

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.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Azo dyes; SC-CO₂; 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-CO₂) is required. The solubility measurements of dyes in SC-CO₂ 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-CO₂ 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).

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:

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

where the parameters of translation are $b \in \mathbb{R}$ and of dilation are $a \in \mathbb{R}$ and $a > 0$ (\mathbb{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

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

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

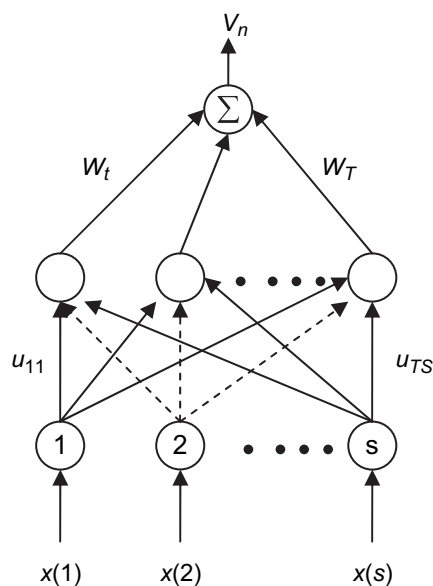


Fig. 1. The WNN topology structure.

values. All those random values are limited in the interval (0, 1).

- (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

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

where h is taken as a Morlet wavelet

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

- (4) Calculation of the WNN parameters:

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

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

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

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

The error function E is taken as

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

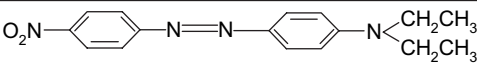
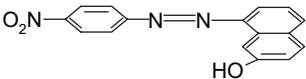
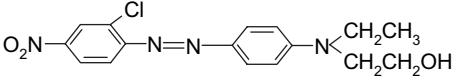
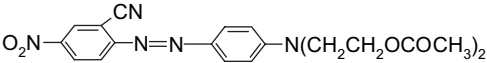
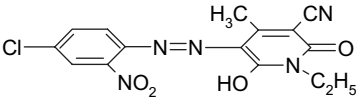
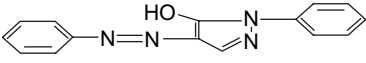
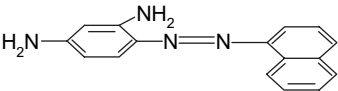
where v_n^T and v_n are the experimental and calculated values, respectively. N stands for the data number of training set, and η

Table 1
Molecular structure of azo dyes

	Dyes	Molecular structure	References
1	Yellow 7		[7–8]
2	Red 1		[7]
3	Orange 3		[10]
4	Blue 79		[11]
5	APAN		[6]
6	Red 153		[12]
7	Yellow 119		[12]
8	Orange 30		[13]
9	Yellow 108		[13]
10	Brown 22		[13]
11	Red 30		[13]
12	Red 167		[13]
13	Violet 91		[13]
14	D1		[15]

(continued on next page)

Table 1(continued)

	Dyes	Molecular structure	References
15	D2		[15]
16	D3		[15]
17	Red 13		[22]
18	Red 82		[23]
19	Modified Yellow 119		[23]
20	Yellow 16		[24]
21	Solvent Brown 1		[30]

and α being the learning rate and the momentum term, respectively.

(5) The WNN parameters were changed until the network output satisfies the error criteria.

3. Data and methodology

3.1. Data set

The structures of the 21 azo dyes with the references to their experimental solubility values are given in Table 1.

3.2. Descriptor selection and calculations

HyperChem (version 6) software was used to calculate the quantum chemical and geometrical descriptors. Before calculation of the descriptors, optimization of the molecular structures was carried out by semi-empirical AM1 method using the Fletcher–Reeves algorithm until the root mean square gradient of 0.01 was obtained. The calculated descriptors were electronic energy, nuclear energy, total dipole, highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), surface area, volume, polarizability, etc. In addition some topological descriptors such as number of double bonds, number of heteroatoms and number of carbon atoms were also calculated. The selection of relevant descriptors, which relate solubility to the molecular structure, is an important step in the construction of

a predictive model. The most feasible descriptors were selected stepwisely by MLR method. In this method, the selected descriptors are those, which construct a model with minimum error. The selected descriptors with their mean

