A Fuzzy ARTMAP Based on Quantitative Structure−Property Relationships (QSPRs) for Predicting Aqueous Solubility of Organic Compounds

Denise Yaffe,[†] Yoram Cohen,*[,†] Gabriela Espinosa,[‡] Alex Arenas,[§] and Francesc Giralt[‡]

Department of Chemical Engineering, University of California, Los Angeles, Los Angeles, California 90095-1592, and Departament d'Enginyeria Química, ETSEQ and Departament d'Enginyeria Informàtica, ETSE, Universitat Rovira i Virgili, 43006 Tarragona, Catalunya, Spain

Received February 22, 2001

Quantitative structure−property relationships (QSPRs) for estimating aqueous solubility of organic compounds at 25 °C were developed based on a fuzzy ARTMAP and a back-propagation neural networks using a heterogeneous set of 515 organic compounds. A set of molecular descriptors, developed from PM3 semiempirical MO-theory and topological descriptors (first-, second-, third-, and fourth-order molecular connectivity indices), were used as input parameters to the neural networks. Quantum chemical input descriptors included average polarizability, dipole moment, resonance energy, exchange energy, electron−nuclear attraction energy, and nuclear−nuclear (core−core) repulsion energy. The fuzzy ARTMAP/QSPR correlated aqueous solubility (S, mol/L) for a range of −11.62 to 4.31 logS with average absolute errors of 0.02 and 0.14 logS units for the overall and validation data sets, respectively. The optimal 11−13−1 back-propagation/QSPR model was less accurate, for the same solubility range, and exhibited larger average absolute errors of 0.29 and 0.28 logS units for the overall and validation sets, respectively. The fuzzy ARTMAP-based QSPR approach was shown to be superior to other back-propagation and multiple linear regression/QSPR models for aqueous solubility of organic compounds.

I. INTRODUCTION

Predictions of aqueous solubilities of organic compounds, without prior knowledge of experimentally determined information, are of particular interest in industrial and environmental applications. For example, aqueous solubility of organic compounds is an important factor for assessing the mass distribution of chemicals in the environment and for determining bioavailability of candidate drugs in drug discovery and development.

As a complement to experimental aqueous solubility data, numerous approaches for estimating aqueous solubility have been proposed in the literature.[1] Most of these approaches rely on correlations of solubility with experimentally derived parameters such as melting and boiling point temperatures, chromatographic retention times, activity coefficient (at infinite dilution) molar volume (derived from liquid density), partition coefficient, or parachor (derived from density and surface tension). Group contribution approaches for estimating aqueous solubility have also been proposed.[2−5] Other popular methods include quantitative structure−property relationships (QSPRs) whereby molecular structural characteristics (e.g., geometric and electronic), expressed in terms of appropriate molecular descriptors, are correlated with the desired physicochemical properties. Due to the wide range of aqueous solubility, S (mol/L), for organic compounds it is common to report and correlate the aqueous solubility as logS, and this form is also adopted in the present work.

Correlations of solubility with other physicochemical parameters can be useful provided that data for the correlating parameters are available. An example of such an approach was reported by Yalkowsky and Valvani[6] who correlated aqueous solubility, at 25 °C, for 167 organic compounds (mono- and polyhalogenated aromatics, alcohols, alkyl *p*-hydroxybenzoates, and alkyl *p*-aminobenzoates) using melting point, octanol/water partition coefficient (K_{ow}), and entropy of fusion as input parameters. The proposed logS correlation, with solubility values ranging over 9 orders of magnitude, performed with an average absolute error of 0.5 logS units. In a later study, Isnard and Lambert,[7] for a set of 300 structurally diverse compounds, developed aqueous solubility correlations for liquids (melting point < 25° C) and solids (melting point > 25° C) using K_{ow}, and both K_{ow} and melting point, respectively, with reported standard deviations for liquids and solids of 0.466 and 0.582 logS units, respectively. Warne et al.[8] used melting point and the approximate sigma electron density term (ASED) to correlate the solubility of a data set of 16 compounds consisting of alkanes, alkyl substituted benzenes, and naphthalenes ($-5.1 \leq \log S \leq -1.6$) with an average absolute error of about 64%. Correlations of logS with normal boiling points have also been proposed for specific chemical families, such as chlorobenzenes,[9] alkanes, and cycloalkanes.[10] For example, Miller et al.[9] correlated the solubility of 12 chlorobenzenes with boiling point with an average absolute logS error of 7.16%. In a similar study, Yaw et al.[10] proposed a limited logS correlation with boiling temperature for liquid alkanes (at 25 °C) with five or more carbon atoms. The correlation, based on 26 alkanes, exhibited an average absolute error and standard deviation of 0.05 (1.04%) and 0.05 (1.1%) logS units, respectively. In another study a linear multiparameter logS correlation, based on 531 heterogeneous compounds ($12.79 \leq \log S \leq -0.51$), proposed by Ruelle and Kesselring[11]

* Corresponding author phone: (310)825-8766; fax: (310)477-3868; e-mail: yoram@ucla.edu.
[†] University of California.
[‡] ETSEQ.
[§] ETSE.

10.1021/ci010323u CCC: $20.00
Published on Web 08/25/2001

using melting point, molar volume, and an additional term accounting for solvation effects as input performed with an average absolute error of 0.37 l logS units.

Group contribution approaches for estimating aqueous solubility have also been proposed in the literature.[2−5] For example, Wakita et al.[2] proposed a group contribution correlation based on 40 fragment terms derived from 307 liquid compounds, two additional fragment terms derived from 15 aliphatic solids, and five additional fragment terms derived from 112 aromatic solids. The fragments for solid compounds also included melting point correction factors, one for aliphatics and another for aromatics. Aqueous solubility for 314 aliphatic and aromatic liquids ($-5.24 \leq \log S \leq 0.68$) were correlated with a standard deviation and average absolute error of 0.25 and 020 logS, respectively. Predictions for aliphatic solids ($-0.01 \leq \log S \leq 0.17$) were with a standard deviation of 0.151 logS units. These authors also reported a solubility correlation for 134 aromatic solids ($-10.49 \leq \log S \leq 0.34$) with a standard deviation and average absolute error of 0.65 and 0.58 logS units, respectively. In another study, Suzuki[3] proposed a correlation for estimating solubilities ($-10.49 \leq \log S \leq 1.96$) based on octanol/water partition coefficients ($\log K_{ow}$), which were calculated from 10 individual atom type fragments that were derived from 497 compounds and melting points. Suzuki[3] reported a standard deviation of 0.505 logS units and an average absolute error of 0.39 logS units for the above correlation. Klopman et al.[4] proposed two group contribution models that focused on pharmaceutical drugs, the first consisting of 45 fragments and one constant (based on 496 organic compounds) and the second consisting of 33 fragments (based on 483 compounds containing halogen, nitrogen, sulfur, oxygen). Klopman et al.[4] reported average standard deviations, based on 10 cross-validation test sets, of 0.503 and 0.55 logS units for model 1 and model 2, respectively. A test of the above models 1 and 2 with a 13 and 21 compound data sets, respectively, yielded standard deviations of 0.58 and 1.25 logS units, respectively. Finally, Kühne et al.[5] proposed a logS correlation, with a group contribution term (based on 55 fragments) and two melting point terms derived from a heterogeneous set of 694 organic compounds ($-11.62 \leq \log S \leq 1.81$), that performed with an average absolute error of 0.38 logS units. Kühne et al.[5] compared their group contribution model with those proposed by Wakita et al.,[2] Suzuki,[3] and Klopman et al.,[4] based on a heterogeneous set of 694 test compounds representing various functional groups. The average absolute aqueous solubility estimation errors for the Wakita et al.,[2] Suzuki,[5] and models 1 and 2 of Klopman et al.,[5] in terms of logS units were reported to be 0.55 (based on 543 from test set), 0.85 (based on 672 from the test set), 0.50 (based on 614 from the test set), and 0.56 (for the complete test set), respectively.[5]

A variety of QSPRs approaches have been reported in the literature to estimate logS for organic compounds.[12−22] The premise of QSPRs is that there is a unique relationship between molecular chemical descriptors and physicochemical properties. Given a selected set of molecular descriptors, one searches for optimized correlations between the descriptors and the desired chemical specific property using partial least-squares or artificial neural networks to build the QSPR model.[12−16,23−32] Medir and Giralt,[17] for example, correlated the inverse of solubility (i.e. $1/x_s = \ln \gamma^{\infty}$, where γ^{∞} is the activity coefficients at infinity dilution and x_s is the solute mole fraction at solubility) for hydrocarbons (42 aliphatic, 12 aromatic, and 30 polycyclic aromatic hydrocarbons, PAHs). The correlations were based on molecular input descriptors which included a zero-order molecular connectivity index, dipole moment, number of carbon atoms, and acentric factor. Standard deviations for the above correlation for the aliphatic, aromatic, and polycyclic aromatic hydrocarbons were 0.336 to 1.26, 0.18, and $0.558-0.545 \ln \gamma^{\infty}$ $u(1/x_s)$ units, respectively. A simpler linear regression model for logS was proposed by Nirmalakhandan and Speece[18,19] using as inputs the zero-order valence molecular connectivity index and a polarizability parameter estimated from group contribution.[18,19] The above model, developed based on a set of 145 compounds,[18] performed with an overall standard error of 0.311 logS. In a later study, Nirmalakhandan and Speece[19] broadened the applicability of the model to include PAHs, PCBs, chlorophenols, dioxins, and ketones by modifying the polarizability parameter. This latter model was tested using a diverse set of 325 compounds ($-9.32 \leq \log S \leq -3.03$) and performed with an overall standard error of 0.33 logS units. A correlation specific to PCBs, based on 71 compounds, was proposed by Patil,[20] using the first-order valence molecular connectivity index as a molecular descriptor, with a reported average absolute error of 0.45 logS units and logS standard deviation of 9.4%. Patil[20] reported that the inclusion of melting point as an additional parameter offered modest reductions of the average absolute error of 0.30 logS units, corresponding to a 8.7% logS error. Another PCB-specific linear logS correlation was recently proposed by Makino[14] using a set of six descriptors (molecular weight, heat of formation, solvent accessible surface area, ionization potential, electron affinity, and dipole moment), based on a data set of 136 PCB congeners ($-10.32 \leq \log S \leq -5.33$) with a reported standard deviation and average absolute error of 0.225 and 0.1681 logS units, respectively.[14] A more recent study by Huibers and Katritzky[15] reported a logS correlation for 109 hydrocarbons and 132 halogenated hydrocarbons ($-10.41 \leq \log S \leq 0.51$) with a set of molecular descriptors that included molecular volume, bonding information content[33] (a topological descriptor encoding both the degree of branching and number of different atom types in the molecule), and atomic charge weighted partial negative surface area[22] (sum of atomic surface areas weighted both by atomic surface charge and partial atomic charge for negatively charged surfaces only); the above model performed with a standard error of 0.386 logS units. In a related study, Katritzky et al.[16] proposed a linear correlation model for logS, based on 411 compounds ($-6.44 \leq \log S \leq 1.57$) using the following six descriptors: a most negative partial charge, relative negative charged surface area, number of electrons, average bond order of nitrogen, zero-order structural information content, and fractional hydrogen donor surface area; the above model performed with a standard error of 0.573 logS units.

In recent years, neural networks (NNs) have gained popularity as an alternative approach to developing QSPRs. The advantage of using artificial neural networks is in their inherent capability of modeling nonlinear relationships between chemical descriptors and physicochemical properties without a priori specification of the analytical form of the QSPR model.[12,13.21,27,28] For example, back-propagation neural

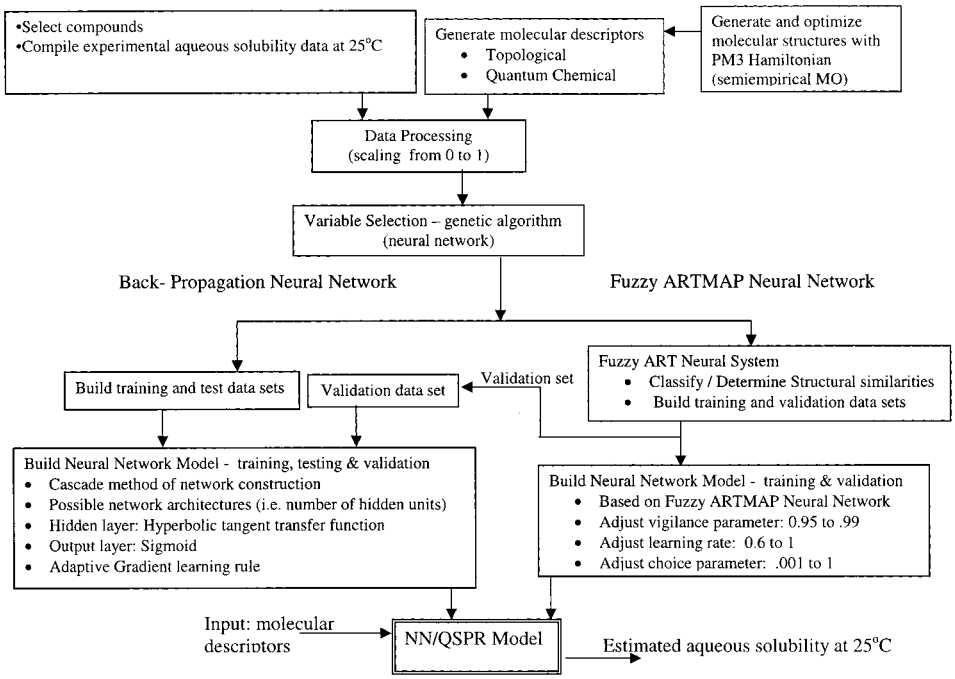

Figure 1. Process flow diagram for developing fuzzy ARTMAP and back-propagation QSPR/neural networks for the prediction of aqueous solubility.

network based logS QSPR models have been proposed by Huuskonen et al.,[21] Sutter and Jurs,[12] and Mitchell and Jurs.[13] Huuskonen et al.,[21] using 14 atom-type electropological indices and nine topological indices, developed a logS QSPR model based on a 23−5−1 neural network using a data set of 211 pharmaceutical drugs ($-5.82 \leq \log S \leq 5.45$) with a reported average error, based on a test set of 51 compounds, of 0.53 logS units. Sutter and Jurs[12] proposed a logS QSPR model based on a 9−3−1 neural network architecture, developed using a diverse set of 140 compounds ($-10.83 \leq \log S \leq 0.28$), using electrical, topological, and geometrical descriptors as inputs, with reported root-mean square (rms) errors for cross-validation and predictions sets of 0.282 and 0.222 logS units, respectively. In the above study polychlorinated biphenyls contributed significantly to the rms error; however, upon their removal from the data sets, and with a change in four of the nine original input descriptors, the performance of the retrained neural network-based logS model improved with rms errors for the cross-validation and prediction sets decreasing to 0.151 and 0.166 logS units, respectively. In a more recent study, Mitchell and Jurs[13] presented two models based on a 9−6−1 neural network architecture, for predicting logS for 332 diverse organic compounds ($-12 \leq \log S \leq 2$) using two different sets of input descriptors. The molecular descriptors were based on topological, geometric, charge partial surface area, and encoded hydrogen bonding input parameters. The first model, with nine descriptors, resulted in rms errors for the training, cross-validation, and prediction sets of 0.460, 0.455, and 0.446 logS units, respectively. The second model[13] had a lower rms of 0.394, 0.358, and 0.343 logS units, for the training, cross-validation, and prediction sets, respectively. The above study clearly demonstrated that the accuracy of the logS QSPR depends on the selection of input descriptors.

To date, published neural network-based logS QSPRs have been developed using back-propagation neural networks.[12,13,21,27] Recent work by Espinosa et al.[28] suggests that it may be possible to improve neural network-based QSPRs by using a cognitive classifier fuzzy ARTMAP neural network. The approach, which was demonstrated for estimating the boiling point temperatures of aliphatic hydrocarbons, was shown to be superior to the back-propagation neural network approach as well as other statistical QSPR correlations reported in the literature. The application of fuzzy ARTMAP networks for QSPR development has several advantages owing to their capability to classify and analyze noisy information with fuzzy logic and to avoid the plasticity-stability dilemma of standard back-propagation architectures.[34−39]

The purpose of the current study was to investigate the feasibility of developing a fuzzy ARTMAP QSPR model for aqueous solubility based on a heterogeneous set of organic compounds using chemical descriptors obtained from PM3 semiempirical MO-theory calculations. The performance of the fuzzy ARTMAP model was compared to a back-propagation neural network based QSPR, developed with the same set of input descriptors, and with other published multiple linear regression (MLR) and back-propagation neural network QSPR models.

