiments are listed in Table 1. The experimental conditions for the batch experiments are as follows: temperature=25°C; pH=2; and biomass concentration (W) = 1g/l. The liquid phase equilibrium concentrations of Fe(III) and Cr(VI) (C_e) listed in Table 1 were estimated from the original data given by Aksu *et al.* (1997) using the following mass balance equation:

$$C_{\mathrm{e}} = C_{\mathrm{o}} - Wq_{\mathrm{e}} \tag{3}$$

where C_o and q_e are the liquid phase initial concentration and biosorbent phase equilibrium concentration of each metal. The results in Table 1 indicate that the uptake of one metal was suppressed in the presence of the other. Mutual suppression of uptake is commonly observed in multimetal systems due to competition between the metal ions for the same binding sites.

In the modeling approach of this paper, the first step was to identify an appropriate neural network topology. The neural network adopted for this study consisted of two input neurons representing the solution phase equilibrium concentrations ($C_{e,Fe}$ and $C_{e,Cr}$) and two output neurons representing the biosorbent phase equilibrium concentrations ($q_{e,Fe}$ and $q_{e,Cr}$). In addition, a single hidden layer was selected because feedforward neural networks with one hidden layer containing a sufficiently large number of hidden neurons have been shown to be capable of providing accurate approximations to any continuous nonlinear function (Hornik, 1991). Unfortunately, there is currently no universal guideline for determining the optimal number of hidden neurons. The selection of the number of neurons in the hidden layer is often the result of empirical rules combined with trial and error. The neural network with a 2-*i*-2 structure depicted in Figure 1, where *i* is the number of neurons in the hidden layer, was trained using the measured data in Table 1 until a specified maximum number of epochs or training cycles was reached. The number of neurons *i* was changed from 2 to 10 in increments of 1. The

performance of the neural network was assessed in terms of the RMS error given by Eq. (2). The final RMS errors are plotted against the number of neurons in the hidden layer in Figure 2. It can be seen that neural networks with hidden neurons exceeding five appeared to yield little advantage in terms of the reduction of the RMS error for Fe(III) and Cr(VI). Because too many neurons require a relatively large computation time and can introduce overfitting of data which can lead to poor predictive performance, a simple neural network with a 2-5-2 topology was used in this study.

Table 1
Experimental data for simultaneous biosorption of Fe(III) and Cr(VI) on *C. vulgaris* (Aksu *et al.*, 1997)

Data pattern	$C_{o,Fe}$ (mg/l)	$C_{o,Cr}$ (mg/l)	$C_{e,Fe}$ (mg/l)	$C_{e,Cr}$ (mg/l)	$q_{e,Fe}$ (mg/g)	$q_{e,Cr}$ (mg/g)
1.1	25.0	25.2	18.1	17.7	6.9	7.5
1.2	47.5	27.5	38.0	20.4	9.5	7.1
1.3	77.5	24.9	66.1	18.4	11.4	6.5
1.4	102.5	25.4	89.5	20.2	13.0	5.2
1.5	150.0	24.5	133.6	19.7	16.4	4.8
1.6	252.5	24.3	234.5	20.5	18.0	3.8
2.1	27.5	49.9	21.5	31.8	6.0	18.1
2.2	50.0	49.4	41.0	32.5	9.0	16.9
2.3	75.0	50.6	64.3	34.8	10.7	15.8
2.4	97.5	50.3	85.5	35.7	12.0	14.6
2.5	147.5	49.7	133.4	37.3	14.1	12.4
2.6	250.0	50.1	231.3	40.0	18.7	10.1
3.1	22.5	100.3	19.3	76.3	3.2	24.0
3.2	47.5	99.8	41.7	77.1	5.8	22.7
3.3	77.5	99.6	69.3	78.4	8.2	21.2
3.4	100.0	100.0	90.0	79.7	10.0	20.3
3.5	150.0	99.4	137.6	80.9	12.4	18.5
3.6	247.5	99.2	232.0	82.2	15.5	17.0
4.1	27.5	149.6	24.6	122.9	2.9	26.7
4.2	52.5	150.2	47.6	124.2	4.9	26.0
4.3	72.5	149.1	65.8	123.9	6.7	25.2
4.4	97.5	150.7	89.1	126.2	8.4	24.5
4.5	152.5	148.5	141.7	125.2	10.8	23.3
4.6	252.5	151.3	238.9	129.8	13.6	21.5