Table 2

The calculated values of different descriptors, studied in this work

Dye	LUMO (eV)	Volume (Å ³)	Polarizability (Å ³)	Number of unsaturated bonds
1 Yellow 7	−1.005	978.79	36.35	11
2 Red 1	−1.359	936.24	33.32	8
3 Orange 3	−1.375	717.51	25.35	8
4 Blue 79	−1.980	1541.36	55.35	12
5 APAN	−0.941	795.27	31.41	9
6 Red 153	−1.496	1076.77	41.55	9
7 Yellow 119	−1.638	885.58	31.78	9
8 Orange 30	−1.519	1186.71	42.79	10
9 Yellow 108	−1.607	1027.09	36.91	10
10 Brown 22	−1.466	1022.23	37.82	8
11 Red 30	−1.485	998.15	35.89	8
12 Red 167	−1.648	1360.42	48.51	11
13 Violet 91	−1.682	1369.16	48.42	12
14 D1	−1.428	812.13	29.02	8
15 D2	−1.415	903.65	32.69	8
16 D3	−1.742	820.74	31.90	10
17 Red 82	−1.714	1238.38	43.32	11
18 Red 13	−1.558	962.80	35.25	8
19 Yellow 16	−0.654	848.89	31.45	9
20 Modified Yellow 119	−1.434	939.74	33.71	9
21 Solvent Brown 1	−0.560	809.87	32.13	9

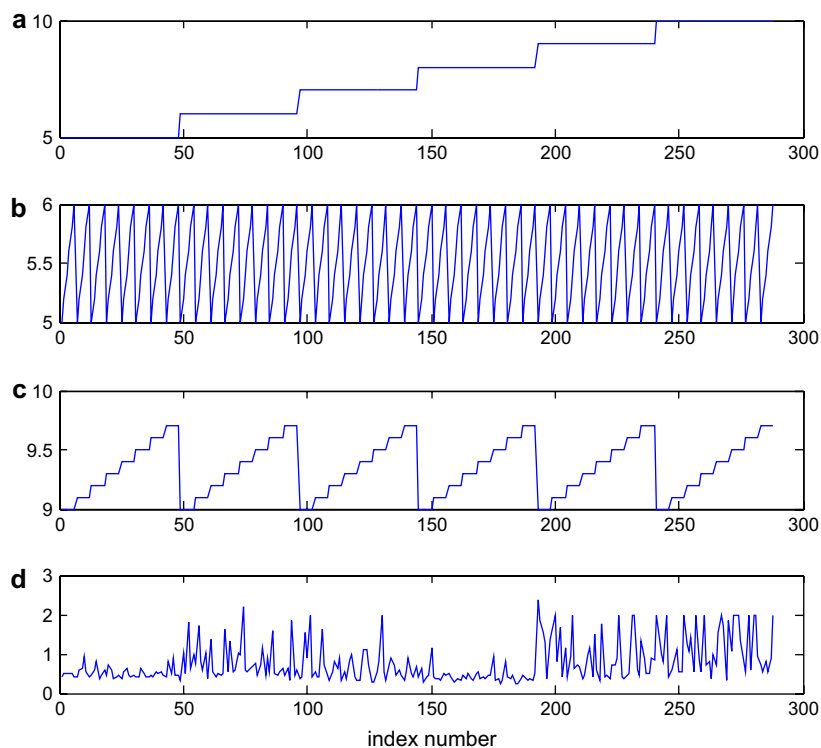


Fig. 2. Plots of (a) the number of neurons in the hidden layer against index number; (b) the momentum values against the index number; (c) the learning rate values against the index number; (d) the RMSE for the prediction set against the index number.

Table 3
Experimental and calculated values of $\log(S)$ using MLR, ANN and WNN models for the prediction set