II. METHODOLOGY

Data Set and Molecular Descriptors. The method utilized in the present study to develop QSPRs is summarized in Figure 1. The aqueous solubility data at 25 °C for the diverse set of 515 compounds considered in this study were compiled from the literature.[11−16,19,23,40] The heterogeneous set of compounds includes aromatic (polycyclic aromatic) and aliphatic (normal, branched, cyclic) hydrocarbons, halogens, polychlorinated biphenyls, mercaptans, sulfides, anilines, pyridines, alcohols, carboxylic acids, aldehydes, amines, ketones, and esters (see Table 1). Aqueous solubility

Table 1. Molecular Descriptors and Experimental Aqueous Solubility Data at 25 °C

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
1,1,1,2-tetrachlorodifluoroethane	25	1.13	49.22	−86.63	−33.38	−10837.59	5495.16	2.89	3.74	1.60	0.00	3.31
1,1,1,2-tetrachloroethane	19	1.21	47.38	−73.19	−34.87	−6245.29	3172.28	2.85	3.52	1.36	0.00	2.18
1,1,1-trichloroethane	16	1.38	40.13	−73.77	−36.11	−4580.62	2336.30	2.20	3.62	0.00	0.00	2.00
1,1,1-trichloropropane	19	1.42	46.76	−106.54	−52.77	−6454.49	3288.48	2.76	3.48	1.20	0.00	1.89
1,1,2,2,-tetrachlorodifluoroethane	25	0.00	51.21	−85.38	−33.16	−10793.69	5472.55	2.89	3.46	1.75	0.00	3.19
1,1,2,2-tetrabromoethane	19	1.01	60.04	−65.00	−33.11	−6015.60	3052.42	4.86	7.06	5.13	0.00	2.72
1,1,2,2-tetrachloroethane	19	1.52	46.34	−72.84	−34.75	−6216.74	3158.20	2.95	2.99	1.71	0.00	1.74
1,1,2-trichloroethane	16	1.08	38.51	−73.11	−35.88	−4445.43	2269.66	2.52	2.13	1.05	0.00	1.48
1,1,2-trichloropropane	19	1.39	45.77	−105.90	−52.54	−6353.86	3239.74	2.87	2.86	1.61	0.00	1.89
1,1,3,4,4-pentachloro-1,2-butadiene	26	0.90	82.03	−119.84	−54.89	−10430.07	5287.51	3.79	3.48	1.56	0.65	4.23
1,1,3-trimethylcyclopentane	24	0.04	54.14	−261.59	−133.48	−10839.62	5533.59	3.60	4.01	2.21	1.78	4.48
1,1-dichlorobutane	19	1.66	44.78	−139.21	−70.47	−6352.68	3248.95	2.92	2.45	0.94	0.65	2.40
1,1-dichloroethane	13	1.62	30.76	−73.70	−37.15	−3094.89	1589.69	1.88	2.05	0.00	0.00	1.29
1,1-dichloroethene	12	0.78	31.89	−64.43	−31.18	−2571.75	1320.92	1.49	1.44	0.00	0.00	1.63
1,1-dichlorotetrafluoroethane	25	1.34	33.22	−101.54	−34.26	−11133.76	5666.48	2.14	2.23	0.75	0.00	2.93
1,1-difluoroethane	13	2.12	15.34	−87.57	−37.89	−3295.02	1705.79	1.01	0.52	0.00	0.00	0.57
1,1-difluoroethene	12	1.59	16.88	−78.12	−31.67	−2719.66	1413.38	0.73	0.34	0.00	0.00	2.55
1,1-diphenylethylene	34	0.04	114.82	−377.15	−181.96	−18244.33	9300.94	4.67	3.21	2.24	1.45	4.44
1,2,3,4-tetrahydronaphthalene	26	0.48	75.34	−288.48	−141.29	−12063.08	6163.87	4.03	2.98	2.26	1.68	4.37
1,2,3,5-tetrafluorobenzene	27	1.55	51.42	−195.00	−77.97	−12301.57	6282.65	2.41	1.68	0.95	0.52	2.31
1,2,3,5-tetramethylbenzene	27	0.26	79.44	−298.91	−147.66	−12635.74	6451.84	3.65	3.02	2.07	1.25	3.76
1,2,3-trichloropropane	19	2.29	44.11	−105.69	−52.47	−6214.64	3170.03	3.07	2.14	1.70	0.37	1.93
1,2,3-trimethylbenzene	24	0.52	70.68	−265.49	−130.86	−10420.12	5330.38	3.24	2.52	1.88	0.90	3.20
1,2,4,5-tetrafluorobenzene	27	0.00	51.64	−194.88	−77.96	−12278.93	6271.18	2.41	1.68	0.96	0.50	2.38
1,2,4,5-tetramethylbenzene	27	0.02	79.74	−298.97	−147.68	−12610.12	6439.02	3.65	3.02	2.11	1.10	3.84
1,2,4-trichlorobenzene	24	0.67	76.04	−168.40	−77.98	−9681.27	4930.54	3.44	2.81	1.85	1.00	3.59
1,2,4-trimethylbenzene	24	0.23	71.24	−265.56	−130.91	−10280.73	5260.67	3.24	2.59	1.66	0.89	3.32
1,2:5,6-dibenzanthracene (DB[a,h])	51	0.01	218.12	−574.98	−271.18	−35724.76	18137.19	7.63	5.87	4.61	3.51	7.44
1,2:7,8-dibenzanthracene	51	0.04	214.66	−574.99	−271.18	−35948.11	18248.99	7.63	5.87	4.61	3.51	8.66
1,2-benzofluorene	40	0.36	151.93	−449.68	−213.30	−24310.70	12368.87	6.02	4.65	3.72	2.91	6.68
1,2-dibromo-3-chloropropane	19	2.45	50.35	−101.53	−51.55	−6146.61	3133.69	4.14	3.15	3.37	0.64	2.38
1,2-dichlorobenzene	21	1.35	64.47	−167.34	−78.84	−7854.59	4012.73	2.96	2.23	1.58	0.71	2.97
1,2-dichloroethane	13	0.01	30.35	−73.36	−37.04	−2988.66	1537.31	2.10	1.13	0.64	0.00	1.06
1,2-dichloropropane	16	0.37	37.75	−106.17	−53.69	−4655.73	2386.60	2.44	1.99	0.99	0.00	1.61
1,2-dichlorotetrafluoroethane	25	1.33	33.23	−101.54	−34.26	−11133.28	5666.24	2.14	2.23	0.75	0.00	2.74
1,2-dichoroethene (trans)	12	0.00	32.74	−64.30	−31.08	−2473.12	1272.78	1.64	0.75	0.43	0.00	1.44
1,2-diethylbenzene	27	0.53	76.68	−297.35	−147.20	−12855.28	6562.17	3.95	2.48	1.88	1.25	3.28
1,2-difluorobenzene	21	2.74	48.42	−179.86	−79.19	−8043.96	4125.32	2.21	1.41	0.82	0.44	2.00
1,2-diiodobenzene	21	0.76	80.16	−158.90	−77.08	−7514.25	3836.81	4.33	3.70	3.67	1.20	4.24
1,2-diiodoethene	12	0.00	53.34	−55.84	−29.25	−2267.03	1166.00	3.22	1.67	2.08	0.00	7.37
1,3,5-trichlorobenzene	24	0.00	76.51	−168.33	−77.94	−9571.91	4877.22	3.43	2.89	1.48	1.39	4.48
1,3,5-trimethylbenzene	24	0.00	71.01	−265.52	−130.97	−10198.32	5219.43	3.23	2.67	1.37	1.20	3.40
1,3-butadiene	11	0.00	35.26	−117.21	−59.84	−2449.24	1279.62	1.15	0.47	0.17	0.00	2.13
1,3-dichlorobenzene	21	0.88	65.47	−167.21	−78.77	−7720.45	3947.05	2.95	2.31	1.26	0.90	3.07
1,3-dichloropropane	16	1.50	36.98	−106.15	−53.70	−4460.15	2288.74	2.60	1.49	0.80	0.45	1.62
1,3-difluorobenzene	21	1.57	48.32	−179.72	−79.11	−7921.08	4064.91	2.20	1.44	0.79	0.46	2.00
1,3-dimethylnaphthalene	30	0.36	101.82	−334.37	−161.81	−15112.34	7711.13	4.23	3.30	2.21	1.69	4.29
1,4,5-trimethylnaphthalene	33	0.30	110.09	−367.46	−178.37	−18050.57	9194.51	4.65	3.70	2.68	1.94	4.91
1,4-cyclohexadiene	16	0.00	44.66	−175.19	−87.03	−5178.16	2670.02	2.30	1.41	0.88	0.54	2.06
1,4-cyclopentadiene	13	0.53	36.59	−141.10	−70.06	−3569.04	1851.19	1.82	1.09	0.66	0.39	2.06
1,4-dichlorobenzene	21	0.00	65.99	−167.22	−78.77	−7697.52	3935.70	2.95	2.31	1.31	0.68	3.27
1,4-dichlorobutane	19	0.02	44.28	−138.88	−70.35	−6078.57	3017.07	3.10	1.84	1.05	0.57	0.92
1,4-diethylbenzene	27	0.09	77.34	−297.54	−147.45	−12291.18	6280.96	3.94	2.52	1.84	1.01	3.75
1,4-difluorobenzene	21	0.00	48.48	−179.66	−79.12	−7896.84	4052.66	2.20	1.44	0.80	0.43	1.97
1,4-dimethylnaphthalene	30	0.01	101.17	−334.31	−161.77	−15260.35	7785.13	4.24	3.24	2.36	1.60	4.14
1,4-pentadiene	14	0.08	39.50	−150.04	−76.52	−3792.36	1966.28	1.63	0.81	0.33	0.12	2.09
1,5-dichloropentane	22	1.52	51.34	−171.61	−87.00	−7824.17	4001.43	3.60	2.19	1.30	0.74	3.05
1,5-dimethylnaphthalene	30	0.01	100.98	−334.30	−161.76	−15251.25	7780.64	4.24	3.24	2.36	1.62	4.74
1,5-hexadiene	17	0.15	46.40	−182.80	−93.18	−5312.35	2740.90	2.13	1.15	0.57	0.24	2.69
1,6-heptadiene	20	0.08	53.70	−215.57	−109.85	−6966.89	3582.90	2.63	1.51	0.82	0.41	3.34
1,6-heptadiyne	19	0.41	53.76	−207.64	−103.69	−7054.50	3621.96	2.82	1.79	1.12	0.70	1.75
10-butylbenz[*a*]anthracene	54	0.46	206.70	−604.10	−290.29	−37864.83	19215.90	8.19	6.15	4.59	3.34	7.55
10-ethylbenz[*a*]anthracene	48	0.46	191.65	−538.54	−256.96	−32063.27	16285.40	7.19	5.40	4.17	3.02	6.78
10-methylbenz[*a*]anthracene	45	0.37	184.08	−505.82	−240.32	−29059.09	14768.67	6.63	5.21	3.86	2.87	6.64
10-pentylbenz[*a*]anthracene	57	0.39	214.16	−636.86	−306.91	−40774.34	20685.40	8.69	6.50	4.84	3.54	8.57
12-methylbenz[*a*]anthracene	45	0.29	181.99	−505.34	−240.10	−29848.52	15163.46	6.64	5.11	4.01	3.09	6.68
1-bromo-2-chloroethane	13	0.21	34.82	−71.17	−36.54	−2938.97	1512.11	2.69	1.55	1.11	0.00	1.32
1-bromobutane	16	1.82	40.18	−137.29	−71.17	−4648.46	2391.56	3.09	1.83	1.05	0.69	2.20
1-bromoheptane	25	1.85	61.58	−235.56	−121.16	−9871.38	5047.57	4.59	2.89	1.80	1.09	4.42
1-bromohexane	22	1.84	54.47	−202.81	−104.50	−8013.44	4103.83	4.09	2.54	1.55	0.92	3.81
1-bromonaphthalene	27	1.13	95.91	−265.76	−126.70	−12460.09	6361.02	4.30	3.31	2.43	1.70	4.35
1-bromopentane	19	1.84	47.36	−170.05	−87.83	−6267.21	3215.92	3.59	2.19	1.30	0.74	3.08