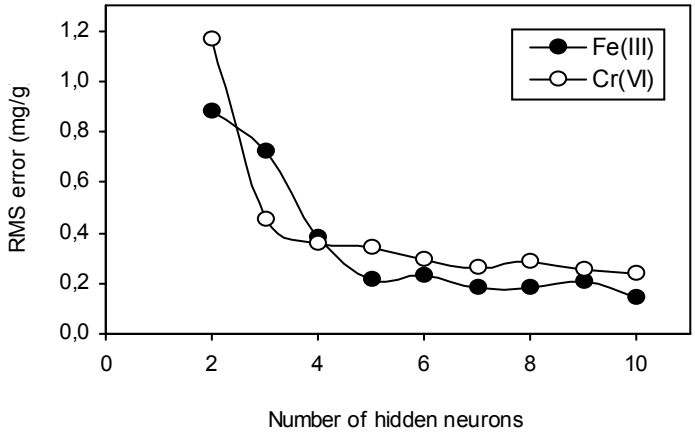


Figure 2. RMS errors in the estimates of neural network trained with the 24 data patterns listed in Table 1 as a function of number of hidden neurons

The experimental $q_{e,Fe}$ and $q_{e,Cr}$ values versus the corresponding network outputs calculated by using the 2-5-2 network model are shown in Figure 3. A line of unit slope, i.e., the line of perfect fit with points corresponding to zero RMS error is also shown in Figure 3. These plots therefore visualize the performance of the model in an obvious way. The results in Figure 3 demonstrate that the neural network model provided a very accurate description of the experimental data, indicating that the model was successful in capturing the nonlinear relationships between the C_e and q_e values of Fe(III) and Cr(VI). Note that the RMS errors in the model's estimates of $q_{e,Fe}$ and $q_{e,Cr}$ are very small, as shown in Table 2, and thus an excellent approximation has been obtained.

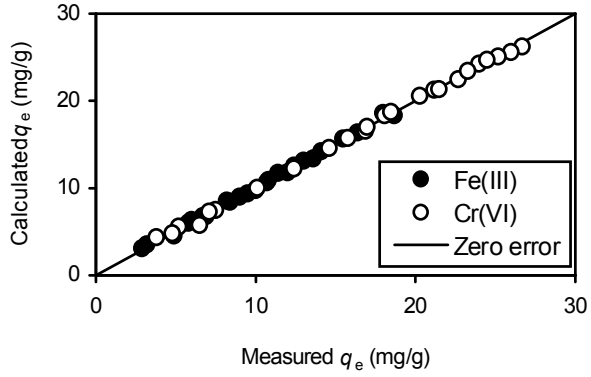


Figure 3. Neural network-calculated Fe(III) and Cr(VI) uptake values versus actual uptake values

Table 2
Root mean square errors of 2-5-2 neural network and two Langmuir based models in correlating the 24 data patterns listed in Table 1

Model	RMS error (mg/g)	
	Fe(III)	Cr(VI)
Neural network	0.232	0.292
Competitive Langmuir[a]	1.131	1.285
Competitive Langmuir-Freundlich[a]	0.797	1.162

[a]RMS errors were estimated using model-calculated values of $q_{e,Fe}$ and $q_{e,Cr}$ taken from Aksu *et al.* (1997).

The performance of the neural network was compared with the performance of traditional equilibrium isotherm models. Aksu *et al.* (1997) fitted a competitive Langmuir model and a competitive Langmuir-Freundlich model to the data in Table 1 using $C_{e,Fe}$ and $C_{e,Cr}$ as the independent variables (equivalent to our network input variables) and $q_{e,Fe}$ and $q_{e,Cr}$ as the dependent variables (equivalent to our network output variables). Figure 4 displays the experimental $q_{e,Fe}$ and $q_{e,Cr}$ values versus the corresponding model-calculated values taken from Aksu *et al.* (1997). A comparison of Figures 3 and 4 indicates that the performance of the two Langmuir based models for fitting the experimental binary equilibrium data was inferior to the performance of the neural network model. Another look at comparative performance is provided by Table 2 that lists the RMS errors of the neural network and the two Langmuir based models. It can be seen from this comparison that the neural network approach with significantly smaller RMS errors easily outperformed the traditional modeling approach based on Langmuir-type models.

Although neural networks are capable of correlating or fitting multimetal equilibrium data accurately, the important feature of neural networks is their ability to predict unseen data (data not used in training the network). From a practical standpoint, the ability to predict multimetal equilibrium data could significantly reduce the amount of experimentation required. To demonstrate the ability of the 2-5-2 network model to predict unseen data, the 24 data patterns in Table 1 were split into two groups. The first group comprising 16 data patterns was used for training the neural network and is termed the 'training set'. The second group

comprising eight data patterns was used to test the predictive capability of the trained network and is termed the 'test set'.