Dye	<i>P</i> (bar)	<i>T</i> (K)	Exp. $\log(S)$ (mol/mol)	Calc. (MLR) $\log(S)$ (mol/mol)	Absolute error (MLR)	Calc. (ANN) $\log(S)$ (mol/mol)	Absolute error (ANN)	Calc. (WNN) $\log(S)$ (mol/mol)	Absolute error (WNN)
Yellow 7	150	353.15	−6.478	−5.668	0.810	−5.714	0.764	−5.896	0.582
	100	393.15	−6.320	−5.925	0.395	−6.555	−0.235	−6.577	−0.257
	152	333.15	−5.684	−5.689	−0.005	−5.648	0.036	−5.707	−0.023
Red 1	200	313.15	−6.191	−5.669	0.522	−6.226	−0.035	−6.034	0.157
	100	393.15	−6.627	−6.184	0.443	−7.190	−0.563	−7.133	−0.506
Orange 3	149.8	323.7	−5.156	−5.925	−0.769	−5.268	−0.112	−5.513	−0.357
	203.9	353.7	−5.004	−5.519	−0.515	−5.210	−0.206	−5.143	−0.139
	250.6	383.7	−4.719	−5.161	−0.442	−4.652	0.067	−4.629	0.090
Blue 79	148.1	323.7	−5.962	−6.612	−0.650	−6.261	−0.299	−5.905	0.057
	196.7	353.7	−5.852	−6.241	−0.389	−5.930	−0.078	−5.905	−0.053
APAN	250	353.15	−4.489	−5.115	−0.626	−4.307	0.182	−4.432	0.057
	300	373.15	−4.153	−4.753	−0.600	−4.151	0.002	−4.174	−0.021
Red 153	150	353.2	−7.523	−6.342	1.181	−6.979	0.544	−7.043	0.481
Yellow 119	200	353.2	−6.349	−5.830	0.519	−6.188	0.161	−6.420	−0.071
	250	393.2	−5.589	−5.433	0.156	−5.612	−0.023	−5.630	−0.041
Orange 30	350	323.15	−4.140	−4.8624	−0.722	−4.380	−0.240	−4.353	−0.213
Yellow 108	200	323.15	−5.077	−5.867	−0.790	−5.334	−0.257	−5.329	−0.252
Brown 22	350	323.15	−4.812	−4.915	−0.103	−4.748	0.064	−5.186	−0.374
Red 30	350	323.15	−5.372	−4.830	0.542	−4.873	0.499	−5.185	0.187
Red 167	200	373.15	−6.013	−5.840	0.173	−6.018	−0.005	−6.096	−0.083
D1	152	308	−5.562	−6.034	−0.472	−5.623	−0.061	−5.540	0.022
	334	308	−5.335	−4.844	0.491	−5.392	−0.057	−5.318	0.017
	213	328	−5.374	−5.600	−0.226	−5.289	0.085	−5.302	0.072
	355	328	−5.007	−4.672	0.335	−5.001	0.006	−4.980	0.027
D2	304	308	−5.585	−5.087	0.498	−5.533	0.052	−5.718	−0.133
	122	328	−6.292	−6.243	0.049	−6.212	0.080	−6.276	0.016
	213	328	−5.654	−5.648	0.006	−5.598	0.056	−5.770	−0.116
	355	348	−4.955	−4.684	0.271	−4.926	0.029	−4.866	0.089
D3	243	328	−5.839	−5.859	−0.020	−5.771	0.068	−5.682	0.157
	213	348	−5.903	−6.020	−0.117	−5.852	0.051	−5.821	0.082

effects in parentheses were pressure (1.5174), temperature (0.6020), volume of the molecule (2.7756), LUMO energy (−1.3148), polarizability (−3.0756) and number of unsaturated bonds in the molecule (0.2683). The mean effect describes the influence of each descriptor on the MLR model. Selection of pressure and temperature as the experimental descriptors is important because the density of the supercritical fluid, which is the key parameter to the solubility of different compounds, is related to both temperature and pressure of the supercritical gas. Polarizability and LUMO energy can be explained as follows: the polarization of a molecule by an external electric field is given in terms of the n th order susceptibility tensors of the molecular bulk. The first order term is referred to the polarizability; the second order term is called the first hyperpolarizability, etc. Thus, the most significant property of the molecular polarizability is its relation to the molecular bulk or molar volume [44]. The LUMO energy is related to the second electronic affinity of the molecule. Presumably, this descriptor is involved in the formation of charge–transfer complexes between the solute and the solvent. The occupation of a higher energy orbital might lower the electron–electron repulsion and results in a lower overall molecular energy. The picture of this interaction, in terms of molecular orbital (MO), is nearly delocalization of the lone pair of electrons lying on a nonbonding MO of the donor atom, in this case the oxygen in carbon dioxide, into a vacant antibonding MO of an acceptor molecule. This electronic delocalization creates a bond polarity and an electrostatic attraction which favors the solubility; therefore, the larger the extent of this interaction the higher the solubility [45]. The volume of the molecule is another selected descriptor and its effect on solubility has been reported previously [32,33,35]. The number of unsaturated bonds in the molecules is the other relevant descriptor, which enhances the correlation coefficient between the observed and the predicted solubilities of the model. The numerical values of these three descriptors are given in Table 2. All computations were carried out on a Pentium 4, 1.5 GHz PC computer. Multiple linear regression analysis of molecular descriptors and logarithm of the solubility were carried out using the stepwise strategy in the SPSS (9.0, for Windows) software. The ANN and WNN algorithms were written in MATLAB 6.1 (Math Works) by the authors.

4. Results and discussion

The data set consisted of experimental values of the solubility of the 21 azo dyes at different temperatures and pressures. The data set of 16 azo dyes was randomly divided into three data sets; training, prediction and validation sets. These sets consist of 139, 30 and 82 data, respectively. It is necessary to emphasize that the data of the validation set have not contributed to the optimization of the model and is only used for evaluation of the performance of the model.