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
1-bromopropane	13	1.81	32.94	−104.54	−54.50	−3179.11	1641.91	2.59	1.48	0.98	0.00	1.70
1-butanethiol	15	2.38	39.08	−140.17	−71.42	−4401.67	2270.50	2.65	1.52	0.83	0.47	2.18
1-butene	12	0.21	30.32	−128.69	−66.44	−2949.97	1533.13	1.52	0.70	0.29	0.00	2.40
1-chloro-1,1-difluoroethane	16	2.02	24.57	−87.13	−36.66	−4788.00	2458.59	1.44	1.44	0.00	0.00	1.20
1-chloro-2-fluoroethane	13	0.27	22.08	−80.23	−37.45	−3080.01	1589.87	1.57	0.76	0.21	0.00	0.51
1-chloro-2-methylpropane	16	1.51	35.82	−139.46	−71.70	−4897.42	2515.70	2.36	1.86	0.92	0.00	2.00
1-chlorobutane	16	1.56	36.47	−139.47	−71.68	−4699.57	2417.19	2.51	1.42	0.75	0.40	2.03
1-chlorohexane	22	1.58	50.68	−205.00	−105.01	−8067.29	4130.72	3.51	2.13	1.25	0.71	3.12
1-chloronaphthalene	27	0.91	93.35	−268.56	−127.34	−12556.71	6409.95	3.89	2.87	2.07	1.43	3.93
1-chlorooctane	28	1.58	64.88	−270.47	−138.33	−11882.28	6067.62	4.51	2.83	1.75	1.06	4.48
1-chloropentane	19	1.57	43.60	−172.23	−88.35	−6319.31	3241.95	3.01	1.77	1.00	0.53	2.73
1-choropropane	13	1.55	29.32	−106.72	−55.02	−3228.10	1666.54	2.01	1.07	0.57	0.00	1.46
1-decanol	34	1.41	74.65	−355.92	−178.17	−16317.42	8326.15	5.02	3.20	2.01	1.25	3.63
1-decene	30	0.21	73.06	−325.24	−166.43	−13393.87	6842.71	4.52	2.85	1.76	1.07	5.51
1-ethylnaphthalene	30	0.31	99.72	−333.75	−161.57	−15148.71	7730.28	4.38	2.99	2.26	1.60	4.17
1-heptanol	25	1.41	53.45	−257.59	−128.17	−10183.81	5215.52	3.52	2.14	1.26	0.72	1.95
1-heptyne	20	0.37	49.93	−216.29	−110.08	−6955.42	3577.34	2.85	1.66	0.92	0.48	3.01
1-hexene	18	0.21	44.67	−194.20	−99.77	−5927.66	3051.21	2.52	1.43	0.76	0.35	3.23
1-hexyne	17	0.37	42.78	−183.52	−93.41	−5297.46	2733.66	2.35	1.31	0.67	0.28	2.36
1-iodopropane	13	1.77	36.63	−103.60	−54.32	−3133.58	1618.77	2.97	1.75	1.25	0.00	2.29
1-methoxy-2-propanol	19	0.61	36.28	−177.45	−83.29	−6935.30	3564.18	1.94	1.30	0.41	0.24	−1.05
1-methylnaphthalene	27	0.27	92.33	−300.92	−145.00	−12755.00	6518.57	3.82	2.80	2.01	1.39	3.70
1-methylbenz[*a*]anthracene	45	0.31	181.94	−505.47	−240.13	−29644.35	15061.83	3.82	2.80	2.01	1.39	6.64
1-methylcyclohexene	20	0.16	51.61	−219.16	−110.35	−7716.14	3955.45	3.05	2.30	1.52	1.00	3.27
1-nitropropane	18	4.21	34.44	−160.27	−71.86	−6224.60	3226.05	1.85	1.03	0.38	0.17	0.77
1-nonanol	31	1.41	67.60	−323.12	−161.50	−14190.30	7247.94	4.52	2.84	1.76	1.07	3.13
1-nonene	27	0.21	65.97	−292.48	−149.77	−11378.72	5820.54	4.02	2.49	1.51	0.89	5.05
1-nonyne	26	0.38	64.15	−281.81	−143.41	−10626.30	5442.06	3.85	2.37	1.42	0.83	4.26
1-octanol	28	1.40	60.52	−290.37	−144.84	−12143.75	6210.09	4.02	2.49	1.51	0.89	2.38
1-octene	24	0.21	58.88	−259.72	−133.10	−9454.92	4844.05	3.52	2.14	1.26	0.72	4.44
1-octyne	23	0.37	57.03	−249.04	−126.74	−8737.90	4483.22	3.35	2.01	1.17	0.65	3.61
1-pentanol	19	1.41	39.32	−192.12	−94.85	−6571.27	3380.08	2.52	1.43	0.76	0.36	0.60
1-pentene	15	0.21	37.55	−161.44	−83.11	−4357.51	2251.54	2.02	1.08	0.49	0.20	2.68
1-pentyne	14	0.36	35.61	−150.76	−76.75	−3782.51	1961.49	1.85	0.95	0.39	0.14	1.64
1-propanol	13	1.43	25.23	−126.58	−61.51	−3464.72	1797.55	1.52	0.72	0.22	0.00	1.22
1-propene	9	0.23	23.25	−95.99	−49.76	−1739.75	913.54	0.99	0.41	0.00	0.00	2.32
2,2,2-trifluoroethanol	19	3.48	18.86	−114.80	−42.50	−6900.91	3544.03	1.24	0.77	0.18	0.00	0.23
2,2′,3,3′,4,6′-hexachlorobiphenyl	47	1.61	158.75	−329.04	−150.11	−31382.31	15866.11	6.97	5.79	4.76	2.85	6.96
2,2′,3,4,4′,5′-hexachlorobiphenyl	47	1.02	162.76	−329.03	−150.08	−30646.85	15499.92	6.96	5.85	4.66	2.75	8.32
2,2,3-trimethylbutane	22	0.05	49.06	−238.38	−122.90	−9325.69	4766.26	2.94	3.52	1.73	0.00	4.36
2,2′,4,4′,6,6′-hexabromobiphenyl	47	0.00	174.04	−312.46	−146.38	−30735.96	15539.09	9.44	8.73	5.68	6.86	9.00
2,2′,4,4′,6,6′-hexachlorobiphenyl	47	0.03	160.83	−328.86	−150.02	−31394.08	15874.59	6.96	5.99	3.89	3.57	8.71
2,2′,4,5,5′-pentachloro-1,1′-biphenyl	44	0.70	150.83	−327.79	−150.82	−27448.42	13898.26	6.47	5.43	3.82	2.47	7.44
2,2,4-trimethylpentane	25	0.08	56.36	−271.05	−139.58	−11336.85	5786.07	3.42	4.16	1.02	1.22	4.70
2,2′,5,5′-tetrabromo-1,1′-biphenyl	41	0.05	148.79	−315.76	−149.13	−24260.21	12298.55	7.65	6.78	4.51	3.05	8.06
2,2,5,5-tetramethyl-3-hexyne	29	0.07	72.73	−315.48	−160.56	−13734.35	7006.62	3.75	4.75	0.81	0.38	3.69
2,2′,5,6′-tetrachloro-1,1′-biphenyl	41	0.79	136.64	−326.62	−151.64	−25136.11	12736.79	6.00	4.88	3.35	2.43	6.79
2,2′,5-trichloro-1,1′-biphenyl	38	0.98	131.27	−325.76	−152.05	−21860.50	11097.97	5.51	4.35	2.98	2.01	6.02
2,2,5-trimethylhexane	28	0.04	63.35	−303.91	−156.24	−13448.29	6856.55	3.92	4.49	1.47	0.72	5.05
2,2′,6,6′-tetrachloro-1,1′-biphenyl	41	0.04	134.71	−326.57	−151.67	−25683.35	13008.43	6.00	4.82	3.43	2.64	7.39
2,2′3,3′-tetrachloro-1,1′-biphenyl	41	1.76	135.97	−326.79	−151.72	−24853.53	12594.09	6.00	4.77	3.83	2.23	7.27
2,2-dimethy-3-hexyne	23	0.08	59.04	−250.29	−127.15	−9100.89	4661.94	3.06	2.91	0.64	0.35	3.03
2,2-dimethylbutane	19	0.07	42.24	−205.62	−106.32	−7134.06	3656.11	2.56	2.91	1.06	0.00	3.56
2,2-dimethylpentane	22	0.06	49.41	−238.39	−122.97	−9020.15	4613.67	3.06	3.31	1.00	0.75	4.36
2,2-dimethylpropane	16	0.00	35.06	−172.84	−89.78	−5259.24	2704.07	2.00	3.00	0.00	0.00	3.33
2,2-dimethylpropanoic acid	21	1.88	41.06	−201.61	−93.93	−8522.54	4376.69	2.18	2.56	0.64	0.00	0.67
2,3,4-trimethylpentane	25	0.04	56.57	−271.14	−139.48	−11414.93	5825.59	3.55	3.35	2.10	0.77	4.70
2,3,5,6-tetrachloro-1,1′-biphenyl	41	0.02	143.59	−326.63	−151.32	−24865.82	12602.48	6.00	4.78	3.75	2.46	7.32
2,3,6-trichloronaphthalene	33	0.76	115.27	−270.75	−125.65	−17284.93	8782.33	4.85	3.89	3.04	1.88	7.14
2,3-benzofluorene	40	0.35	154.10	−449.69	−213.36	−24086.39	12256.60	6.02	4.69	3.66	2.83	8.04
2,3-dichloro-2-methylbutane	22	0.25	51.87	−171.83	−86.97	−8843.43	4511.13	3.09	3.81	1.93	0.00	2.69
2,3-dichlorobutane	19	0.43	45.13	−139.02	−70.29	−6621.53	3385.27	2.80	2.73	1.52	0.00	2.70
2,3-dichloropropene	15	0.38	38.18	−96.17	−47.48	−4032.22	2068.93	2.07	1.45	0.74	0.00	1.71
2,3-dimethyl-1,3-butadiene	17	0.00	48.95	−183.81	−93.36	−5591.57	2879.12	1.96	1.56	0.73	0.00	2.40
2,3-dimethyl-2-butanol	22	1.50	45.63	−224.47	−111.39	−9343.04	4779.62	2.67	2.81	1.41	0.00	0.37
2,3-dimethylbutane	19	0.00	42.63	−205.57	−106.27	−7038.28	3608.55	2.64	2.49	1.33	0.00	3.58
2,3-dimethylnaphthalene	30	0.47	101.55	−334.41	−161.74	−15223.32	7766.67	4.23	3.28	2.36	1.47	4.72
2,3-dimethylpentane	22	0.04	49.71	−238.37	−122.87	−9014.85	4611.43	3.18	2.63	1.78	0.47	4.28
2,3-dimethylpyridine	21	1.91	60.56	−224.80	−107.69	−8285.31	4253.60	2.69	1.91	1.27	0.56	−0.38
2,4,4′-trichloro-1,1′-biphenyl	38	0.75	138.12	−325.82	−152.22	−21152.02	10744.59	5.51	4.41	2.88	2.00	6.21
2,4,6-tribromobiphenyl	38	0.54	136.86	−317.41	−150.64	−21545.38	10935.58	6.76	5.73	3.78	4.13	7.39
2,4,6-trichloro-1,1′-biphenyl	38	0.12	134.67	−325.71	−152.13	−21767.72	11051.16	5.51	4.36	2.88	2.46	6.14

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
2,4′-dichloro-1,1′-biphenyl	35	1.49	119.08	−324.61	−153.31	−18792.69	9558.63	5.03	3.83	2.61	1.69	5.28
2,4-dichloro-1,1′-biphenyl	35	0.83	124.28	−324.72	−153.11	−18792.87	9559.65	5.03	3.83	2.56	1.86	5.25
2,4-dimethyl-3-pentanone	24	2.44	52.44	−245.98	−121.77	−10477.81	5359.95	3.09	2.71	1.14	0.67	1.30
2,4-dimethylpentane	22	0.03	49.70	−238.25	−122.94	−8897.90	4552.71	3.13	3.02	0.94	0.94	4.26
2,4-dimethylpyridine	21	2.08	60.70	−224.74	−107.69	−8169.49	4195.95	2.68	1.97	1.05	0.69	−0.51
2,5-cyclohexadiene-1−4-dione	20	0.02	54.40	−189.19	−84.75	−7430.00	3830.47	2.23	1.47	0.82	0.44	0.99
2,5-dichloro-1,1′-biphenyl	35	0.28	122.12	−324.48	−153.11	−18841.87	9584.05	5.03	3.83	2.58	1.73	5.36
2,5-dimethylpyridine	21	1.87	61.30	−224.78	−107.72	−8148.58	4185.15	2.68	1.98	1.06	0.54	−0.40
2,6-dimethylnaphthalene	30	0.01	100.98	−334.30	−161.76	−15251.25	7780.64	4.24	3.24	2.36	1.62	4.89
2,6-dichloro-1,1′-biphenyl	35	1.03	120.66	−324.56	−152.96	−19287.16	9805.87	5.04	3.78	2.65	1.99	5.21
2,6-dichlorophenol	24	1.34	69.83	−188.58	−84.39	−10088.71	5150.55	3.10	2.38	1.55	0.89	1.79
2,6-diethylaniline	30	1.28	86.37	−326.21	−158.31	−15137.10	7720.18	4.15	2.70	2.05	1.28	2.35
2,6-dimethyl-4-heptanone	30	2.48	66.24	−311.38	−155.09	−14647.99	7473.26	4.04	3.73	1.27	0.99	1.73
2-bromofluorobenzene	21	2.14	58.54	−170.96	−78.43	−7883.55	4034.89	3.00	2.27	1.47	0.72	2.70
2-bromonaphthalene	27	1.35	98.33	−265.79	−126.70	−12286.54	6274.94	4.30	3.41	2.20	1.54	4.40
2-bromopropane	13	2.04	33.96	−104.97	−54.51	−3276.60	1691.63	2.29	2.84	0.00	0.00	1.59
2-bromotoluene	21	0.98	65.59	−196.98	−95.84	−7980.99	4086.67	3.31	2.60	1.96	0.84	2.23
2-butanol	16	1.50	32.06	−159.13	−78.11	−5138.42	2649.42	1.95	1.26	0.59	0.00	−0.39
2-butanone	15	2.70	31.69	−147.78	−71.78	−4472.51	2313.29	1.76	1.06	0.50	0.00	−0.49
2-chloro-1,1,1-trifluoroethane	19	2.17	22.18	−95.00	−36.07	−6654.20	3407.35	1.72	1.02	0.45	0.00	1.15
2-chloro-1-nitrobenzene	26	5.38	70.28	−219.58	−95.30	−11924.01	6109.94	2.89	2.04	1.30	0.76	2.55
2-chloro-2-methylbutane	19	1.66	43.58	−172.45	−88.36	−6885.63	3525.06	2.63	3.09	1.11	0.00	2.51
2-chloro-2-methylpropane	16	1.72	36.82	−139.68	−71.73	−5020.21	2577.78	2.07	3.20	0.00	0.00	2.20
2-chlorobutane	16	1.68	36.95	−139.48	−71.65	−4881.12	2508.40	2.35	1.93	0.87	0.00	1.96
2-chloroethanol	13	1.41	25.27	−93.20	−43.51	−3223.72	1668.55	1.62	0.79	0.25	0.00	−1.09
2-chloronaphthalene	27	1.07	95.54	−268.53	−127.33	−12374.03	6318.96	3.88	2.93	1.95	1.34	4.14
2-chloropentane	19	1.70	44.14	−172.34	−88.29	−6579.69	3372.76	2.85	2.31	0.90	0.62	2.63
2-chloropropane	13	1.66	29.82	−106.77	−55.01	−3321.99	1713.76	1.81	1.88	0.00	0.00	1.41
2-chloropyridine	18	2.44	54.27	−158.93	−73.27	−6084.49	3128.03	2.34	1.54	0.85	0.47	0.75
2-ethylpyridine	21	1.67	59.75	−224.05	−107.58	−8160.02	4190.23	2.83	1.67	1.09	0.61	−0.50
2-ethyl-1,3-hexandiol	31	1.37	63.68	−309.37	−149.71	−16006.43	8162.04	4.05	2.71	1.89	1.05	2.81
2-ethylbutyric acid	24	1.89	48.91	−234.34	−110.39	−10419.19	5340.77	2.95	1.85	1.33	0.64	0.81
2-ethylnaphthalene	30	0.39	100.87	−333.75	−161.68	−14900.95	7606.15	4.38	3.04	2.22	1.46	4.29
2-fluorochlorobenzene	21	2.04	56.66	−173.58	−79.00	−7950.10	4069.28	2.58	1.82	1.13	0.57	2.42
2-heptanone	24	2.66	52.80	−246.05	−121.77	−9674.50	4958.65	3.26	2.16	1.13	0.62	1.45
2-heptene	21	0.04	53.55	−227.55	−116.48	−7576.25	3889.58	3.03	1.71	0.96	0.49	3.82
2-heptyne	20	0.08	52.91	−217.81	−110.34	−6812.23	3504.35	2.81	1.53	0.83	0.39	2.64
2-hexanol	22	1.52	46.34	−224.72	−111.40	−8650.69	4435.01	2.95	1.99	0.98	0.50	1.42
2-hexanone	21	2.66	45.76	−213.30	−105.10	−7827.40	4020.39	2.76	1.81	0.88	0.43	0.79
2-hexene	18	0.05	45.65	−194.74	−99.85	−5858.94	3016.56	2.53	1.36	0.69	0.33	3.10
2-iodepropane	13	2.08	38.15	−104.03	−54.32	−3224.29	1665.21	2.60	3.46	0.00	0.00	2.09
2-methyl-1,3-butadiene	14	0.19	39.36	−150.62	−76.61	−3900.42	2019.60	1.55	1.05	0.35	0.00	2.03
2-methyl-1-butanol	19	1.35	39.11	−192.06	−94.79	−6908.40	3547.75	2.42	1.70	1.01	0.13	0.47
2-methyl-1-pentene	18	0.37	44.55	−194.85	−99.86	−6086.61	3130.32	2.41	1.71	0.68	0.43	3.03
2-methyl-2-butanol	19	1.56	38.80	−191.85	−94.71	−7173.19	3681.47	2.28	2.17	0.87	0.00	−0.10
2-methyl-2-butene	15	0.22	39.03	−162.69	−83.27	−4463.68	2303.59	1.87	1.37	0.58	0.00	2.56
2-methyl-2-pentanol	22	1.55	46.03	−224.62	−111.36	−9067.96	4643.41	2.78	2.56	0.86	0.61	0.50
2-methyl-3-hexyne	20	0.04	52.39	−217.57	−110.41	−6962.60	3578.88	2.75	1.83	0.57	0.31	2.59
2-methyl-3-pentanone	21	2.61	45.78	−213.31	−105.09	−8204.02	4208.38	2.71	1.97	0.99	0.41	0.81
2-methylaziridine	12	1.79	27.63	−123.17	−60.05	−3075.84	1602.08	1.63	1.31	0.41	0.45	−1.24
2-methyldecalin	32	0.01	75.24	−351.28	−177.18	−17597.94	8953.55	5.36	4.72	3.80	3.11	6.57
2-methylheptane	25	0.04	56.94	−271.15	−139.65	−10516.15	5376.87	3.77	2.89	1.39	0.80	5.08
2-methylhexane	22	0.04	49.86	−238.39	−122.99	−8623.12	4415.81	3.27	2.54	1.14	0.61	4.60
2-methylnaphthalene	27	0.31	93.38	−300.99	−145.03	−12583.86	6432.93	3.82	2.85	1.90	1.31	3.76
2-methylpentane	19	0.02	42.70	−205.60	−106.36	−6807.34	3493.26	2.77	2.18	0.87	0.58	3.74
2-methylpyridine	18	1.75	52.34	−191.29	−90.90	−6293.68	3243.41	2.27	1.47	0.81	0.45	1.03
2-nitropropane	18	4.16	34.20	−160.09	−71.92	−6442.80	3335.07	1.74	1.32	0.40	0.00	1.75
2-nitrotoluene	26	5.00	68.72	−251.94	−112.70	−12114.15	6217.94	2.88	2.01	1.28	0.75	0.72
2-nonanone	30	2.62	66.94	−311.56	−155.09	−13651.69	6976.43	4.26	2.87	1.63	0.98	2.57
2-octanone	27	2.63	59.90	−278.80	−138.43	−11618.62	5945.25	3.76	2.51	1.38	0.80	2.05
2-pentanol	19	1.53	39.28	−191.96	−94.74	−6847.26	3518.64	2.45	1.64	0.71	0.42	0.30
2-pentanone	18	2.68	38.71	−180.54	−88.44	−6084.54	3134.33	2.26	1.45	0.60	0.35	0.30
2-propanone	12	2.78	24.53	−115.02	−55.10	−2977.19	1551.32	1.20	0.91	0.00	0.00	−1.24
2-undecanone	36	2.62	81.06	−377.09	−188.42	−17949.79	9154.72	5.26	3.57	2.13	1.33	3.94
3,3′,4,4′-tetrachloro-1,1′-biphenyl	41	0.00	149.44	−327.05	−151.52	−23636.07	11989.61	5.99	4.88	3.65	1.95	8.73
3,3′,5,5′-tetrachloro-1,1′-biphenyl	41	0.00	145.45	−326.81	−151.40	−23632.09	11990.59	5.98	5.06	2.93	2.49	8.37
3,3′,5-trichloro-1,1′-biphenyl	38	1.05	132.63	−325.68	−152.19	−21367.87	10852.56	5.51	4.41	2.86	1.96	6.01
3,3-dimethyl-2-butanone	21	2.71	45.07	−213.23	−105.09	−8475.90	4344.85	2.45	2.81	1.06	0.00	0.72
3,3-dimethylpentane	22	0.06	49.35	−238.32	−122.92	−9160.53	4683.61	3.12	2.87	1.91	0.25	4.23
3,4-dichloro-1,1-biphenyl	35	1.57	124.94	−324.86	−153.21	−18574.54	9449.77	5.03	3.81	2.76	1.60	6.39
3,4-dichlorotoluene	24	1.65	73.65	−200.77	−95.63	−9887.48	5043.71	3.37	2.73	1.81	0.95	3.79