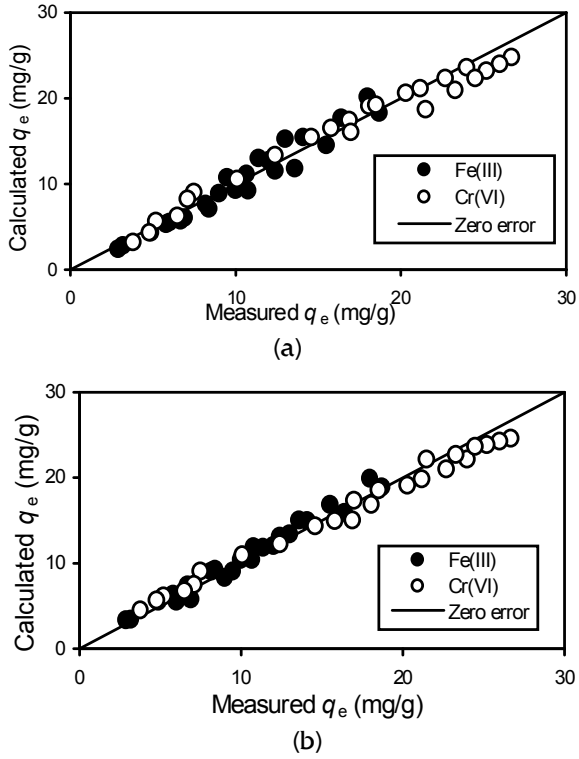


Figure 4. Calculated Fe(III) and Cr(VI) uptake values using Langmuir based models versus actual uptake values. (a) Competitive Langmuir model; (b) Competitive Langmuir-Freundlich model

Figure 5 shows the network-calculated outputs for a training set comprising data patterns *x*.1, *x*.2, *x*.5, and *x*.6 and a test set comprising data patterns *x*.3 and *x*.4 where *x* = 1-4 (see first column of Table 1) plotted against the corresponding experimental data. The solid circles represent the network-*trained* outputs while the open circles denote the network-*predicted* outputs for input variables not belonging to the training set. It is evident that the network model not only fitted the training data very well but also provided predictions of the test data very close to those measured experimentally. There was indeed little difference in the RMS error values for the training and test sets, as shown in part A of Table 3.

Table 3
Root mean square errors of 2-5-2 neural network model in correlating and predicting the 24 data patterns listed in Table 1

	RMS error (mg/g)	
Data set	Fe(III)	Cr(VI)
Part A		
Training set (*x*.1, *x*.2, *x*.5, *x*.6)[a]	0.359	0.403
Test set (*x*.3, *x*.4)[a]	0.481	0.588
Part B		
Training set (*x*.2, *x*.3, *x*.4, *x*.5)[a]	0.311	0.416
Test set (*x*.1, *x*.6)[a]	1.835	2.792

[a] See first column of Table 1; *x* = 1-4.

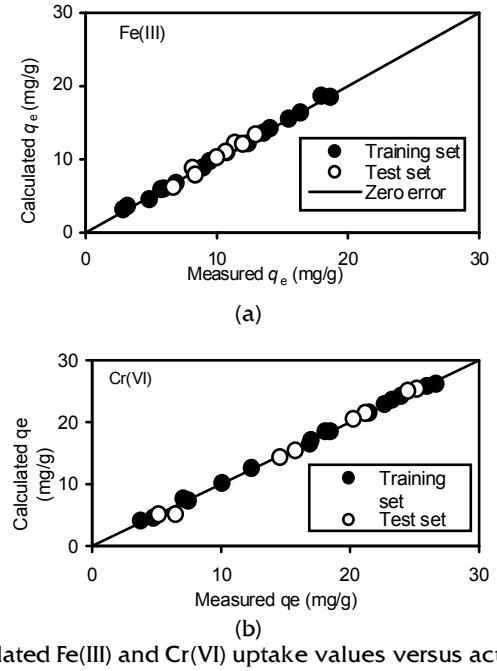
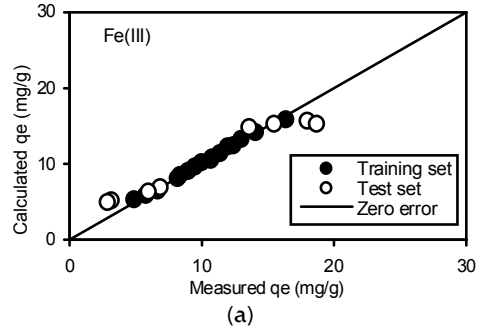