4.1. Wavelet neural network model

The network architecture consisted of six neurons in the input layer corresponding to the six mentioned descriptors. The output layer had one neuron that predicts the solubility. The number of neurons in the hidden layer is unknown and needs to be optimized. In addition to the number of neurons in the hidden layer, the learning rate, the momentum and the number of iterations should also be optimized. In this work, the number of neurons in the hidden layer and other parameters except the number of iterations were simultaneously optimized. The procedure was performed by the following way. A MATLAB program was written to change the number of neurons in the hidden layer from 2 to 11, the learning rate from 0.001 to 0.1 with a step of 0.001 and the momentum

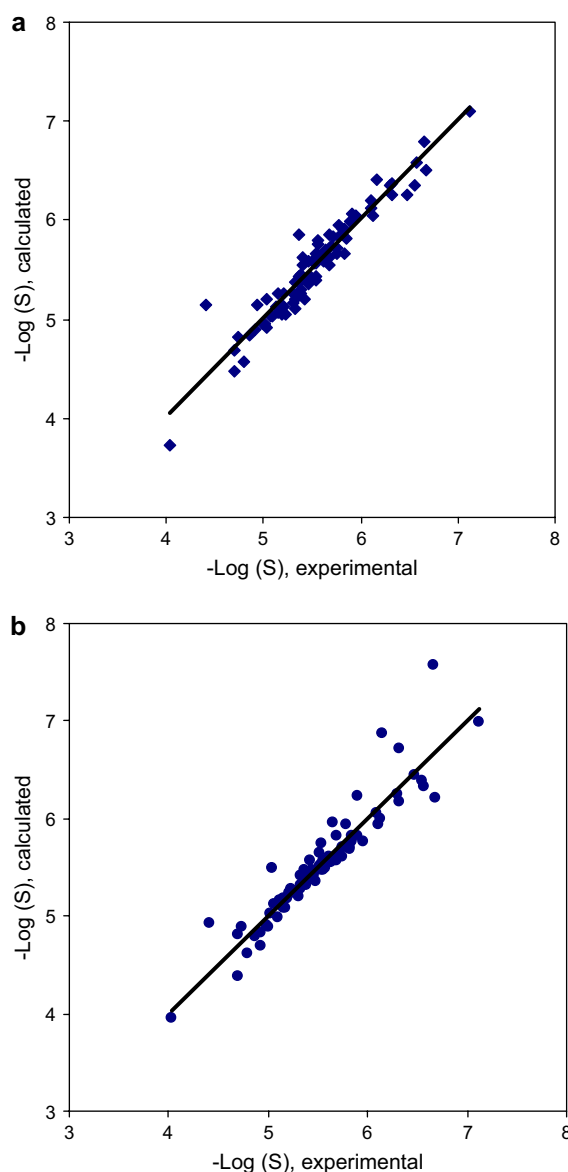


Fig. 3. Scatter plot of the calculated $-\log(S)$ against the experimental $-\log(S)$ for the validation set; (a) WNN model ($R^2 = 0.9189$); (b) ANN model ($R^2 = 0.8774$).

Table 4
Experimental and calculated values of $\log(S)$ using MLR, ANN and WNN models for the validation set