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
3,4-dimethylpyridine	21	2.41	60.45	−224.90	−107.62	−8274.18	4249.25	2.68	1.96	1.30	0.54	−0.36
3,5-dimethylpyridine	21	2.19	60.76	−224.91	−107.71	−8169.70	4196.64	2.67	2.04	1.01	0.71	−0.38
3-bromo-1-propene	12	1.69	34.18	−93.05	−47.92	−2650.08	1373.85	2.20	1.09	0.57	0.00	1.50
3-bromofluorobenzene	21	1.39	59.13	−170.76	−78.34	−7739.77	3964.87	2.99	2.35	1.28	0.83	2.67
3-chloro-1-propene	12	1.43	30.47	−95.28	−48.43	−2700.56	1399.47	1.62	0.75	0.33	0.00	1.28
3-chloroiodobenzene	21	0.79	74.00	−163.20	−77.92	−7576.15	3872.75	3.64	3.10	1.68	1.36	3.55
3-chlorophenol	21	1.94	60.70	−187.47	−85.22	−7976.31	4091.17	2.61	1.92	1.04	0.66	0.69
3-chloropropene	12	1.43	30.47	−95.28	−48.43	−2700.56	1399.47	1.62	0.75	0.33	0.00	1.10
3-ethylphenol	24	0.88	66.66	−252.51	−119.52	−10233.31	5243.68	3.11	2.02	1.32	0.78	−0.54
3-heptanol	25	1.60	53.35	−257.41	−128.05	−10696.44	5471.91	3.49	2.20	1.36	0.64	1.46
3-heptanone	24	2.58	52.98	−246.06	−121.79	−9783.61	5012.73	3.33	2.00	1.20	0.55	1.42
3-hexanone	21	2.59	45.90	−213.34	−105.12	−7916.14	4064.42	2.83	1.64	0.92	0.46	0.83
3-hexyne	17	0.03	45.42	−184.94	−93.70	−5154.19	2660.47	2.37	1.06	0.48	0.25	1.99
3-methyl-1-butene	15	0.16	37.08	−161.39	−83.15	−4486.05	2315.32	1.90	1.48	0.47	0.00	2.73
3-methyl-2-butanol	19	1.48	39.03	−191.86	−94.76	−7065.92	3626.78	2.32	1.98	0.96	0.00	0.16
3-methyl-2-butanone	18	2.60	38.25	−180.49	−88.49	−6299.99	3241.88	2.15	1.77	0.81	0.00	0.15
3-methyl-3-pentanol	22	1.56	45.90	−224.56	−111.34	−9220.15	4719.03	2.84	2.20	1.52	0.25	0.38
3-methylcholanthrene	50	0.76	199.38	−563.17	−267.61	−35304.24	17916.07	7.68	6.20	5.10	4.08	7.92
3-methylheptane	25	0.08	56.94	−271.17	−139.60	−10666.02	5451.86	3.81	2.66	1.75	0.76	5.16
3-methylhexane	22	0.08	49.88	−238.39	−122.94	−8733.98	4471.26	3.31	2.30	1.48	0.70	4.58
3-methylpentane	19	0.07	42.74	−205.64	−106.28	−6904.02	3541.78	2.81	1.92	1.39	0.29	3.84
3-methylpropylbenzene	27	0.27	77.40	−297.57	−147.47	−12178.28	6224.91	3.88	2.74	1.62	1.16	3.96
3-methylpyridine	18	2.07	52.21	−191.40	−90.89	−6291.20	3242.99	2.26	1.53	0.81	0.45	1.03
3-methylthiophene	16	0.95	50.77	−151.76	−74.63	−4860.80	2506.15	2.82	2.08	1.46	1.00	2.39
3-pentanol	19	1.61	39.19	−191.90	−94.73	−6928.49	3559.25	2.49	1.47	0.94	0.29	0.23
3-pentanone	18	2.61	38.86	−180.57	−88.46	−6152.44	3167.54	2.33	1.25	0.79	0.25	0.28
4,4′-bromo-1,1′-biphenyl	35	0.00	132.56	−319.24	−151.89	−18077.69	9202.67	5.85	4.84	3.08	1.81	7.74
4,4′-dichloro-1,1′-biphenyl	35	0.01	128.40	−324.78	−153.14	−18257.86	9293.24	5.03	3.88	2.52	1.54	6.56
4,4′-dimethyl-1,1′-biphenyl	35	0.01	122.96	−389.65	−188.54	−18673.25	9518.33	4.89	3.73	2.44	1.49	6.02
4-bromo-1,1′biphenyl	32	1.37	117.01	−320.96	−153.40	−15963.17	8140.40	4.96	3.79	2.48	1.54	5.55
4-bromo-1-butene	15	1.63	41.60	−125.88	−64.59	−4055.23	2092.09	2.70	1.56	0.77	0.40	2.25
4-bromoiodobenzene	21	0.46	77.38	−160.43	−77.29	−7476.86	3822.68	4.05	3.58	2.04	1.05	4.56
4-bromophenol	21	1.52	63.15	−184.64	−84.57	−7872.75	4038.80	3.03	2.39	1.35	0.70	1.09
4-bromotoluene	21	1.49	66.60	−196.90	−95.83	−7827.71	4010.36	3.30	2.71	1.54	0.80	3.19
4-chloro-2-nitrophenol	29	3.64	76.39	−241.43	−100.57	−14365.34	7358.03	3.03	2.25	1.31	0.76	3.09
4-chloroaniline	21	1.97	66.59	−195.12	−90.68	−7928.32	4063.08	2.68	1.99	1.12	0.59	1.51
4-chloroiodobenzene	21	0.22	74.61	−163.18	−77.91	−7555.57	3862.65	3.64	3.10	1.76	0.91	4.03
4-chlorophenol	21	1.39	60.76	−187.37	−85.19	−7950.97	4078.36	2.61	1.91	1.08	0.57	0.73
4-ethenylcyclohexene	22	0.14	59.27	−239.90	−120.33	−9083.29	4651.27	3.21	2.28	1.64	1.03	3.34
4-fluoro-10-methylbenz[*a*]anthracene	48	2.13	186.22	−513.03	−239.60	−32163.66	16335.45	6.74	5.32	3.94	2.92	7.72
4-fluoroiodobenzene	21	0.95	64.83	−169.41	−78.08	−7654.34	3920.50	3.26	2.66	1.51	0.78	3.13
4-heptanone	24	2.58	52.99	−246.05	−121.77	−9801.15	5021.86	3.33	2.04	1.06	0.68	1.55
4-methylpyridine	18	2.27	51.97	−191.43	−90.87	−6289.85	3242.69	2.26	1.52	0.85	0.43	0.10
4-methyl-1,1′-biphenyl	32	0.33	112.37	−356.16	−171.73	−16259.06	8297.38	4.48	3.23	2.16	1.38	4.62
4-methyl-1-pentene	18	0.25	44.26	−194.14	−99.77	−6129.35	3151.91	2.38	1.92	0.64	0.33	3.24
4-methyl-2-methoxybenzene	24	1.04	68.80	−251.45	−119.54	−10251.70	5250.10	2.93	2.02	1.26	0.67	1.82
4-methyl-2-pentanone	21	2.62	45.28	−213.20	−105.09	−8122.99	4167.72	2.62	2.30	0.70	0.57	0.72
4-methyl-2-pentylacetate	30	1.77	64.75	−298.47	−143.67	−15059.39	7685.06	3.69	3.05	1.17	1.00	2.05
4-methyloctane	28	0.08	64.08	−303.91	−156.26	−12732.37	6499.53	4.31	3.04	1.83	1.19	6.05
4-methylphenol	21	1.18	59.63	−219.78	−102.90	−8161.48	4193.11	2.54	1.84	1.03	0.54	0.70
4-tertbutylphenol	30	1.20	80.56	−317.77	−152.89	−15192.42	7751.20	3.79	3.80	1.73	0.96	−4.31
5,6-dimethylchrysene	48	0.51	186.32	−538.96	−256.80	−33475.25	16990.90	7.07	5.50	4.49	3.33	7.01
5-fluoro-7-methylbenz[*a*]anthracene	48	1.59	184.58	−512.96	−239.48	−32960.18	16733.94	6.75	5.22	4.10	3.08	7.72
5-methyl-2-hexanone	24	2.67	52.52	−246.03	−121.75	−9991.87	5117.20	3.12	2.63	1.07	0.49	1.33
5-methylbenzo[*a*]pyrene	49	0.28	206.82	−553.25	−261.16	−34721.82	17622.95	7.39	5.91	4.69	3.75	8.52
5-methylchrysene	45	0.34	177.64	−506.09	−240.23	−29689.08	15084.13	6.64	5.12	3.97	3.04	6.59
6-chloro-10-methylbenz[*a*]anthracene	48	1.19	196.68	−506.87	−239.43	−31924.39	16207.09	7.11	5.79	4.15	3.06	8.70
6-fluoro-7-methylbenz[*a*]anthracene	48	1.28	184.25	−512.92	−239.45	−33183.34	16845.43	6.75	5.22	4.08	3.13	7.72
6-methylbenzo[*a*]pyrene	49	0.39	208.62	−553.32	−261.08	−34873.01	17698.51	7.39	5.85	4.80	3.81	8.53
6-methylchrysene	45	0.28	177.96	−505.62	−240.10	−29854.50	15167.01	6.64	5.12	3.95	3.09	6.57
7,12-dimethyl-1,2-benzanthracene	48	0.29	194.93	−539.25	−257.12	−31985.33	16245.54	7.04	5.71	4.10	3.05	7.02
7-ethylbenz[*a*]anthracene	48	0.32	189.63	−538.43	−256.79	−33094.84	16801.42	7.20	5.31	4.24	3.29	6.80
7-methylbenz[*a*]anthracene	45	0.33	182.41	−505.75	−240.20	−29631.15	15054.85	6.64	5.10	4.04	3.04	7.35
9,10-dimethylanthracene	39	0.02	149.43	−436.47	−209.21	−23788.03	12097.80	5.65	4.34	3.47	2.56	6.57
9-chloro-7-methylbenz[*a*]anthracene	48	1.06	197.08	−506.74	−239.29	−32434.07	16462.29	7.12	5.68	4.33	3.23	7.44
9-methylbenz[*a*]anthracene	45	0.37	184.29	−505.80	−240.32	−29028.19	14753.20	6.63	5.21	3.86	2.87	6.56
a,a,a-trichlorotoluene	27	1.84	81.14	−197.52	−93.79	−12187.50	6201.23	3.86	4.14	1.75	1.00	2.51
a,a,a-trifluorotoluene	27	3.11	54.18	−219.09	−94.96	−12629.82	6452.61	2.73	1.86	1.10	0.63	2.51
acenaphthene	29	0.56	98.45	−324.75	−155.50	−14680.42	7492.23	4.45	3.43	2.74	2.19	4.63
acenaphthylene	28	0.30	101.78	−313.19	−148.85	−13763.02	7029.83	4.15	3.13	2.41	1.86	3.96
acetaldehyde	9	2.46	17.30	−81.95	−38.35	−1692.96	893.44	0.81	0.24	0.00	0.00	−1.36

