Available online at www.sciencedirect.com
SCIENCE DIRECT
Dyes and Pigments 73 (2007) 230–238
DYES and PIGMENTS
www.elsevier.com/locate/dyepig

Solubility prediction of 21 azo dyes in supercritical carbon dioxide using wavelet neural network

R. Tabaraki, T. Khayamian*, A.A. Ensafi

Department of Chemistry, College of Chemistry, Isfahan University of Technology, Isfahan 84154, Iran

Received 11 July 2005; received in revised form 18 October 2005; accepted 14 December 2005
Available online 20 February 2006

Abstract

The solubility of 21 azo dyes in supercritical carbon dioxide was related to the six descriptors over a wide range of pressures (100–355 bar) and temperatures (308–413 K). The wavelet neural network (WNN) model was constructed with six descriptors as an input layer, eight neurons as a hidden layer and a neuron as an output layer. The descriptors consisted of temperature, pressure, LUMO energy, polarizability, volume of the molecule and number of unsaturated bonds and they were selected based on stepwise feature selection from different descriptors using multiple linear regression (MLR) method. The WNN architecture and its parameters were optimized simultaneously. The data were randomly divided into the training, prediction and validation sets. The RMSE and mean absolute errors in WNN model were 0.220 and 0.158 for prediction set and 0.156 and 0.114 for validation set. In addition, the prediction ability of the model was also evaluated for five azo dyes, the molecules and data of which were not in any previous data sets.

The performance of the WNN model was also compared with artificial neural network (ANN) and MLR models.

Keywords: Azo dyes; SC-CO_2; Solubility; WNN

1. Introduction

Supercritical fluid dyeing (SFD) is an alternative dyeing process, which is able to replace the conventional wet process. In this process, water, surfactants, dispersing agents and drying process are eliminated. Therefore, this method will not deliver a lot of wastewater to the environment. Moreover, a lot of energy (roughly 50%) can be saved [1–5]. However, in order to apply this technique, the knowledge of the solubility of disperse dyes in supercritical carbon dioxide (SC-CO_2) is required. The solubility measurements of dyes in SC-CO_2 have been conducted by many researchers [6–24]. These experimental data are commonly to correlate with theoretical or semi-empirical models [25–34]. In addition to those hard modeling methods, some software methods such as artificial neural network (ANN) [35] and wavelet neural network (WNN) [36] have also been used for the prediction of solubility in supercritical conditions.

The purpose of this work is (1) prediction of solubility of various azo dyes in supercritical carbon dioxide using WNN, (2) simultaneous optimization of the WNN architecture and its parameters, and (3) evaluation of the performance of the model using two data sets; a validation set, in which data were not used in construction of the model and a data set consisted of five azo dyes, the molecules and data of which were new for the model. This model can be used to predict the solubility of newly synthesized azo dyes in SC-CO_2 as a primary estimation.

2. Theory

2.1. Wavelet

Wavelet is a type of transformation that retains both time and frequency information of the signal [37]. In chemical

* Corresponding author. Tel.: +98 311 391 2351; fax: +98 311 391 2350.
E-mail address: taghi@cc.iut.ac.ir (T. Khayamian).

0143-7208/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dyepig.2005.12.003

studies, the time domain can be replaced by other domains such as wavelength. In Fourier transform, only the sine and cosine functions can be chosen as the basis functions. However, wavelet transformation (WT) has versatile basis functions to be selected based on the type of the signal analyzed. In WT, all basis function $\psi_{a,b}(x)$ can be derived from a mother wavelet $\psi(x)$ through the following dilation and translation processes:

$$\psi_{a,b}(x) = a^{-1/2}\psi\left(\frac{x-b}{a}\right) \quad a,b\in R \text{ and } a>0 \tag{1}$$

where the parameters of translation are $b\in R$ and of dilation are $a\in R$ and $a>0$ (R denotes real number). The mother wavelet, $\psi(x)$, is a single fixed function such as Morlet function from which, all basis functions are generated.

The continuous wavelet transformation of a signal function such as $f(x)$ is given by

$$W_f(a,b) = \int_0^{+\infty} \psi^* a,b(x) f(x)\mathrm{d}x \tag{2}$$

where the superscript * represents the complex conjugate. From Eq. (2) $W_f(a,b)$ can be computed through the convolution product.