Dye	<i>P</i> (bar)	<i>T</i> (K)	Exp. $\log(S)$ (mol/mol)	Calc. (MLR) $\log(S)$ (mol/mol)	Absolute error (MLR)	Calc. (ANN) $\log(S)$ (mol/mol)	Absolute error (ANN)	Calc. (WNN) $\log(S)$ (mol/mol)	Absolute error (WNN)
Yellow 7	200	313.15	−5.527	−5.410	0.117	−5.638	−0.111	−5.605	−0.078
	250	393.15	−4.695	−4.944	−0.249	−4.806	−0.111	−4.481	0.214
	172	353.15	−5.582	−5.524	0.058	−5.482	0.100	−5.643	−0.061
Red 1	150	353.15	−6.674	−5.926	0.748	−6.218	0.456	−6.499	0.175
	250	393.15	−4.791	−5.202	−0.411	−4.606	0.185	−4.573	0.218
Orange 3	247.8	353.7	−4.739	−5.232	−0.493	−4.891	−0.152	−4.819	−0.080
	148.1	413.7	−5.357	−5.779	−0.422	−5.473	−0.116	−5.863	−0.506
	310.6	413.7	−4.034	−4.716	−0.682	−3.949	0.085	−3.723	0.311
Blue 79	202	383.7	−6.156	−6.154	0.002	−6.878	−0.722	−6.415	−0.259
	305.1	383.7	−5.036	−5.480	−0.444	−5.485	−0.449	−5.201	−0.165
APAN	200	373.15	−5.022	−5.407	−0.385	−5.018	0.004	−4.913	0.109
Red 153	200	393.2	−6.575	−5.945	0.630	−6.323	0.252	−6.576	−0.001
Yellow 119	150	373.2	−7.127	−6.122	1.005	−6.992	0.135	−7.094	0.033
Orange 30	200	373.2	−5.757	−5.756	0.001	−5.615	0.142	−5.724	0.033
Yellow 108	350	373.15	−4.416	−4.798	−0.382	−4.926	−0.510	−5.148	−0.732
Brown 22	300	323.15	−4.932	−5.242	−0.310	−4.690	0.242	−5.151	−0.219
Red 30	300	373.15	−4.690	−5.070	−0.380	−4.387	0.303	−4.684	0.006
Violet 91	200	373.15	−6.481	−5.808	0.673	−6.440	0.041	−6.266	0.215
D1	243	308	−5.442	−5.439	0.003	−5.470	−0.028	−5.412	0.030
	122	328	−5.699	−6.196	−0.497	−5.819	−0.120	−5.743	−0.044
	274	348	−5.010	−5.166	−0.156	−4.884	0.126	−4.961	0.049
	304	328	−5.151	−5.005	0.146	−5.078	0.073	−5.075	0.076
	122	318	−5.836	−6.213	−0.377	−5.770	0.066	−5.671	0.165
	152	318	−5.674	−6.017	−0.343	−5.576	0.098	−5.543	0.131
	182	318	−5.548	−5.821	−0.273	−5.479	0.069	−5.441	0.107
	213	318	−5.456	−5.618	−0.162	−5.407	0.049	−5.359	0.097
	243	318	−5.376	−5.422	−0.046	−5.352	0.024	−5.300	0.076
	274	318	−5.377	−5.219	0.158	−5.306	0.071	−5.252	0.125
	304	318	−5.324	−5.023	0.301	−5.269	0.055	−5.209	0.115
	334	318	−5.295	−4.826	0.469	−5.238	0.057	−5.163	0.132
	355	318	−5.206	−4.689	0.517	−5.220	−0.014	−5.124	0.082
	122	338	−5.782	−6.178	−0.396	−5.940	−0.158	−5.836	−0.054
	152	338	−5.682	−5.982	−0.300	−5.561	0.121	−5.623	0.059
	182	338	−5.472	−5.786	−0.314	−5.357	0.115	−5.434	0.038
	213	338	−5.306	−5.583	−0.277	−5.205	0.101	−5.270	0.036
	243	338	−5.182	−5.387	−0.205	−5.086	0.096	−5.142	0.040
	274	338	−5.094	−5.184	−0.090	−4.983	0.111	−5.038	0.056
	304	338	−4.987	−4.988	−0.001	−4.900	0.087	−4.958	0.029
	334	338	−4.921	−4.791	0.130	−4.829	0.092	−4.891	0.030
	355	338	−4.861	−4.654	0.207	−4.785	0.076	−4.846	0.015
D2	152	308	−5.951	−6.082	−0.131	−5.767	0.184	−6.046	−0.095
	243	308	−5.676	−5.486	0.190	−5.597	0.079	−5.861	−0.185
	182	348	−5.818	−5.816	0.002	−5.690	0.128	−5.903	−0.085
	274	348	−5.151	−5.214	−0.063	−5.187	−0.036	−5.264	−0.113
	122	318	−6.097	−6.260	−0.163	−6.064	0.033	−6.199	−0.102
	152	318	−5.903	−6.064	−0.161	−5.822	0.081	−6.065	−0.162
	182	318	−5.780	−5.868	−0.088	−5.711	0.069	−5.949	−0.169
	213	318	−5.706	−5.665	0.041	−5.633	0.073	−5.843	−0.137
	243	318	−5.564	−5.469	0.095	−5.573	−0.009	−5.750	−0.186
	274	318	−5.541	−5.266	0.275	−5.521	0.02	−5.657	−0.116
	304	318	−5.533	−5.070	0.463	−5.480	0.053	−5.563	−0.030
	334	318	−5.388	−4.874	0.514	−5.445	−0.057	−5.459	−0.071
	355	318	−5.360	−4.736	0.624	−5.424	−0.064	−5.377	−0.017
	122	338	−6.553	−6.226	0.327	−6.393	0.16	−6.361	0.192
	152	338	−6.102	−6.029	0.073	−5.938	0.164	−6.122	−0.020
	182	338	−5.830	−5.833	−0.003	−5.711	0.119	−5.905	−0.075
	213	338	−5.627	−5.630	−0.003	−5.548	0.079	−5.707	−0.080
	243	338	−5.409	−5.434	−0.025	−5.423	−0.014	−5.539	−0.130
	274	338	−5.322	−5.231	0.091	−5.316	0.006	−5.385	−0.063
	304	338	−5.206	−5.035	0.171	−5.230	−0.024	−5.251	−0.045
	334	338	−5.127	−4.839	0.288	−5.156	−0.029	−5.122	0.005
	355	338	−5.068	−4.701	0.367	−5.112	−0.044	−5.033	0.035

(continued on next page)

Tabel 4 (continued)