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
acetamide	12	3.31	24.71	−110.93	−49.51	−2991.34	1561.49	0.99	0.61	0.00	0.00	−1.08
acetic acid	12	1.83	20.52	−103.59	−43.86	−3005.22	1573.86	0.93	0.52	0.00	0.00	1.22
acetonitrile	8	3.21	18.63	−77.35	−37.08	−1292.85	688.59	0.72	0.22	0.00	0.00	−1.39
acetophenone	23	2.79	65.37	−238.60	−112.72	−9485.20	4866.67	2.86	1.92	1.18	0.67	1.29
a-chlorotoluene	21	1.33	61.49	−198.02	−96.02	−7901.70	4049.49	3.06	1.89	1.31	0.74	2.43
allyl alcohol (2-propen-1-ol)	12	1.48	26.12	−115.14	−54.92	−2930.84	1527.95	1.13	0.47	0.13	0.00	−0.74
amyl propionate (pentyl)	30	1.83	65.25	−298.61	−143.84	−14046.16	7179.02	3.96	2.28	1.30	0.64	2.25
aniline	18	1.30	55.86	−194.05	−91.57	−6286.64	3237.94	2.20	1.41	0.80	0.45	0.41
anthracene	33	0.00	131.41	−369.69	−175.83	−17604.92	8978.31	4.81	3.55	2.61	1.88	6.35
benz[*a*]anthracene	42	0.02	173.21	−472.36	−223.51	−26202.95	13326.77	6.22	4.71	3.61	2.70	7.21
benz[*b*]anthracene	42	0.01	187.16	−471.85	−223.41	−25818.94	13134.36	6.21	4.75	3.57	2.63	8.60
benzene	15	0.00	45.60	−165.18	−80.55	−4580.33	2368.49	2.00	1.15	0.67	0.38	1.64
benzo[*a*]pyrene	46	0.02	198.33	−519.91	−244.40	−31242.58	15869.32	6.97	5.45	4.37	3.49	7.80
benzo[*b*]fluoranthene	46	0.14	193.18	−518.95	−244.40	−30577.04	15535.97	6.98	5.42	4.39	3.55	8.23
benzo[*e*]pyrene	46	0.02	184.50	−520.22	−244.39	−31654.85	16075.90	6.98	5.42	4.39	3.56	7.60
benzo[*ghi*]perylene	50	0.03	207.39	−567.71	−265.35	−36691.75	18617.00	7.72	6.20	5.14	4.24	9.03
benzo[*j*]fluoranthene	46	0.36	187.47	−518.90	−244.23	−31070.03	15782.21	6.98	5.42	4.39	3.56	8.00
benzo[*k*]fluoranthene	46	0.27	188.39	−519.09	−244.37	−30931.67	15713.59	6.97	5.45	4.37	3.48	8.49
benzoaldehyde	20	2.69	58.35	−205.73	−96.02	−7398.80	3808.07	2.44	1.53	0.94	0.53	3.24
benzoic acid	23	2.25	61.38	−227.26	−101.54	−9543.82	4904.94	2.59	1.67	1.02	0.58	1.81
benzonitrile	19	3.61	61.85	−201.20	−94.87	−6653.81	3433.15	2.38	1.48	0.90	0.51	1.01
benzyl acetate	29	1.85	77.08	−291.63	−134.85	−13847.26	7082.19	3.46	2.27	1.28	0.76	1.69
biphenyl	29	0.00	102.00	−322.71	−154.92	−13942.30	7125.09	4.07	2.73	1.88	1.26	4.31
bromochloromethane	10	1.51	26.20	−38.94	−20.12	−1678.46	864.12	2.19	1.57	0.00	0.00	0.89
bromoethane	10	1.85	25.99	−71.81	−37.82	−1899.20	986.70	2.09	1.39	0.00	0.00	1.09
bromomethane	7	1.55	17.07	−38.69	−21.19	−834.66	437.83	1.96	0.00	0.00	0.00	0.80
butanol	16	1.42	32.28	−159.33	−78.18	−4944.52	2552.02	2.02	1.08	0.51	0.16	0.07
butyl acetate	24	1.91	50.81	−233.03	−110.50	−9887.22	5070.10	2.90	1.69	0.80	0.36	1.27
butylamine	16	1.34	34.76	−167.08	−83.60	−4937.80	2546.32	2.12	1.14	0.56	0.20	−1.14
butylbenzene	27	0.34	76.07	−296.88	−147.33	−12029.18	109.69	3.97	2.59	1.66	1.03	4.06
butyraldehyde	15	2.54	31.72	−147.46	−71.68	−4324.38	2239.94	1.85	0.96	0.41	0.12	0.01
chlorobenzene	18	0.93	55.14	−166.21	−79.66	−6056.66	3110.78	2.48	1.73	0.98	0.56	2.44
chlorodifluoromethane	13	1.52	17.11	−54.63	−20.02	−3057.01	1576.69	1.09	0.58	0.00	0.00	1.47
chloroethane	10	1.55	22.24	−73.96	−38.34	−1941.24	1007.94	1.51	0.80	0.00	0.00	1.06
chloroethene	9	0.91	22.96	−63.38	−32.02	−1514.59	790.14	1.06	0.46	0.00	0.00	1.30
chloromethane	7	1.38	14.02	−41.20	−21.72	−875.79	459.44	1.13	0.00	0.00	0.00	0.98
chloropentafluoroethane	25	1.10	25.29	−109.26	−34.72	−11277.80	5750.41	1.76	1.47	0.54	0.00	2.79
cholanthrene	47	0.59	188.81	−529.64	−250.78	−32070.18	16285.36	7.27	5.75	4.72	3.84	7.85
chrysene	42	0.00	168.13	−472.61	−223.50	−26409.99	13430.53	6.23	4.67	3.65	2.77	8.06
cis-2-pentene	15	0.04	38.37	−161.97	−83.18	−4303.97	2224.42	2.03	0.98	0.47	0.24	2.54
cis-1,3-dichloro-1-propene	15	1.29	39.90	−96.13	−47.45	−3841.73	1974.16	2.20	1.08	0.53	0.30	2.05
cis-1,2-dichloroethene	12	0.00	32.74	−64.30	−31.08	−2473.14	1272.79	1.64	0.75	0.43	0.00	1.44
cis-1,2-dimethylcyclohexane	24	0.01	55.26	−262.05	−133.40	−10942.09	5585.29	3.80	3.24	2.54	1.50	4.44
coronene	54	0.00	227.15	−615.31	−286.34	−41996.34	21291.99	8.46	6.98	5.89	4.89	9.33
cumene	24	0.25	68.10	−263.95	−130.69	−10318.83	5280.48	3.35	2.57	1.47	0.84	3.29
cycloheptane	21	0.02	48.56	−229.02	−116.69	−8602.12	4401.11	3.50	2.47	1.75	1.24	3.51
cycloheptene	20	0.21	49.67	−218.30	−110.20	−7826.37	4010.79	3.15	2.11	1.41	0.94	3.16
cyclohexane	18	0.00	41.86	−196.54	−100.10	−6584.58	3378.54	3.00	2.12	1.50	1.06	3.18
cyclohexanol	21	1.50	45.08	−215.58	−105.19	−8592.48	4402.52	3.07	2.29	1.57	1.08	0.43
cyclohexanone	20	2.81	44.80	−204.18	−98.83	−7799.65	4004.93	2.91	2.10	1.41	0.95	0.63
cyclohexene	17	0.17	43.64	−185.81	−93.60	−5813.51	2990.23	2.65	1.76	1.16	0.76	2.59
cyclooctane	24	0.00	55.09	−261.52	−133.26	−10769.16	5498.84	4.00	2.83	2.00	1.41	4.15
cyclopentane	15	0.01	33.86	−163.30	−83.45	−4739.86	2441.68	2.50	1.77	1.25	0.88	2.59
cyclopentanone	17	2.71	37.02	−171.28	−82.18	−5792.17	2986.36	2.41	1.75	1.16	0.77	0.96
cyclopentene	14	0.15	35.87	−152.24	−76.79	−4132.44	2135.36	2.15	1.40	0.91	0.59	2.10
cyclopropane	9	0.00	20.85	−97.70	−49.72	−1826.20	960.23	1.50	1.06	0.00	0.53	2.04
decachlorobiphenyl	59	0.05	220.38	−331.77	−146.63	−46168.65	23270.07	8.91	7.66	7.21	4.30	11.62
decalin (decahydronaphthalene)	29	0.00	68.41	−318.52	−160.50	−15051.97	7666.25	4.97	4.09	3.47	2.86	5.19
decan-2-one	33	2.62	74.00	−344.34	−171.76	−15765.05	8047.74	4.76	3.22	1.88	1.15	3.30
decanal	33	2.60	74.34	−343.89	−171.65	−15452.55	7893.02	4.85	3.08	1.93	1.18	3.41
decane	31	0.00	71.26	−336.62	−173.02	−14170.01	7233.20	4.91	3.12	1.96	1.21	6.44
di(*n*-butyl) ether	28	1.12	62.05	−289.16	−144.82	−12330.85	6300.79	3.99	2.32	1.28	0.59	2.64
di(*n*-propyl) ether	22	1.11	47.83	−223.62	−111.48	−8469.24	4340.63	2.99	1.61	0.70	0.39	1.32
dibromodifluoromethane	16	0.54	33.63	−52.03	−18.40	−4467.60	2284.51	2.34	3.48	0.00	0.00	1.89
dibromomethane	10	1.45	27.97	−36.34	−19.60	−1650.54	848.21	2.77	2.72	0.00	0.00	1.16
dichlorodifluoromethane	16	0.80	26.94	−54.18	−18.89	−4564.44	2334.43	1.51	1.57	0.00	0.00	2.64
dichlorofluoromethane	13	1.25	25.13	−47.09	−19.60	−2944.88	1510.94	1.53	1.23	0.00	0.00	1.03
dichloromethane	10	1.36	22.36	−40.85	−20.50	−1709.59	880.63	1.60	0.91	0.00	0.00	0.82
diethyl sulfide	16	1.94	43.08	−147.75	−76.44	−4559.23	2351.28	3.14	2.34	1.22	0.61	1.25
diethylamine	16	1.18	35.68	−166.59	−83.49	−4958.82	2556.89	2.12	0.96	0.50	0.25	−1.14
diethyl ether	16	1.15	33.58	−158.12	−78.15	−5047.89	2600.43	1.99	0.78	0.41	0.20	0.09

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
difluoromethane	10	1.81	8.19	−55.32	−21.30	−1880.31	982.52	0.53	0.10	0.00	0.00	1.81
diisobutylamine	28	1.08	63.39	−297.58	−150.16	−12917.35	6593.38	3.83	3.45	1.11	0.66	1.77
diisopropylamine	22	1.22	49.42	−231.91	−116.69	−8964.03	4588.43	2.89	2.48	0.67	0.67	−0.04
diisopropyl ether	22	1.31	47.40	−223.27	−111.34	−9093.69	4652.50	2.78	2.23	0.54	0.54	1.06
dimethylamine	10	1.27	21.09	−101.36	−50.20	−2183.97	1140.92	1.00	0.50	0.00	0.00	−1.56
dimethyl ether	10	1.26	18.64	−93.19	−44.89	−2213.09	1154.98	0.82	0.41	0.00	0.00	−0.55
dimethyl sulfide	10	1.96	26.75	−82.45	−43.13	−1894.71	989.59	2.44	1.22	0.00	0.00	0.74
di-*n*-butylamine	28	1.15	64.01	−297.61	−150.15	−12176.31	6224.32	4.12	2.46	1.41	0.73	1.57
di-*n*-propylamine	22	1.16	49.82	−232.05	−116.81	−8346.47	4280.07	3.12	1.75	0.85	0.48	0.24
diphenylmethane	32	0.17	102.60	−355.24	−171.66	−16406.95	8371.20	4.53	3.15	2.09	1.43	4.62
diphenylmethane	35	0.23	118.51	−387.95	−188.27	−19428.12	9896.15	4.98	3.54	2.60	1.68	4.08
dodecane	37	0.00	85.50	−401.98	−206.33	−18474.35	9414.30	5.91	3.83	2.46	1.56	7.67
ethane	7	0.00	14.95	−74.57	−39.70	−1121.68	592.68	1.00	0.00	0.00	0.00	2.70
ethanol	10	1.45	18.20	−93.85	−44.85	−2161.56	1131.23	1.02	0.32	0.00	0.00	−1.34
ethene	6	0.00	15.01	−62.63	−33.05	−791.09	424.61	0.50	0.00	0.00	0.00	2.33
ethyl acetate	18	1.91	36.44	−167.65	−77.18	−6255.54	3224.64	1.90	0.92	0.35	0.20	0.04
ethyl benzene	21	0.33	61.53	−231.36	−114.00	−8127.75	4170.95	2.97	1.84	1.25	0.71	1.36
ethyl benzoate	29	2.09	78.21	−291.31	−134.86	−14003.96	7159.94	3.56	2.09	1.28	0.79	2.32
ethyl butyrate	24	1.81	50.92	−233.06	−110.49	−10014.64	5133.88	2.96	1.56	0.76	0.44	1.37
ethyl formate	15	3.91	29.72	−134.84	−60.61	−4462.05	2312.42	1.47	0.55	0.23	0.07	−0.08
ethyl mercaptan	9	2.33	24.64	−74.67	−38.10	−1726.83	902.48	1.65	0.95	0.00	0.00	0.73
ethylallylamine	18	1.22	44.22	−187.94	−93.57	−5935.52	3057.55	2.23	1.10	0.54	0.25	1.35
ethylcyclohexane	24	0.05	55.73	−262.14	−133.35	−10780.60	5505.13	3.93	2.91	2.30	1.60	4.25
fluorene	31	0.37	108.81	−347.34	−165.66	−16027.27	8178.05	4.61	3.49	2.71	2.08	5.00
fluoromethane	7	1.44	7.92	−48.31	−22.13	−932.80	494.86	0.38	0.00	0.00	0.00	1.28
formaldehyde	6	2.16	9.92	−49.13	−21.73	−723.52	391.92	0.29	0.00	0.00	0.00	−1.12
formic acid	9	3.94	13.66	−70.85	−27.35	−1703.52	904.87	0.49	0.11	0.00	0.00	1.34
furan	13	0.22	32.09	−129.20	−59.25	−3634.00	1887.76	1.47	0.79	0.43	0.23	0.83
heptanal	24	2.59	53.03	−245.70	−121.67	−9424.75	4834.97	3.35	2.02	1.18	0.65	1.96
heptanoic acid	27	1.78	55.94	−267.31	−127.17	−11677.97	5984.86	3.49	2.20	1.24	0.70	1.66
hexabromobenzene	33	0.01	116.95	−155.20	−72.15	−16380.68	8275.51	7.38	6.63	9.09	4.54	9.74
hexachloro-1,3-butadiene	29	0.00	98.24	−122.80	−54.64	−13165.02	6653.39	4.15	3.80	2.51	0.92	4.92
hexachloro-1-propene	27	1.07	78.94	−97.73	−43.69	−11814.40	5970.90	3.90	4.67	2.31	0.96	4.16
hexachlorobenzene	33	0.00	108.70	−172.13	−75.76	−16969.77	8581.96	4.90	4.15	4.00	2.00	6.78
hexachlorocyclopentadiene	31	0.78	97.73	−147.02	−65.12	−15346.50	7757.37	4.65	4.66	4.25	2.13	5.18
hexachloroethane	25	0.01	64.76	−72.83	−32.72	−10564.99	5337.56	3.65	5.54	2.88	0.00	4.49
hexadecane	49	0.00	113.81	−533.07	−273.00	−27880.82	14175.93	7.91	5.24	3.46	2.27	8.40
hexamethylbenzene	33	0.01	96.21	−365.62	−181.00	−18086.43	9204.26	4.50	3.75	3.38	1.69	5.23
hexanal	21	2.58	45.94	−212.97	−105.01	−7605.03	3910.25	2.85	1.66	0.93	0.48	1.22
hexane	19	0.00	43.01	−205.61	−106.37	−6562.01	3370.96	2.91	1.71	0.96	0.50	3.96
hexanoic acid	24	1.80	48.86	−234.55	−110.50	−9728.78	4995.54	2.99	1.85	0.99	0.53	1.05
hydroquinone	21	2.22	56.33	−207.30	−91.63	−8205.24	4220.83	2.27	1.52	0.85	0.45	0.17
indane	23	0.45	67.86	−255.36	−124.56	−9796.59	5016.14	3.53	2.62	2.01	1.51	3.13
iodobenzene	18	0.79	63.21	−162.17	−78.79	−5922.36	3041.82	3.16	2.52	1.44	0.82	2.95
iodoethane	10	1.83	29.61	−70.91	−37.65	−1860.90	967.23	2.47	1.77	0.00	0.00	1.60
iodomethane	7	1.44	19.95	−37.78	−21.02	−811.08	425.62	2.50	0.00	0.00	0.00	1.00
isoamyl acetate	27	1.87	57.73	−265.90	−127.13	−12245.51	6263.96	3.26	2.52	1.00	0.44	1.81
isoamyl formate	24	3.94	51.17	−233.07	−110.56	−9955.84	5104.77	2.82	2.15	0.88	0.35	1.52
isobutane	13	0.01	28.61	−140.08	−73.06	−3562.33	1841.76	1.73	1.73	0.00	0.00	3.08
isobutanol	16	1.39	31.98	−159.27	−78.19	−5136.07	2646.96	1.88	1.58	0.37	0.00	−0.01
isobutene	12	0.36	30.72	−129.36	−66.55	−3003.53	1559.59	1.35	1.21	0.00	0.00	2.33
isobutyl acetate	24	1.94	50.42	−233.08	−110.52	−10225.32	5238.76	2.76	2.20	0.62	0.28	1.24
isobutylamine	16	1.36	34.54	−166.96	−83.58	−5125.92	2640.27	2.03	1.39	0.64	0.00	−1.14
isobutyl formate	21	3.84	43.76	−200.33	−93.96	−8083.63	4152.96	2.32	1.82	0.50	0.23	1.00
isobutyl methyl ether	19	1.15	39.98	−191.08	−94.86	−6905.12	3543.60	2.26	1.85	0.50	0.33	0.90
isobutylbenzene	27	0.34	76.02	−296.87	−147.26	−12441.96	6357.30	3.83	3.08	1.49	1.09	4.12
isobutyraldehyde	15	2.45	31.39	−147.47	−71.74	−4478.45	2316.44	1.72	1.38	0.27	0.00	−0.09
isopentyl acetate	27	1.88	57.73	−265.92	−127.13	−12249.67	6266.04	3.26	2.52	1.00	0.44	1.06
isophorone	28	3.74	69.91	−291.56	−142.28	−13720.86	7006.44	3.70	3.97	1.82	1.67	0.55
isopropyl alcohol	13	1.52	25.03	−126.44	−61.49	−3563.95	1847.51	1.41	1.09	0.00	0.00	−1.22
isopropyl formate	18	4.02	36.70	−167.41	−77.24	−6286.02	3240.00	1.86	1.28	0.33	0.11	0.63
isopropylamine	13	1.32	27.54	−134.18	−66.94	−3555.53	1839.99	1.49	1.24	0.00	0.00	−1.23
m-bromotoluene	21	1.36	66.45	−196.89	−95.83	−7849.26	4021.14	3.30	2.71	1.47	1.12	3.52
m-chlorotoluene	21	1.12	63.78	−199.63	−96.45	−7930.45	4062.18	2.89	2.23	1.21	0.85	3.52
m-cresol (3-hydroxytoluene)	21	0.96	59.21	−219.81	−102.88	−8186.34	4205.70	2.54	1.84	1.00	0.63	0.68
methanol	7	1.49	10.94	−61.35	−28.21	−1074.40	573.48	0.45	0.00	0.00	0.00	−1.49
methycyclopentane	18	0.04	40.70	−196.07	−100.13	−6524.85	3348.33	2.89	2.39	1.64	1.13	3.30
methyl acetate	15	1.83	28.71	−135.15	−60.55	−4592.41	2378.42	1.32	0.70	0.29	0.00	−0.52
methylamine	7	1.40	13.55	−68.97	−33.62	−1101.99	584.88	0.58	0.00	0.00	0.00	−1.54
methyl butyl ether	19	1.18	40.32	−191.16	−94.85	−6655.98	3419.85	2.40	1.35	0.70	0.29	0.99
methyl butyrate	21	1.74	43.14	−200.59	−93.87	−8059.17	4141.61	2.38	1.33	0.68	0.31	0.83