The applications of wavelet in chemistry have been reviewed in several papers [38] and books [39].

2.2. Wavelet neural networks

Wavelet neural network (WNN) is a novel approach towards the learning function. Wavelet networks, which combine the wavelet theory and feed-forward neural networks, utilize wavelets as the basis function to construct a network. Wavelet function is a local function and influences the networks' output only in some local ranges. The wavelet neural network shows surprising effectiveness in solving the conventional problems of poor convergence or even divergence encountered in other kinds of neural networks [40]. Guo et al. have used WNN for prediction of driving forces of α-cyclodextrin complexation with benzene derivatives [41] and the inclusion of β-cyclodextrin with benzene derivatives [42].

The topological structure of the WNN employed in this study is shown in Fig. 1. The WNN consists of three layers: input layer, hidden layer and output layer. The calculation steps of WNN have been given in Ref. [40]. In brief, the connections between input–hidden units and hidden–output units are called weights u_{ti} and w_t, respectively. A Morlet mother function is used as node activation function for the hidden layer. The dilation and translation parameters, a_t and b_t, of the Morlet function for each node in the hidden layer are different and they need to be optimized. In the WNN, the gradient descend algorithm is employed and the error is minimized by adjusting u_{ti}, w_t, a_t, and b_t parameters [43]. In the WNN, the following steps are carried out:

(1) Initializing the dilation parameter a_t, translation parameter b_t and node connection weights u_{ti}, w_t to some random values. All those random values are limited in the interval (0, 1).

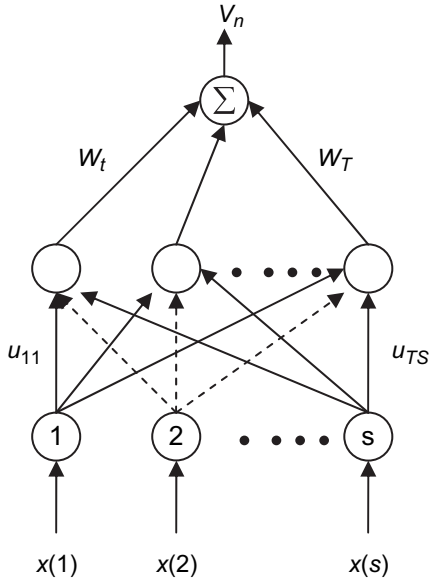

Fig. 1. The WNN topology structure.

(2) Inputting data $X_n(i)$ and corresponding output values v_n^T, where the superscript T represents the target output state.

(3) Propagating the initial signal forward through the network using

$$v_n = \sum_{t=1}^{T} w_t h\left(\frac{\sum_{i=1}^{s} u_{ti}x_n(i) - b_t}{a_t}\right) \tag{3}$$

where h is taken as a Morlet wavelet

$$h(t) = \cos(1.75t)\ \exp\left(-\frac{t^2}{2}\right) \tag{4}$$

(4) Calculation of the WNN parameters:

$$\Delta w_t^{\text{new}} = -\eta\frac{\partial E}{\partial w_t^{\text{old}}} + \alpha\Delta w_t^{\text{old}} \tag{5}$$

$$\Delta u_{ti}^{\text{new}} = -\eta\frac{\partial E}{\partial u_{ti}^{\text{old}}} + \alpha\Delta u_{ti}^{\text{old}} \tag{6}$$

$$\Delta a_t^{\text{new}} = -\eta\frac{\partial E}{\partial a_t^{\text{old}}} + \alpha\Delta a_t^{\text{old}} \tag{7}$$

$$\Delta b_t^{\text{new}} = -\eta\frac{\partial E}{\partial b_t^{\text{old}}} + \alpha\Delta b_t^{\text{old}} \tag{8}$$

The error function E is taken as

$$E = \frac{1}{2}\sum_{n=1}^{N}\left(v_n^T - v_n\right)^2 \tag{9}$$

where v_n^T and v_n are the experimental and calculated values, respectively. N stands for the data number of training set, and η