Dye	P (bar)	T (K)	Exp. log(S) (mol/mol)	Calc. (MLR) log(S) (mol/mol)	Absolute error (MLR)	Calc. (ANN) log(S) (mol/mol)	Absolute error (ANN)	Calc. (WNN) log(S) (mol/mol)	Absolute error (WNN)
D3	152	328	−6.319	−6.454	−0.135	−6.164	0.155	−6.248	0.071
	355	328	−5.561	−5.126	0.435	−5.538	0.023	−5.788	−0.227
	334	348	−5.362	−5.229	0.133	−5.319	0.043	−5.427	−0.065
	152	338	−6.301	−6.437	−0.136	−6.254	0.047	−6.361	−0.060
	182	338	−6.125	−6.240	−0.115	−5.987	0.138	−6.054	0.071
	213	338	−5.845	−6.038	−0.193	−5.829	0.016	−5.815	0.030
	243	338	−5.752	−5.841	−0.089	−5.710	0.042	−5.661	0.091
	274	338	−5.614	−5.638	−0.024	−5.605	0.009	−5.577	0.037
	304	338	−5.511	−5.442	0.069	−5.517	−0.006	−5.560	−0.049
	334	338	−5.455	−5.246	0.209	−5.440	0.015	−5.591	−0.136
	355	338	−5.403	−5.109	0.294	−5.392	0.011	−5.633	−0.230
	122	358	−6.658	−6.598	0.060	−7.569	−0.911	−6.798	−0.140
	152	358	−6.328	−6.402	−0.074	−6.711	−0.383	−6.368	−0.040
	182	358	−5.896	−6.206	−0.310	−6.235	−0.339	−5.983	−0.087
	213	358	−5.654	−6.003	−0.349	−5.951	−0.297	−5.649	0.005
	243	358	−5.545	−5.806	−0.261	−5.742	−0.197	−5.397	0.148
	274	358	−5.420	−5.604	−0.184	−5.559	−0.139	−5.212	0.208
	304	358	−5.322	−5.407	−0.085	−5.406	−0.084	−5.103	0.219
	334	358	−5.228	−5.211	0.017	−5.272	−0.044	−5.053	0.175
	355	358	−5.190	−5.074	0.116	−5.186	0.004	−5.047	0.143

from 0.1 to 0.99 with a step of 0.01. The root mean square errors (RMSE) for training and prediction sets were calculated for all of the possible combination of values for the mentioned variables. Fig. 2 shows changes in the variables against the index number. Each index number corresponds to a combination of variable values and total number of index number is equal to the possible combinations for variable values. In Fig. 2 a subset of index numbers are shown. The RMSE for prediction set for each index number are calculated and the results for a subset of index numbers are shown in Fig. 2d. It was realized that the RMSE for the training and prediction sets are minimum when eight neurons were selected in the hidden layer and the learning rate and the momentum values were 0.056 and 0.95, respectively. Finally, the number of iterations was optimized with the optimum values for the variables. It was realized that after 5800 iterations, the RMSE for prediction set were minimum. The calculated absolute errors for the prediction set are given in Table 3. The performance of the model was evaluated by plotting the estimated vs. experimental solubility values for the validation set (Fig. 3). The absolute errors for the validation set are given in Table 4. The RMSE and mean absolute errors were 0.220 and 0.158 for prediction set and 0.156 and 0.114 for validation set, respectively.

4.2. Artificial neural network method

In order to compare the predictive ability of WNN with another nonlinear model, a similar study was carried out using ANN model. A neural network was constructed with a sigmoid function as the hidden transfer function and a linear function as the output transfer function. A back propagation learning algorithm was employed to adjust the weights. The ANN architecture and its parameters were optimized simultaneously as

the same as WNN. It was realized that the RMSE for the training and prediction sets are minimum when eight neurons were selected in the hidden layer, and when the learning rate and the momentum values were 0.025 and 0.74, respectively. The RMSE for prediction set were minimum after 9000 iterations. The calculated absolute errors for the prediction and validation sets are presented in Tables 3 and 4, respectively. The RMSE and mean absolute errors were, respectively, 0.250 and 0.164 for prediction set and 0.197 and 0.126 for validation set. The performance of the model was also evaluated by plotting the solubility estimated by the model against those of the experimental values for the validation set (Fig. 3).

The results show that the WNN converge faster than ANN and the mean absolute error for the test set is smaller in WNN compared to ANN.

4.3. Prediction of solubilities for new azo dyes

The capability of the model was also evaluated for prediction of the solubility of five azo dyes, their data were not used in any of the previous data sets. These azo dyes were Solvent Brown 1, Red 82, Red 13, Modified Yellow 119 and Yellow 16. The structures of these dyes are given in Table 1. The calculated descriptors for these dyes are presented in Table 2. The constructed WNN model was used to predict the solubility of these dyes. The mean absolute errors were 0.349 for Yellow 16, 1.23 for Solvent Brown 1, 0.61 for Red 82, 0.27 for Modified Yellow 119 and 0.98 for Red 13.