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
methyl formate	12	3.86	21.98	−102.35	−43.97	−3004.64	1568.87	0.88	0.33	0.10	0.00	−0.58
methyl hexanoate	27	1.74	57.35	−266.11	−127.20	−11895.05	6089.18	3.38	2.03	1.21	0.62	1.99
methyl isopropyl ether	16	1.25	33.00	−158.20	−78.15	−5215.88	2684.94	1.80	1.28	0.47	0.00	0.06
methyl mercaptan	6	2.23	18.55	−42.03	−21.46	−739.78	393.30	1.34	0.00	0.00	0.00	0.49
methyl propanoate	18	1.76	36.00	−167.86	−77.22	−6288.21	3240.99	1.88	0.93	0.52	0.14	0.15
methyl propyl ether	13	1.21	26.13	−125.66	−61.52	−3533.36	1829.17	1.40	0.58	0.29	0.00	0.39
methylcyclohexane	21	0.02	48.58	−229.34	−116.77	−8646.66	4423.58	3.39	2.74	1.89	1.31	3.85
m-xylene	21	0.26	62.57	−232.07	−114.16	−8141.18	4176.99	2.82	2.16	1.17	0.81	2.80
n-amyl acetate	27	1.94	57.94	−265.90	−127.17	−11844.98	6063.90	3.40	2.05	1.05	0.57	1.88
naphthalene	24	0.00	83.80	−267.54	−128.23	−10411.15	5332.60	3.40	2.35	1.66	1.13	3.60
n-butane	13	0.00	28.95	−140.09	−73.04	−3470.54	1796.13	1.91	1.00	0.50	0.00	2.98
n-butyl propionate	27	1.83	58.11	−265.90	−127.18	−11991.99	6137.25	3.46	1.93	1.05	0.43	1.94
n-butyl-acetate	24	1.92	50.77	−233.15	−110.51	−9892.56	5072.93	2.90	1.69	0.80	0.36	1.27
n-butylformate	21	3.91	44.11	−200.37	−93.94	−7783.86	4003.84	2.47	1.32	0.68	0.28	1.13
n-ethylaniline	24	1.27	72.04	−258.93	−124.70	−10097.63	5173.72	3.22	1.90	1.15	0.76	1.70
n-heptane	22	0.00	50.06	−238.35	−123.03	−8310.18	4259.57	3.41	2.06	1.21	0.68	4.53
n-hexyl acetate	30	1.92	64.99	−298.67	−143.84	−13891.64	7101.95	3.90	2.40	1.30	0.74	2.45
n-hexylamine	22	1.33	48.88	−232.53	−116.92	−8307.44	4260.31	3.12	1.85	1.06	0.57	0.69
n-hexylbenzene	33	0.35	90.40	−362.41	−180.66	−16246.78	8289.31	4.97	3.30	2.16	1.40	5.21
nitrobenzene	23	5.24	59.94	−218.47	−96.06	−9665.70	4978.59	2.46	1.56	0.94	0.54	1.88
nitromethane	12	3.99	19.98	−95.01	−38.57	−3058.49	1611.26	0.76	0.42	0.00	0.00	0.26
N-methylaniline	21	1.28	64.04	−226.36	−108.11	−8168.76	4194.49	2.66	1.62	1.06	0.60	1.28
N,*N*-dimethyl formamide	15	3.46	33.34	−143.13	−66.02	−4534.96	2348.00	1.39	1.07	0.21	0.00	−1.14
N,*N*-dimethylaniline	24	1.21	72.30	−258.43	−124.53	−10375.90	5312.41	3.03	2.23	1.27	0.73	1.79
nonanal	30	2.59	67.20	−311.29	−155.02	−13384.50	6844.14	4.35	2.72	1.68	1.01	2.58
nonane	28	0.00	64.18	−303.87	−156.36	−12126.10	6196.67	4.41	2.77	1.71	1.03	5.88
nonanoic acid	33	1.76	70.07	−332.81	−160.50	−15829.05	8089.70	4.49	2.91	1.74	1.06	2.87
n-pentyl acetate	27	1.92	57.89	−265.90	−127.17	−11850.61	6066.71	3.40	2.05	1.05	0.57	1.88
n-pentylamine	19	1.33	41.80	−199.83	−100.27	−6560.16	3372.12	2.62	1.50	0.81	0.39	−1.06
n-pentylbenzenz	30	0.34	83.26	−329.67	−164.00	−14099.07	7200.82	4.47	2.94	1.91	1.22	4.59
n-pentylcyclopentane	30	0.05	69.11	−327.02	−166.72	−14498.23	7393.11	4.93	3.65	2.69	1.98	6.08
n-propyl acetate	21	1.91	43.60	−200.39	−93.85	−8026.78	4125.20	2.40	1.34	0.51	0.25	0.73
n-propylamine	13	1.34	27.71	−134.32	−66.94	−3465.01	1795.37	1.62	0.79	0.29	0.00	−1.23
n-propylbenzene	24	0.34	68.88	−264.11	−130.66	−10026.86	5135.31	3.47	2.24	1.38	0.93	3.36
n-propylcyclopentane	24	0.06	54.97	−261.62	−133.39	−10391.83	5310.94	3.93	2.94	2.17	1.71	7.74
o-cresol (2-hydroxytoluene)	21	1.40	58.99	−219.78	−102.83	−8304.57	4265.44	2.55	1.79	1.12	0.56	0.62
octachloronaphthalene	48	0.00	175.20	−275.79	−121.68	−32630.54	16470.07	7.28	6.28	5.88	3.63	9.70
octanal	27	2.59	60.13	−278.52	−138.35	−11364.27	5819.32	3.85	2.37	1.43	0.83	2.36
octane	25	0.00	57.13	−271.10	−139.69	−10167.75	5202.91	3.91	2.41	1.46	0.85	5.11
octanoic acid	30	1.78	63.02	−300.06	−143.83	−13704.94	7013.04	3.99	2.56	1.49	0.88	2.26
p-chloronitrobenzene	26	4.59	71.82	−219.65	−95.24	−11558.36	5930.33	2.89	2.09	1.23	0.66	2.54
pentachlorobenzene	30	0.43	97.46	−170.87	−76.48	−14318.89	7254.46	4.41	3.73	3.20	1.65	5.65
pentafluoroethane	22	0.85	55.74	−72.85	−33.73	−8240.28	4173.01	3.29	4.30	2.22	0.00	2.62
pentane	16	0.00	35.97	−172.85	−89.70	−4941.59	2546.18	2.41	1.35	0.71	0.35	3.28
pentanoic acid	21	1.80	41.82	−201.78	−93.84	−7877.15	4054.85	2.49	1.50	0.74	0.33	1.29
pentylbenzene	30	0.35	83.27	−329.64	−163.99	−14095.78	7199.13	4.47	2.94	1.91	1.22	4.64
perylene	46	0.00	191.44	−519.89	−244.34	−31604.37	16050.52	6.98	5.42	4.39	3.58	8.80
p-ethyltoluene	24	0.04	70.30	−264.82	−130.77	−10138.99	5190.43	3.38	2.34	1.53	0.82	3.10
phenanthrene	33	0.03	123.59	−370.13	−175.88	−17846.75	9099.63	4.82	3.51	2.65	1.95	5.26
phenol	18	1.14	50.76	−186.33	−86.08	−6303.73	3250.32	2.13	1.34	0.76	0.43	0.06
phenylacetylene	19	0.17	66.17	−209.24	−101.18	−6664.37	3433.70	2.45	1.51	0.94	0.53	2.35
propane	10	0.00	21.93	−107.33	−56.37	−2183.70	1138.19	1.41	0.71	0.00	0.00	2.85
propionaldehyde	12	2.51	24.59	−114.69	−55.01	−2911.90	1518.45	1.35	0.57	0.17	0.00	−0.72
propionic acid	15	1.80	27.69	−136.29	−60.52	−4511.46	2341.87	1.49	0.75	0.30	0.00	−1.13
propyl formate	18	3.90	36.95	−167.56	−77.27	−6059.11	3126.32	1.97	0.97	0.39	0.17	0.60
propylene oxide	12	1.83	24.63	−114.19	−54.69	−3086.58	1606.19	1.51	1.14	0.33	0.35	−1.01
propyne	8	0.36	20.97	−85.37	−43.40	−1323.75	702.66	0.79	0.29	0.00	0.00	1.04
p-xylene	21	0.05	62.95	−232.08	−114.17	−8117.10	4164.89	2.82	2.15	1.22	0.64	2.82
pyridine	15	1.93	43.62	−157.92	−74.07	−4596.28	2381.02	1.85	1.02	0.57	0.31	−1.10
pyrrole	13	2.18	35.76	−140.13	−65.44	−3607.98	1874.13	1.58	0.88	0.50	0.28	0.15
pyrrolidine	15	1.46	32.80	−156.97	−77.19	−4755.97	2453.60	2.21	1.46	0.96	0.63	−1.15
quinoline	24	1.85	81.74	−260.14	−121.81	−10450.14	5356.40	3.26	2.20	1.51	1.02	1.33
sec-butyl acetate	24	1.84	50.55	−232.95	−110.41	−10384.60	5318.42	2.84	1.83	0.94	0.37	1.27
sec-butylbenzene	27	0.27	75.50	−296.84	−147.27	−12546.39	6409.17	3.89	2.72	1.98	1.02	3.93
styrene	20	0.02	66.26	−219.95	−107.39	−7418.01	3813.35	2.61	1.61	1.04	0.59	2.53
tert-butyl methyl ether	19	1.28	39.57	−190.89	−94.83	−7225.42	3704.02	2.11	2.32	0.61	0.00	0.24
tert-butylbenzene	27	0.30	74.90	−296.73	−147.28	−12856.42	6563.49	3.66	3.62	1.64	0.94	3.66
tert-amylbenzene	30	0.37	82.31	−329.43	−163.89	−15335.88	7817.84	4.22	3.60	2.44	1.09	4.15
tert-butyl alcohol	16	1.54	31.58	−159.10	−78.18	−5284.02	2721.99	1.72	2.17	0.00	0.00	−1.13
tert-butylamine	16	1.36	34.08	−166.83	−83.62	−5260.11	2707.07	1.79	2.37	0.00	0.00	−1.14
tetrabromomethane	16	0.03	53.31	−34.05	−17.16	−4186.96	2118.40	3.92	11.54	0.00	0.00	3.14

Table 1. (Continued)

compound	no. of filled levels[a]	dipole moment [Debye]	average polarizability [au]	reson energy[b] [EV]	exch energy[c] [EV]	E−N attraction[d] [EV]	N−N repulsion[e] [EV]	$^1\chi^v$	$^2\chi^v$	$^3\chi^v$	$^4\chi^v$	−logS [mol/L]
tetrachloroethene	18	0.00	52.79	−66.79	−29.58	−5453.65	2765.25	2.51	2.41	1.28	0.00	2.54
tetrachloromethane	16	0.01	42.79	−40.77	−18.39	−4370.09	2217.22	2.26	3.85	0.00	0.00	2.31
tetrafluoroethene	18	0.00	19.59	−94.59	−30.73	−5796.43	2977.71	1.01	0.52	0.14	0.00	1.60
tetrafluoromethane	16	0.00	8.02	−72.20	−20.33	−4914.20	2533.84	0.76	0.43	0.00	0.00	3.67
thiophene	13	0.67	42.24	−118.17	−57.80	−3323.01	1723.03	2.41	1.61	1.05	0.68	1.44
thiophenol	16	1.69	47.46	−139.06	−63.29	−4859.74	2510.41	2.50	1.87	1.17	0.72	2.12
toluene	18	0.26	45.60	−198.61	−97.35	−6268.82	3226.81	2.41	1.65	0.94	0.53	2.22
trans-1,2-diphenylethene (*trans*-stilbene)	34	0.00	131.69	−377.35	−181.73	−17314.01	8837.12	4.73	3.15	2.07	1.32	5.80
trans-1,4-dimethylcyclohexane	24	0.00	55.40	−262.12	−133.47	−10665.54	5447.69	3.79	3.37	2.30	1.43	4.47
trans-1−2-dichloroethene	12	0.00	32.74	−64.31	−31.08	−2473.19	1272.82	1.64	0.75	0.43	0.00	1.19
trans-2-pentene	15	0.04	38.37	−161.96	−83.18	−4302.46	2223.65	2.03	0.98	0.47	0.24	2.54
tribromomethane	13	0.95	39.97	−34.81	−18.26	−2810.79	1430.23	3.40	6.66	0.00	0.00	1.91
trichloroethene	15	0.49	42.05	−65.54	−30.32	−3848.17	1961.50	2.07	1.62	0.74	0.00	1.96
trichlorofluoromethane	16	0.56	35.44	−46.82	−18.53	−4446.41	2265.04	1.89	2.57	0.00	0.00	2.10
trichloromethane	13	1.02	32.20	−40.67	−19.39	−2874.80	1467.01	1.96	2.22	0.00	0.00	1.17
trichloronitromethane	21	2.80	46.67	−93.71	−35.23	−7985.88	4085.39	2.25	2.87	0.59	0.00	2.01
trifluorobenzene	24	0.00	49.65	−187.22	−78.46	−9900.26	5068.92	2.30	1.59	0.83	0.52	2.51
triiodomethane	13	0.61	64.07	−31.48	−17.72	−2679.97	1361.96	4.33	10.83	0.00	0.00	3.55
trimethylamine	13	1.15	28.46	−133.48	−66.75	−3582.48	1854.11	1.34	1.34	0.00	0.00	−1.18
triphenylene	42	0.00	161.56	−472.82	−223.45	−26887.23	13669.57	6.23	4.64	3.65	2.86	6.74

[a] Number of doubly filled molecular orbitals. [b] Resonance energy. [c] Exchange energy. [d] Electron-nuclear attraction energy. [e] Nuclear-nuclear attraction energy. [f] $^1\chi^v$, $^2\chi^v$, $^3\chi^v$, and $^4\chi^v$ are the first-, second-, third-, and fourth-order connectivity indices.

data, compiled from the literature, is reported as logS in Table 1, where S is in units of mol/L. Both the logS solubility data and chemical input descriptors were normalized from 0 to 1, i.e., $A_n = (X - X_{min})/(X_{max} - X_{min})$, in which A_n is the normalized variable and X, X_{min}, X_{max} denote the input parameter and the minimum and maximum variable values in each respective input set.

The molecular descriptors for each compound were determined from knowledge of the chemical structure. Two-dimensional molecular structures were drawn using Molecular Modeling Pro 3.01[41] and then converted to three-dimensional molecular structures using the CAChe 3.2 WorkSystem.[42] The geometry of the three-dimensional structures were subsequently optimized using MOPAC,[43] a semiempirical molecular orbital modeling routine, with the PM3 Hamiltonian[44] to arrive at the compounds' minimum energy conformations. The PM3 method is based on the correct inclusion of one-center overlap (i.e. neglecting diatomic differential overlap only). Semiempirical quantum-chemical methods have been developed within the mathematical framework of the molecular orbital theory (SCF MO) but based on simplifications and approximations introduced into the computational procedure. For example, experimental data for atoms and prototype molecular systems have often been used to estimate values of the parameters utilized in semiempirical MO calculations.