Figure 5. Neural network-calculated Fe(III) and Cr(VI) uptake values versus actual uptake values. The training set contained data patterns *x*.1, *x*.2, *x*.5, and *x*.6 while the test set contained data patterns *x*.3 and *x*.4 where *x* = 1-4 (see first column of Table 1)

The predictive capability of the neural network was further assessed by using different data patterns to train the network. Figure 6 displays the network-calculated $q_{e,Fe}$ and $q_{e,Cr}$ values for a different training set comprising data patterns *x*.2, *x*.3, *x*.4, and *x*.5 and a test set comprising data patterns *x*.1 and *x*.6 where *x* = 1-4 plotted against actual uptake of Fe(III) and Cr(VI). Figure 6 shows wide discrepancies between network-predicted outputs (open circles) and measured data despite the seemingly excellent training (solid circles). The training and test set RMS errors are given in part B of Table 3. As can be seen in Table 3, the network model trained on data patterns *x*.2, *x*.3, *x*.4, and *x*.5 gave substantially worse RMS errors in its predictions of Fe(III) and Cr(VI) uptake values than the network model trained on data patterns *x*.1, *x*.2, *x*.5, and *x*.6. This is largely due to the weakness of neural networks in making extrapolations. Since a neural network is essentially a 'black-box' model, it cannot provide any reasonable extrapolation beyond the range of the training data. An inspection of the data patterns in Table 1 reveals that data patterns *x*.1 and *x*.6 contain the lowest and highest values of the input variable $C_{e,Fe}$. Training a neural network using a data set that excludes the lower and upper bounds of the data will result in a network with a poor predictive capability. The results in Figures 5 and 6 demonstrate that neural networks are better at interpolation than extrapolation. It is therefore critical that training data points be carefully chosen to reflect the range and magnitude of the inputs and outputs in order to develop a neural network with good predictive capability.



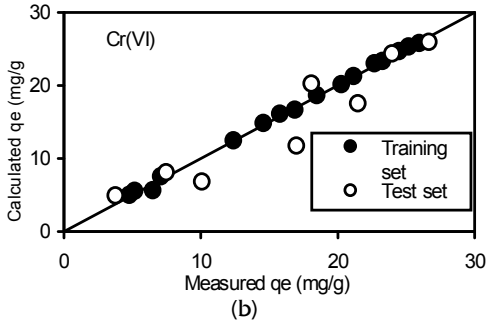


Figure 6. Neural network-calculated Fe(III) and Cr(VI) uptake values versus actual uptake values. The training set contained data patterns *x*.2, *x*.3, *x*.4, and *x*.5 while the test set contained data patterns *x*.1 and *x*.6 where *x* = 1-4 (see first column of Table 1)

A suitably trained neural network with good predictive capability for interpolation can be viewed as a useful tool for saving experimental time and effort. For example, the 2-5-2 network model trained with a subset of available data to produce the excellent predictions in Figure 5 can be used to simulate complete equilibrium isotherms for the binary metal system at different experimental conditions, as shown in Figure 7. It is difficult to produce the isotherms shown in Figure 7 experimentally because the conventional batch techniques employed in multimetal biosorption studies do not generate complete isotherms but rather a collection of paired data points with each point lying on a different isotherm. The extent of competitive biosorption in the Fe(III)/Cr(VI) binary system can be easily deduced from Figure 7 which is not immediately apparent from an examination of the experimental data listed in Table 1.

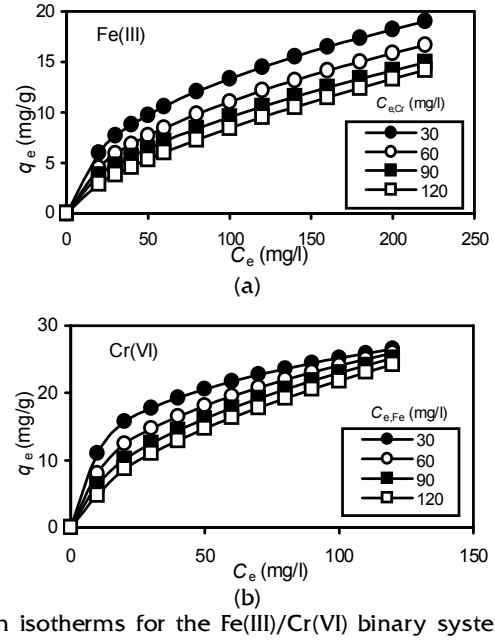


Figure 7. Simulated equilibrium isotherms for the Fe(III)/Cr(VI) binary system obtained from 2-5-2 neural network. (a) The effect of Cr(VI) on the equilibrium isotherms of Fe(III); (b) The effect of Fe(III) on the equilibrium isotherms of Cr(VI)