5. Conclusions

Only one WNN model was constructed to predict the solubility of 16 azo dyes in supercritical carbon dioxide over a wide range of pressures (100–355 bar) and temperatures

(308–413 K). The performance of the model was evaluated by the validation set and also with a data set, the molecules and data of which were new for the model. The ability of the WNN model was also compared with ANN and MLR models. It was demonstrated that WNN is superior to ANN and ANN has a better performance than MLR.

Acknowledgement

The authors acknowledge the Research Council of Isfahan University of Technology and Center of Excellency in Chemistry of Isfahan University of Technology for the support of this work.

References

- [1] Saus W, Knittel D, Schollmeyer E. Dyeing of textile in supercritical carbon dioxide. *Text Res J* 1993;63:135.
- [2] Gebert B, Saus W, Knittel D, Bushman HJ, Shollmeyer E. Dyeing natural fibers with disperse dyes in supercritical carbon dioxide. *Text Res J* 1994;64:371.
- [3] Knittel W, Schollmeyer E. Dyeing with supercritical CO₂, research with different types of dyes. *Melliand Textilber* 1995;76:1092.
- [4] Bach E, Cleve E, Scholleyer E. Dyeing of poly(ethylene terephthalate) fibers in supercritical carbon dioxide. In: von Rohr PhR, Trepp Ch, editors. *High-pressure chemical engineering*. Amsterdam: Elsevier Science; 1996. p. 581–6.
- [5] Chang KH, Bae HK, Shim JJ. Dyeing of PTE textile fibers and films in supercritical carbon dioxide. *Korean J Chem Eng* 1996;13:310.
- [6] Ozcan AS, Clifford AA, Bartle KD, Lewis DM. Solubility of disperse dyes in supercritical carbon dioxide. *J Chem Eng Data* 1997;42:590.
- [7] Joung SN, Yoo KP. Solubility of disperse anthraquinone and azo dyes in supercritical carbon dioxide at 313.15 to 393.15 K and from 10–25 MPa. *J Chem Eng Data* 1998;43:9.
- [8] Guzel B, Akgerman A. Solubility of disperse and mordant dyes in supercritical CO₂. *J Chem Eng Data* 1999;44:83.
- [9] Sung HD, Shim JJ. Solubility of C.I. disperse red 60 and C.I. disperse blue 60 in supercritical carbon dioxide. *J Chem Eng Data* 1999;44:985.
- [10] Lee JW, Min JM, Bae HK. Solubility measurement of disperse dyes in supercritical carbon dioxide. *J Chem Eng Data* 1999;44:684.
- [11] Lee JW, Park MW, Bae HK. Measurement and correlation of dye solubility in supercritical carbon dioxide. *Fluid Phase Equilib* 2001;179:387.
- [12] Lin HM, Liu CY, Chang CH, Chen YT, Lee MJ. Solubility of disperse dyes of blue 79, red 153 and yellow 119 in supercritical carbon dioxide. *J Supercrit Fluids* 2001;21:1.
- [13] Draper SL, Montero GA, Smith B, Beck K. Solubility relationships for disperse dyes in supercritical carbon dioxide. *Dyes Pigments* 2000;45:177.
- [14] Shamsipur M, Karami AR, Yamini Y, Sharghi H. Solubility of some 1-hydroxy-9,10 anthraquinone derivatives in supercritical carbon dioxide. *J Supercrit Fluids* 2004;32:47.
- [15] Fasihi J, Yamini Y, Nourmohammadian F, Bahramifar N. Investigations on the solubilities of some disperse azo dyes in supercritical carbon dioxide. *Dyes Pigments* 2004;63:161.
- [16] Fat'hi MR, Yamini Y, Sharghi H, Shamsipur M. Solubility of some recently synthesized 1,8-dihydroxy-9,10-anthraquinone derivatives in supercritical carbon dioxide. *Talanta* 1999;48:951.
- [17] Mishima K, Matsuyama K, Ishikawa H, Hayashi K, Maeda S. Measurement and correlation of solubilities of azo dyes and anthraquinone in supercritical carbon dioxide. *Fluid Phase Equilib* 2002;194–197:895.
- [18] Kautz CB, Wagner B, Schneider GM. High-pressure solubility of 1,4-bis-(*n*-alkylamino)-9,10-anthraquinones in near and supercritical carbon dioxide. *J Supercrit Fluids* 1998;13:43.
- [19] Tuma D, Wagner B, Schneider GM. Comparative solubility investigations of anthraquinone disperse dyes in near and supercritical fluids. *Fluid Phase Equilib* 2001;182:133.
- [20] Wagner B, Kautz CB, Schneider GM. Investigations on the solubility of anthraquinone dyes in supercritical carbon dioxide by a flow method. *Fluid Phase Equilib* 1999;158–160:707.