During the MOPAC energy minimization, quantum chemical descriptors, derived from the PM3 MO theory were also calculated; these included average molecular polarizability, dipole moment, moments of inertia, ionization potential, number of doubly occupied (filled) MO levels, molecular weight, heat of formation, total energy, electronic energy, nuclear−nuclear (core−core), and energy components partitioned into the individual one-center and two-center terms were calculated. The total energy, in terms of the PM3 MO, is the sum of the total one-center and two-center terms. The one-center energy terms include electron−electron repulsion and electron−nuclear attraction. The two-center energy terms include resonance energy, exchange energy, electron−electron repulsion, electron−nuclear attraction, and nuclear−nuclear repulsion. The total electrostatic (or Coulombic) interaction is equal to the sum of the following two-center energy terms: electron−electron repulsion, electron−nuclear attraction, and nuclear−nuclear repulsion energy. The resonance energy corresponds to the difference in pi electrons completely delocalized and the pi electrons localized in a double bond.[45] Finally, the exchange energy involves two electrons where the energy of attraction is between the nuclei and the overlap charge in the bond.[45] We note that the total energy calculated by semiempirical methods has been shown to be a suitable descriptor in a number of QSPR studies.[46−48] Also, the energy partitioned into one-center and two-center terms have been used to develop molecular structure−biological activity relationships of organic compounds.[49]

Molecular topological descriptors, generated using Molecular Modeling Pro 3.01, included four valance molecular connectivity indices of orders 1, 2, 3, and 4 ($^1\chi^v$, $^2\chi^v$, $^3\chi^v$, $^4\chi^v$) and the second Kappa shape index, $^2\kappa$. Molecular connectivity indices are topological indices that encode two-dimensional structural information into numerical values or indexes.[50,51] The molecular structure is expressed topologically by a hydrogen-suppressed graph. The carbons (and heteroatoms) are represented as vertices, and bonds connecting atoms are represented as edges. Briefly, the connectivity indices $^m\chi^v$ are valance-weighted counts of connected subgraphs. The first-order term $^1\chi^v$ is related to the degree of branching and size of the molecule expressed as the number of non-hydrogen atoms. The second-order term $^2\chi^2$ represents a dissection of the molecular skeleton into "two contiguous bond" fragments. The third-order term $^3\chi^v$ is a weighted count of four atoms (three-bond) fragments representing the potential for rotation around the central bond and is the smallest molecular structure necessary for conformational variability. The $^3\chi^v$ index also reflects the degree of branching at each of the four atoms in the fragment. The fourth-order term, $^4\chi^v$, represents path, cluster, path/cluster, and cyclic subgraphs of four edges. Structural information

from the $^4\chi^v$ index is useful for compounds with at least five carbon atoms in a chain. Finally the kappa 2 shape index,[52] $^2\kappa$, is considered for characterizing the level of branching among isomers.

The set of input descriptors was selected using a nonlinear variable selection method based on a dynamic neural network genetic algorithm. Variable selection operations were performed on 10 different neural networks to statistically identify the set of descriptors based on a frequency distribution (of the descriptors selected in each run). The 11 descriptors that were selected as the final set of input descriptors for QSPR modeling were all at or above the level of selection of 70%. The final set of input parameters, as listed in Table 1, includes the first-, second-, third-, and fourth-order valence connectivity indices ($^1\chi^v$,$^2\chi^v$,$^3\chi^v$,$^4\chi^v$), dipole moment (sum of both point charge and hybrid), number of filled doubly occupied MO levels, first-order average molecular polarizability, resonance energy, exchange energy, electron−nuclear attraction (two-center term), and nuclear−nuclear (core−core) repulsion. We note that the dipole moment was the only parameter that was selected in all 10 runs. Valence molecular connectivity indices ranked within the 70th to 80th percentile range, with typically three out-of four indices being selected at a time. A selection frequency of 70% was exhibited for both the number of filled MO levels and first-order average polarizability. Finally, the energy terms ranked within the 70th to 90th percentile.

Fuzzy ARTMAP Neural System. The fuzzy ARTMAP neural network was used in the development of QSPR for aqueous solubility. The fuzzy ARTMAP architecture was introduced by Carpenter et al.[37] as a classifier for multidimensional data clustering. As shown in Figure 2, the fuzzy ARTMAP network consists of two fuzzy ART modules, *artA* and *artB*, that are linked together via an inter-ART module. Each ART system includes a field, F_o, of nodes that represent a current input vector; a field F_1 that receives both bottom-up input from F_o and top-down input from a field F_2 that represents the active code, or category. A fuzzy ARTMAP system actively searches for recognition categories, or hypotheses, for providing acceptable matches among the top-down expectations (based upon clusters of input features that are deemed to be relevant) and input patterns (bottom-up data).[34−38] The *artA* module learns how to categorize the input patterns (molecular descriptors), while *artB* module develops categories of the target patterns (physical property). During supervised learning, *artA* module receives the molecular descriptors, and *artB* module receives the correct physical property prediction of the input pattern presented to F_oa. The *artA* module attempts the prediction through the map field of the category to which the current target belongs. The inter-ART module, called a *map field*, is an associative learning network and an internal controller that is designed to create a minimal number of *artA* recognition categories, or "hidden units", by following the *match tracking rule*. It works by increasing the vigilance parameter ρ_a of *artA* by the minimal amount needed to correct a predictive error at *artB*. The vigilance parameter calibrates the minimal confidence that *artA* must have in a recognition category, or hypothesis, that is activated by an input pattern in order for *artA* to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing. When the search discovers a category that provides an acceptable match, the system locks into an attentive resonance whereby the input pattern refines the adaptive weights of the category based on any new information that it contains. A category modifies its previous learning only if its top-down expectation matches the input vector well enough to risk changing its defining characteristics. When a prediction is correct, both modules learn their respective inputs by modifying the prototypes of the corresponding category. If a category or hypothesis is not a good or acceptable match, then the selection and learning of another category and top-down expectation is automatically initiated. Otherwise, hypothesis testing selects a new category on which to base learning of a novel event. A predictive failure at *artB* increases the vigilance parameter ρ_a by the least amount needed to initiate hypothesis testing at *artA*. The process continues until the network finds either an *artA* category that predicts the category of the current target correctly or creates a new one and the corresponding map field, with the correctly learned current descriptors-physical property pair.

The fuzzy ARTMAP dynamics are determined by vigilance ρ_a, ρ_b $\rho_{ab} \in [0,1]$, learning rate β_a, $\beta_b \in [0,1]$, and choice $\alpha > 0$, parameters. The vigilance parameter calibrates how well an input pattern (or exemplar) must match the learned prototype or cluster of input features that the category deems to be relevant, for a category to be accepted. Thus, the vigilance controls the degree of generalization. The rate of learning, which determines how the map field weights change through time, is controlled by the learning rate parameter. We note that rare events are learned with rapid learning ($\beta = 1$). The choice parameter controls the fuzzy subsethood of the category choice function and accounts for the noise in the activation of the F_1 layer. When the choice parameter approaches zero, the choice function measures the degree to which the adaptive weight vector is a fuzzy subset of the input vector. Fuzzy ARTMAP is capable of autonomously calibrating how much compression or generalization should occur in each category (set by a range of fuzzy features that best match the input pattern) and for self-organizing stable recognition categories in response to arbitrary sequences of analogue input patterns.

The fuzzy ARTMAP system was recently modified by Giralt et al.[39] to implement predictive capabilities into this neural classifier. In the current study, this modified cognitive architecture was used to establish QSPR models capable of predicting aqueous solubility as in previous back-propagation models. Fuzzy ARTMAP incorporates predictive feedback to control the hypothesis testing cycle, and this embodies characteristics of a self-organizing production system. The basic learning mechanism of this neural system consists of the creation of new categories (equivalent to hidden units in back-propagation) when dissimilar molecular descriptors and different values of the physical property are encountered. The degree of similarity is determined by threshold indicators such as the vigilance and choice parameters. Additional information about fuzzy ART and fuzzy ARTMAP systems can be found elsewhere.[53−58]

The fuzzy ARTMAP based QSPR for aqueous solubility was developed following the methodology described in Figure 1. About 85% (437) of compounds in the complete data set were selected for training by the fuzzy ART classifier to ensure that adequate information was provided to the

Table 2. Experimental and Predicted Solubilities at 25 °C Using Fuzzy ARTMAP and Back-Propagation Neural Network QSPRs

	data sets[a]		logS [mol/L]			absolute logS error	
compound	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
10-butylbenz[*a*]anthracene	tr	te	−7.55	−7.55	−8.18	0.00	0.63
1-bromohexane	tr	te	−3.81	−3.81	−3.12	0.00	0.69
1-chlorooctane	tr	te	−4.48	−4.47	−4.62	0.01	0.14
1-methoxy-2-propanol	tr	te	1.05	1.06	0.89	0.01	0.16
1-methylcyclohexene	tr	te	−3.27	−3.27	−3.19	0.00	0.08
1-nitropropane	tr	te	−0.77	−0.77	−0.52	0.00	0.25
1-nonene	tr	te	−5.05	−5.05	−4.76	0.00	0.29
1-octanol	tr	te	−2.38	−2.38	−2.24	0.00	0.14
1-propanol	tr	te	−1.22	−1.22	−1.08	0.00	0.14
2-bromotoluene	tr	te	−2.23	−2.22	−2.47	0.01	0.24
2-chlorobutane	tr	te	−1.96	−1.96	−1.87	0.00	0.09
2-heptanone	tr	te	−1.45	−1.45	−1.32	0.00	0.13
2-heptyne	tr	te	−2.64	−2.64	−3.52	0.00	0.88
2-methyl-2-butene	tr	te	−2.56	−2.56	−2.21	0.00	0.35
2-methyl-2-pentanol	tr	te	−0.5	−0.50	−0.82	0.00	0.32
2-nitrotoluene	tr	te	−0.72	−0.72	−0.69	0.00	0.03
2-pentanone	tr	te	−0.3	−0.30	−0.46	0.00	0.16
3-chloroiodobenzene	tr	te	−3.55	−3.55	−3.33	0.00	0.22
3-ethylphenol	tr	te	0.54	0.55	−0.72	0.01	1.26
3-pentanol	tr	te	−0.23	−0.23	−0.33	0.00	0.10
4-bromophenol	tr	te	−1.09	−1.09	−1.01	0.00	0.08
4-methyl-2-pentylacetate	tr	te	−2.05	−2.05	−1.90	0.00	0.15
4-methylpyridine	tr	te	−0.103	−0.09	−0.35	0.01	0.25
5-methylchrysene	tr	te	−6.59	−6.59	−6.82	0.00	0.23
7-ethylbenz[*a*]anthracene	tr	te	−6.8	−6.80	−7.09	0.00	0.29
7-methylbenz[*a*]anthracene	tr	te	−7.35	−7.35	−7.61	0.00	0.26
a,a,a-trifluorotoluene	tr	te	−2.51	−2.51	−2.09	0.00	0.42
a-chlorotoluene	tr	te	−2.43	−2.43	−1.83	0.00	0.60
butyl acetate	tr	te	−1.27	−1.27	−0.90	0.00	0.37
coronene	tr	te	−9.33	−9.33	−8.73	0.00	0.60
cyclohexane	tr	te	−3.18	−3.18	−3.48	0.00	0.30
diisobutylamine	tr	te	−1.77	−1.77	−2.04	0.00	0.27
ethane	tr	te	−2.7	−2.69	−2.82	0.01	0.12
ethyl benzene	tr	te	−1.36	−1.35	−1.84	0.01	0.48
ethylcyclohexane	tr	te	−4.25	−4.25	−4.34	0.00	0.09
furan	tr	te	−0.83	−0.82	−0.75	0.01	0.08
hexane	tr	te	−3.96	−3.96	−3.72	0.00	0.24
isoamyl formate	tr	te	−1.52	−1.52	−1.38	0.00	0.14
isobutanol	tr	te	0.009	0.01	*otl*[b]	0.00	*otl*[b]
methanol	tr	te	1.49	1.49	1.61	0.00	0.12
n-ethylaniline	tr	te	−1.7	−1.70	−1.57	0.00	0.13
n-hexylbenzene	tr	te	−5.21	−5.21	−5.18	0.00	0.03
n-pentyl acetate	tr	te	−1.88	−1.88	−1.41	0.00	0.47
n-propylamine	tr	te	1.23	1.23	1.04	0.00	0.19
n-propylbenzene	tr	te	−3.36	−3.36	−3.36	0.00	0.00
n-propylcyclopentane	tr	te	−7.74	−7.74	−7.23	0.00	0.51
o-cresol (2-hydroxytoluene)	tr	te	−0.62	−0.62	−0.64	0.00	0.02
pentachlorobenzene	tr	te	−5.65	−5.65	−5.38	0.00	0.27
pentafluoroethane	tr	te	−2.62	−2.62	−2.63	0.00	0.01
propyne	tr	te	−1.04	−1.03	−1.17	0.01	0.13
tert-amylbenzene	tr	te	−4.15	−4.15	−4.31	0.00	0.16
tert-butyl alcohol	tr	te	1.13	1.13	−0.06	0.00	1.19
tetrafluoromethane	tr	te	−3.67	−3.66	−3.96	0.01	0.29
thiophenol	tr	te	−2.12	−2.12	−2.29	0.00	0.17
10-pentylbenz[*a*]anthracene	tr	tr	−8.57	−8.57	−8.20	0.00	0.37
12-methylbenz[*a*]anthracene	tr	tr	−6.68	−6.68	−7.31	0.00	0.63
1-bromo-2-chloroethane	tr	tr	−1.32	−1.32	−1.86	0.00	0.54
1-bromobutane	tr	tr	−2.2	−2.20	−1.99	0.00	0.21
1-bromoheptane	tr	tr	−4.42	−4.42	−3.78	0.00	0.64
1-bromonaphthalene	tr	tr	−4.35	−4.35	−4.05	0.00	0.30
1-bromopentane	tr	tr	−3.08	−3.08	−2.82	0.00	0.26
1-bromopropane	tr	tr	−1.7	−1.70	−1.49	0.00	0.21
1-butanethiol	tr	tr	−2.18	−2.18	−1.36	0.00	0.82
1-butene	tr	tr	−2.4	−2.39	−1.97	0.01	0.43
1-chloro-2-fluoroethane	tr	tr	−0.51	−0.50	−0.60	0.01	0.09
1-chloro-2-methylpropane	tr	tr	−2	−1.99	−1.93	0.01	0.07
1-chlorobutane	tr	tr	−2.03	−2.03	−1.99	0.00	0.04
1-chlorohexane	tr	tr	−3.12	−3.12	−2.77	0.00	0.35
1-chloronaphthalene	tr	tr	−3.93	−3.93	−3.45	0.00	0.48
1-chloropentane	tr	tr	−2.73	−2.72	−2.46	0.01	0.27
1-choropropane	tr	tr	−1.46	−1.45	−1.03	0.01	0.43
1-decanol	tr	tr	−3.63	−3.63	−3.74	0.00	0.11