4. CONCLUSIONS

This study has demonstrated the feasibility of using a neural network to capture the nonlinear and interacting relationship between the liquid phase and biosorbent phase

equilibrium concentrations of Fe(III) and Cr(VI) in a binary biosorption system. The trained network model substantially outperformed two Langmuir based isotherm models in correlating the binary uptake data. Moreover, the neural network trained with a limited number of data points was capable of predicting fresh data that was not used to train the network. However, it should be noted that the selection of training data is critical to successful prediction using neural networks. The neural network approach outlined here for a binary metal system can readily be extended to systems containing more than two metals if sufficient experimental data is available.

REFERENCES

Aksu, Z., Acikel, U., Kutsal, T., Application of multicomponent adsorption isotherms to simultaneous biosorption of iron(III) and chromium(VI) on *C. vulgaris*. J. Chem. Technol. Biotechnol., 1997, **70**, 368-378.

Baughman, D.R., Liu, Y.A., Neural networks in bioprocessing and chemical engineering. Academic Press, 1995, San Diego.

Bulsari, A.B., Neural networks for chemical engineers. Elsevier, 1995, Amsterdam.

Carksy, M., Do, D.D., Neural network modeling of adsorption of binary vapour mixtures. Adsorption, 1999, **5**, 183-192.

Chang, J.S., Chen, C.C., Quantitative analysis and equilibrium models of selective adsorption in multimetal systems using a bacterial biosorbent. Sep. Sci. Technol., 1998, **33**, 611-632.

Chong, K.H., Volesky, B., Description of two-metal biosorption equilibria by Langmuir-type models. Biotechnol. Bioeng., 1995, **47**, 451-460.

Hornik, K., Approximation capabilities of multilayer feedforward networks. Neural Networks, 1991, 4, 251-257.

Hornik, K., Stinchcombe, M., White, H., Multilayer feedforward networks are universal approximators. Neural Networks, 1989, **2**, 359-366.

Kaewsarn, P., Yu, Q., Binary adsorption of copper(II) and cadmium(II) from aqueous solutions by biomass of marine alga *Durvillaea Potatorum*. Sep. Sci. Technol., 1999, **34**, 1595-1605.

Klimmer, S., Stan, H.J., Wilke, A., Bunke, G., Buchholz, R., Comparative analysis of the biosorption of cadmium, lead, nickel, and zinc by algae. Environ. Sci. Technol., 2001, **35**, 4283-4288.

Lee, M.G., Lim, J.H., Kam, S.-K., Biosorption characteristics in the mixed heavy metal solution by biosorbents of marine brown algae. Korean J. Chem. Eng., 2002, **19**, 277-284.

Puranik, P.R., Paknikar, K.M., Influence of co-cations on biosorption of lead and zinc – a comparative evaluation in binary and multimetal systems. Bioresource Technol., 1999, **70**, 269-276.

Sag, Y., Kaya, A., Kutsal, T., Lead, copper and zinc biosorption from bicomponent systems modeled by empirical Freundlich isotherm. Appl. Microbiol. Biotechnol., 2000, **53**, 338-341.

Sanchez, A., Ballester, A., Blazquez, M.L., Gonzalez, F., Munoz, J., Hammaini, A., Biosorption of copper and zinc by *Cymodocea nodosa*. FEMS Microbiol. Rev., 1999, **23**, 527-536.

Sato, A., Sha, Z., Palosaari, S., Neural networks for chemical engineering unit operations. Chem. Eng. Technol., 1999, 22, 732-739.

Schiewer, S., Volesky, B., Modeling multi-metal ion exchange in biosorption. Environ. Sci. Technol., 1996, **30**, 2921-2927.

Syu, M.-J., Tsai, G.-J., Tsao, G.T., Artificial neural network modeling of adsorptive separation. Adv. Biochem. Eng./Biotechnol., 1993, **49**, 97-122.

Yang, M., Hubble, J., Lockett, A.D., Rathbone, R.R., A neural network model for prediction of binary adsorption using single solute and limited binary solute adsorption data. Sep. Sci. Technol., 1996, **31**, 1259-1265.