- [21] Tuma D, Schneider GH. High-pressure solubility of disperse dyes in near and supercritical fluids: measurements up to 100 MPa by a static method. *J Supercrit Fluids* 1998;13:37.
- [22] Shinoda T, Tamura K. Solubility of C.I. disperse red 1 and C.I. disperse red 13 in supercritical carbon dioxide. *Fluid Phase Equilib* 2003;213:115.
- [23] Lin HM, Ho CC, Lee MJ. Solubilities of disperse dyes of blue 79:1, red 82 and modified yellow 119 in supercritical carbon dioxide and nitrous oxide. *J Supercrit Fluids* 2004;32:105.
- [24] Tamura K, Shinoda T. Binary and ternary solubilities of disperse dyes and their blend 13 in supercritical carbon dioxide. *Fluid Phase Equilib* 2004;219:25.
- [25] Chrastil J. Solubility of solids and liquids in supercritical gases. *J Phys Chem* 1982;86:3016.
- [26] Mendez-Santiago J, Teja AS. The solubility of solids in supercritical fluids. *Fluid Phase Equilib* 1999;158–160:501.
- [27] Del Valle JM, Aguilera JM. An improved equation for predicting the solubility of vegetable oils in supercritical CO₂. *Ind Eng Chem Res* 1998;27:1551.
- [28] Gordillo MD, Blanco MA, Molero A, Martinez de la Ossa E. Solubility of the antibiotic penicillin G in supercritical carbon dioxide. *J Supercrit Fluids* 1999;15:183.
- [29] Jouyban A, Chan H, Foster NR. Mathematical representation of solute solubility in supercritical carbon dioxide using empirical expressions. *J Supercrit Fluids* 2002;24:19.
- [30] Ferri A, Banchemo M, Manna L, Sicardi S. A new correlation of solubilities of azoic compounds and anthraquinone derivatives in supercritical carbon dioxide. *J Supercrit Fluids* 2004;32:27.
- [31] Peng DY, Robinson DB. A new two constant equation of state. *Ind Eng Chem Fundam* 1976;15:59.
- [32] Politzer P, Murray JS, Lane P, Brink T. Relationships between solute molecular properties and solubility in supercritical CO₂. *J Phys Chem* 1993;97:729.
- [33] Politzer P, Lane P, Murray JS, Brink T. Investigation of relationships between solute molecule surface electrostatic potentials and solubility in supercritical fluids. *J Phys Chem* 1992;96:7938.
- [34] Huang Z, Kawi S, Chiew YC. Application of the perturbed Lennard-Jones chain equation of state to solute solubility in supercritical carbon dioxide. *Fluid Phase Equilib* 2004;216:111.
- [35] Battersby P, Dean JR, Tomlinson WR, Hitchen SM, Myers P. Predicting solubility in supercritical fluid extraction using a neural network. *Analyst* 1994;119:925.
- [36] Khayamian T, Esteki M. Prediction of solubility for polycyclic aromatic hydrocarbons in supercritical carbon dioxide using wavelet neural networks in quantitative structure property relationship. *J Supercrit Fluids* 2004;32:73.
- [37] Daubechies I, editor. *Ten lectures on wavelets*. Philadelphia: SIAM Press; 1992.
- [38] Leung AK, Chau F, Gao J. A review on application of wavelet transform techniques in chemical analysis: 1989–1997. *Chemom Intell Lab Syst* 1998;43:165.
- [39] Walczak B. *Wavelets in chemistry*. Amsterdam: Elsevier Science; 2000.
- [40] Zhang X, Qi J, Zhang R, Liu M, Hu Z, Xue H, et al. Prediction of programmed-temperature retention values of naphthas by wavelet neural network. *Comput Chem* 2001;25:125.
- [41] Guo Q, Liu L, Cai W, Liu Y. Driving force prediction for inclusion complexation of α -cyclodextrin with benzene derivatives by a wavelet neural network. *Chem Phys Lett* 1998;514:518.
- [42] Liu L, Guo Q. Wavelet neural network and its application to inclusion of β -cyclodextrin with benzene derivatives. *J Chem Inf Comput Sci* 1999;39:133.
- [43] Tabaraki R, Khayamian T, Ensafi AA. Principle component-wavelet neural networks as a new multivariate calibration method. *Anal Lett* 2005;38:1–13.
- [44] Karelson M, Lobanov VS, Katritzky AR. Quantum-chemical descriptors in QSAR/QSPR studies. *Chem Rev* 1996;96:1027.
- [45] Delgado EJ, Alderete JB, Matamala AR, Jana GA. On the aggregation state and QSPR models. The solubility of herbicides as a case study. *J Chem Inf Comput Sci* 2004;44:958.