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
1-ethylnaphthalene	tr	tr	−4.17	−4.17	−4.65	0.00	0.48
1-heptanol	tr	tr	−1.95	−1.94	−1.59	0.01	0.36
1-heptyne	tr	tr	−3.01	−3.01	−2.91	0.00	0.10
1-hexene	tr	tr	−3.23	−3.23	−2.83	0.00	0.40
1-hexyne	tr	tr	−2.36	−2.35	−2.20	0.01	0.16
1-iodopropane	tr	tr	−2.29	−2.29	−1.92	0.00	0.37
1-methylbenz[*a*]anthracene	tr	tr	−6.64	−6.64	−5.70	0.01	0.94
1-methylnaphthalene	tr	tr	−3.7	−3.69	−4.01	0.00	0.31
1-nonyne	tr	tr	−4.26	−4.25	−3.94	0.01	0.32
1-octene	tr	tr	−4.44	−4.44	−3.85	0.00	0.59
1-octyne	tr	tr	−3.61	−3.60	−3.23	0.01	0.38
1-pentanol	tr	tr	−0.6	−0.60	−0.59	0.00	0.01
1-pentene	tr	tr	−2.68	−2.67	−2.31	0.01	0.37
1-pentyne	tr	tr	−1.64	−1.63	−1.70	0.01	0.06
1-propene	tr	tr	−2.32	−2.32	−1.88	0.00	0.44
2-bromofluorobenzene	tr	tr	−2.7	−2.69	−2.33	0.01	0.37
2-bromopropane	tr	tr	−1.59	−1.59	−1.44	0.00	0.15
2-butanol	tr	tr	0.39	0.39	−0.28	0.00	0.67
2-butanone	tr	tr	0.49	0.49	−0.11	0.00	0.60
2-chloro-2-methylbutane	tr	tr	−2.51	−2.51	−2.34	0.00	0.17
2-chloro-2-methylpropane	tr	tr	−2.2	−2.20	−1.97	0.00	0.23
2-chloroethanol	tr	tr	1.09	1.09	0.81	0.00	0.28
2-chloropropane	tr	tr	−1.41	−1.41	−1.25	0.00	0.16
2-chloropyridine	tr	tr	−0.75	−0.74	−0.89	0.01	0.14
2-ethylbutyric acid	tr	tr	−0.81	−2.81	−0.98	0.00	0.17
2-ethylpyridine	tr	tr	0.5	−0.80	−0.82	0.01	1.32
2-fluorochlorobenzene	tr	tr	−2.42	−2.42	−2.66	0.00	0.24
2-heptene	tr	tr	−3.82	−3.81	−3.84	0.01	0.02
2-hexanol	tr	tr	−1.42	−1.41	−0.96	0.01	0.46
2-hexanone	tr	tr	−0.79	−0.79	−0.84	0.00	0.05
2-iodopropane	tr	tr	−2.09	−2.09	−1.84	0.00	0.25
2-methyl-1-butanol	tr	tr	−0.47	−0.47	−0.61	0.00	0.14
2-methyl-1-pentene	tr	tr	−3.03	−3.03	−2.81	0.00	0.22
2-methyl-2-butanol	tr	tr	0.1	0.10	−0.31	0.00	0.41
2-methyl-3-hexyne	tr	tr	−2.59	−2.58	−2.75	0.01	0.16
2-methyl-3-pentanone	tr	tr	−0.81	−0.80	−0.84	0.01	0.03
2-methylaziridine	tr	tr	1.24	1.24	1.01	0.00	0.23
2-methyldecalin	tr	tr	−6.57	−6.56	−6.16	0.01	0.41
2-methylheptane	tr	tr	−5.08	−5.08	−4.43	0.00	0.65
2-methylhexane	tr	tr	−4.6	−4.60	−3.94	0.00	0.66
2-methylpentane	tr	tr	−3.74	−3.74	−3.63	0.00	0.11
2-nitropropane	tr	tr	−1.75	−1.74	−1.49	0.01	0.26
2-nonanone	tr	tr	−2.57	−2.56	−2.49	0.01	0.08
2-pentanol	tr	tr	−0.3	−0.30	−0.50	0.00	0.20
2-propanone	tr	tr	1.24	1.24	0.13	0.00	1.11
2-undecanone	tr	tr	−3.94	−3.93	−3.92	0.01	0.02
3-bromo-1-propene	tr	tr	−1.5	−1.50	−1.21	0.00	0.29
3-bromofluorobenzene	tr	tr	−2.67	−2.67	−2.12	0.00	0.55
3-chlorophenol	tr	tr	−0.69	−0.69	−0.82	0.00	0.13
3-heptanol	tr	tr	−1.46	−1.45	−1.51	0.01	0.05
3-heptanone	tr	tr	−1.42	−1.41	−1.34	0.01	0.08
3-hexanone	tr	tr	−0.83	−0.82	−0.87	0.01	0.04
3-hexyne	tr	tr	−1.99	−1.99	−2.18	0.00	0.19
3-methyl-1-butene	tr	tr	−2.73	−2.72	−2.39	0.01	0.34
3-methyl-2-butanol	tr	tr	−0.16	−0.15	−0.32	0.01	0.16
3-methyl-2-butanone	tr	tr	−0.15	−0.15	−0.38	0.00	0.23
3-methylcholanthrene	tr	tr	−7.92	−7.92	−8.08	0.00	0.16
3-methylheptane	tr	tr	−5.16	−5.16	−5.36	0.00	0.20
3-methylpropylbenzene	tr	tr	−3.96	−3.96	−3.98	0.00	0.02
3-methylpyridine	tr	tr	−1.03	−1.03	−1.17	0.00	0.14
3-methylthiophene	tr	tr	−2.39	−2.39	−1.95	0.00	0.44
3-pentanone	tr	tr	−0.28	−0.28	−0.45	0.00	0.17
4-bromo-1-butene	tr	tr	−2.25	−2.25	−1.64	0.00	0.61
4-bromoiodobenzene	tr	tr	−4.56	−4.56	−4.22	0.00	0.34
4-bromotoluene	tr	tr	−3.19	−3.18	−3.29	0.01	0.10
4-chloroaniline	tr	tr	−1.51	−1.50	−1.46	0.01	0.05
4-chloroiodobenzene	tr	tr	−4.03	−4.03	−4.18	0.00	0.15
4-chlorophenol	tr	tr	−0.73	−0.72	−1.04	0.01	0.31
4-ethenylcyclohexene	tr	tr	−3.34	−3.33	−3.47	0.01	0.13
4-fluoro-10-methylbenz[*a*]anthracene	tr	tr	−7.72	−7.72	−7.57	0.00	0.15
4-fluoroiodobenzene	tr	tr	−3.13	−3.13	−2.85	0.00	0.28
4-heptanone	tr	tr	−1.55	−1.55	−1.36	0.00	0.19

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
4-methyl-1-pentene	tr	tr	−3.24	−3.23	−3.46	0.01	0.22
4-methyl-2-methoxybenzene	tr	tr	−1.82	−1.81	−1.36	0.01	0.46
4-methyloctane	tr	tr	−6.05	−6.05	−5.81	0.00	0.24
4-methylphenol	tr	tr	−0.7	−0.69	−0.76	0.01	0.06
4-tertbutylphenol	tr	tr	4.31	4.31	4.37	0.00	0.06
5-fluoro-7-methylbenz[*a*]anthracene	tr	tr	−7.72	−7.72	−7.60	0.00	0.12
5-methylbenzo[*a*]pyrene	tr	tr	−8.52	−8.52	−8.13	0.00	0.39
6-chloro-10-methylbenz[*a*]anthracene	tr	tr	−8.7	−8.70	−7.95	0.00	0.75
9-chloro-7-methylbenz[*a*]anthracene	tr	tr	−7.44	−7.44	−7.98	0.00	0.54
9-methylbenz[*a*]anthracene	tr	tr	−6.56	−6.56	−6.70	0.00	0.14
a,a,a-trichlorotoluene	tr	tr	−2.51	−2.51	−2.17	0.00	0.34
acenaphthene	tr	tr	−4.63	−4.62	−4.24	0.01	0.39
acenaphthylene	tr	tr	−3.96	−3.96	−4.04	0.00	0.08
acetaldehyde	tr	tr	1.36	1.36	1.39	0.00	0.03
acetamide	tr	tr	1.08	1.09	1.19	0.01	0.11
acetic acid	tr	tr	−1.22	−1.22	−1.41	0.00	0.19
acetonitrile	tr	tr	1.39	1.39	1.37	0.00	0.02
acetophenone	tr	tr	−1.29	−1.29	−1.13	0.00	0.16
allyl alcohol (2-propen-1-ol)	tr	tr	0.74	0.74	0.80	0.00	0.06
aniline	tr	tr	−0.41	−0.41	−0.35	0.00	0.06
anthracene	tr	tr	−6.35	−6.35	−5.88	0.00	0.47
benz[*a*]anthracene	tr	tr	−7.21	−7.21	−7.35	0.00	0.14
benz[*b*]anthracene	tr	tr	−8.6	−8.60	−7.46	0.00	1.14
benzaldehyde	tr	tr	−3.24	−3.23	−2.70	0.01	0.54
benzene	tr	tr	−1.64	−1.63	−1.69	0.01	0.05
benzo[*a*]pyrene	tr	tr	−7.8	−7.80	−7.96	0.00	0.16
benzo[*ghi*]perylene	tr	tr	−9.03	−9.03	−8.29	0.00	0.74
benzo[*j*]fluoranthene	tr	tr	−8	−8.00	−7.81	0.00	0.19
benzo[*k*]fluoranthene	tr	tr	−8.49	−8.49	−7.83	0.00	0.66
benzoic acid	tr	tr	−1.81	−1.81	−1.78	0.00	0.03
benzonitrile	tr	tr	−1.01	−1.00	−0.82	0.01	0.19
benzyl acetate	tr	tr	−1.69	−1.69	−1.65	0.00	0.04
biphenyl	tr	tr	−4.31	−4.31	−4.90	0.00	0.59
bromochloromethane	tr	tr	−0.89	−0.89	−1.07	0.00	0.18
bromoethane	tr	tr	−1.09	−1.09	−1.24	0.00	0.15
bromomethane	tr	tr	−0.8	−0.80	−1.15	0.00	0.35
butanol	tr	tr	−0.07	−0.06	−0.22	0.01	0.15
butylamine	tr	tr	1.14	1.15	*otl*[b]	0.01	*otl*[b]
butyraldehyde	tr	tr	−0.01	−0.01	*otl*[b]	0.00	*otl*[b]
chlorobenzene	tr	tr	−2.44	−2.43	−2.33	0.01	0.11
chlorodifluoromethane	tr	tr	−1.47	−1.47	−0.92	0.00	0.55
chloroethane	tr	tr	−1.06	−1.05	−0.95	0.01	0.11
chloroethene	tr	tr	−1.3	−1.29	−1.24	0.01	0.06
chloromethane	tr	tr	−0.98	−0.98	−0.76	0.00	0.22
chloropentafluoroethane	tr	tr	−2.79	−2.79	−2.01	0.00	0.78
cholanthrene	tr	tr	−7.85	−7.85	−7.90	0.00	0.05
chrysene	tr	tr	−8.06	−8.06	−7.32	0.00	0.74
cycloheptane	tr	tr	−3.51	−3.51	−3.98	0.00	0.47
cyclohexanol	tr	tr	−0.43	−0.43	−0.64	0.00	0.21
cyclohexanone	tr	tr	−0.63	−0.62	−0.98	0.01	0.35
cyclooctane	tr	tr	−4.15	−4.15	−4.55	0.00	0.40
cyclopentane	tr	tr	−2.59	−2.58	−2.91	0.01	0.32
cyclopentanone	tr	tr	−0.96	−0.96	−0.80	0.00	0.16
cyclopentene	tr	tr	−2.1	−2.09	−2.23	0.01	0.13
cyclopropane	tr	tr	−2.04	−2.03	−1.87	0.01	0.17
decachlorobiphenyl	tr	tr	−11.62	−11.62	−9.82	0.00	1.80
decalin (decahydronaphthalene)	tr	tr	−5.19	−5.18	−5.66	0.01	0.47
decan-2-one	tr	tr	−3.3	−3.30	−3.18	0.00	0.12
decanal	tr	tr	−3.41	−3.40	−3.28	0.01	0.13
di(*n*-butyl) ether	tr	tr	−2.64	−2.64	−2.55	0.00	0.09
di(*n*-propyl) ether	tr	tr	−1.32	−1.32	−1.30	0.00	0.02
dibromodifluoromethane	tr	tr	−1.89	−1.88	−1.81	0.01	0.08
dibromomethane	tr	tr	−1.16	−1.15	−1.26	0.01	0.10
dichlorodifluoromethane	tr	tr	−2.64	−2.64	−2.49	0.00	0.15
dichlorofluoromethane	tr	tr	−1.03	−1.03	−1.13	0.00	0.10
dichloromethane	tr	tr	−0.82	−0.82	−1.02	0.00	0.20
diethyl sulfide	tr	tr	−1.25	−1.24	−1.34	0.01	0.09
diethyl ether	tr	tr	−0.09	−0.09	−0.26	0.00	0.17
difluoromethane	tr	tr	−1.81	−1.81	*otl*[b]	0.00	*otl*[b]
diisopropylamine	tr	tr	0.04	0.04	0.01	0.00	0.03
diisopropyl ether	tr	tr	−1.06	−1.05	−0.93	0.01	0.13
dimethylamine	tr	tr	1.56	1.56	1.38	0.00	0.18

Table 2. (Continued)

	data sets[a]		logS [mol/L]			absolute logS error	
compound	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
dimethyl ether	tr	tr	0.55	0.55	0.44	0.00	0.11
dimethyl sulfide	tr	tr	−0.74	−0.74	−0.83	0.00	0.09
di-*n*-butylamine	tr	tr	−1.57	−1.57	−1.65	0.00	0.08
di-*n*-propylamine	tr	tr	−0.24	−0.23	*otl*[b]	0.01	*otl*[b]
diphenylmethane	tr	tr	−4.08	−4.62	−5.66	0.00	1.58
dodecane	tr	tr	−7.67	−7.67	−6.94	0.00	0.73
ethanol	tr	tr	1.34	1.34	1.35	0.00	0.01
ethene	tr	tr	−2.33	−2.32	−2.23	0.01	0.10
ethyl acetate	tr	tr	−0.04	−0.04	−0.10	0.00	0.06
ethyl benzoate	tr	tr	−2.32	−2.32	−2.13	0.00	0.19
ethyl mercaptan	tr	tr	−0.73	−0.72	−0.67	0.01	0.06
ethylallylamine	tr	tr	−1.35	−1.35	−1.41	0.00	0.06
ethyl butyrate	tr	tr	−1.37	−1.37	−1.56	0.00	0.19
fluorene	tr	tr	−5	−5.00	−4.85	0.00	0.15
fluoromethane	tr	tr	−1.28	−1.27	−1.36	0.01	0.08
formaldehyde	tr	tr	1.12	1.13	0.93	0.01	0.19
formic acid	tr	tr	−1.34	−1.34	−1.31	0.00	0.03
heptanal	tr	tr	−1.96	−1.96	−1.73	0.00	0.23
heptanoic acid	tr	tr	−1.66	−1.66	−1.71	0.00	0.05
hexabromobenzene	tr	tr	−9.74	−9.74	−8.49	0.00	1.25
hexachloro-1-propene	tr	tr	−4.16	−4.15	−4.17	0.01	0.01
hexachlorobenzene	tr	tr	−6.78	−6.78	−6.88	0.00	0.10
hexachlorocyclopentadiene	tr	tr	−5.18	−5.18	−5.33	0.00	0.15
hexachloroethane	tr	tr	−4.49	−4.49	−4.54	0.00	0.05
hexadecane	tr	tr	−8.4	−8.40	−7.67	0.00	0.73
hexamethylbenzene	tr	tr	−5.23	−5.23	−5.27	0.00	0.04
hexanal	tr	tr	−1.22	−1.22	−0.89	0.00	0.33
hexanoic acid	tr	tr	−1.05	−1.05	−0.95	0.00	0.10
hydroquinone	tr	tr	−0.17	−0.17	−0.26	0.00	0.09
indane	tr	tr	−3.13	−3.13	−2.95	0.00	0.18
iodobenzene	tr	tr	−2.95	−2.95	−2.84	0.00	0.11
iodoethane	tr	tr	−1.6	−1.59	−1.68	0.01	0.08
iodomethane	tr	tr	−1	−1.00	−1.14	0.00	0.14
isoamyl acetate	tr	tr	−1.81	−1.81	−1.73	0.00	0.08
isobutane	tr	tr	−3.08	−3.08	−2.82	0.00	0.26
isobutyl acetate	tr	tr	−1.24	−1.24	−0.95	0.00	0.29
isobutylamine	tr	tr	1.14	1.15	1.24	0.01	0.10
isobutylbenzene	tr	tr	−4.12	−4.12	−3.87	0.00	0.25
isobutyraldehyde	tr	tr	0.09	0.10	−0.11	0.01	0.20
isopentyl acetate	tr	tr	−1.06	−1.05	−1.23	0.01	0.17
isophorone	tr	tr	−0.55	−0.55	−0.82	0.00	0.27
isopropyl alcohol	tr	tr	1.22	1.23	1.16	0.01	0.06
isopropyl formate	tr	tr	−0.63	−0.62	−0.47	0.01	0.16
m-bromotoluene	tr	tr	−3.52	−3.51	−3.27	0.01	0.25
m-chlorotoluene	tr	tr	−3.52	−3.51	−3.70	0.01	0.18
methyl acetate	tr	tr	0.52	0.52	0.26	0.00	0.26
methylamine	tr	tr	1.54	1.54	1.60	0.00	0.06
methyl butyl ether	tr	tr	−0.99	−0.98	−0.65	0.01	0.34
methyl butyrate	tr	tr	−0.83	−0.82	−0.72	0.01	0.11
methyl formate	tr	tr	0.58	0.58	0.32	0.00	0.26
methyl hexanoate	tr	tr	−1.99	−1.99	−1.74	0.00	0.25
methyl isopropyl ether	tr	tr	−0.06	−0.06	−0.18	0.00	0.12
methyl mercaptan	tr	tr	−0.49	−0.49	−0.55	0.00	0.06
methyl propanoate	tr	tr	−0.15	−0.15	−0.07	0.00	0.08
methyl propyl ether	tr	tr	−0.39	−0.39	0.22	0.00	0.61
methylcyclohexane	tr	tr	−3.85	−3.84	−3.83	0.01	0.02
naphthalene	tr	tr	−3.6	−3.60	−4.11	0.00	0.51
n-butyl propionate	tr	tr	−1.94	−1.94	−1.83	0.00	0.11
n-butyl acetate	tr	tr	−1.27	−1.27	−0.90	0.00	0.37
n-butylformate	tr	tr	−1.13	−1.13	−0.83	0.00	0.30
n-heptane	tr	tr	−4.53	−4.53	−4.20	0.00	0.33
n-hexyl acetate	tr	tr	−2.45	−2.45	−2.00	0.00	0.45
n-hexylamine	tr	tr	−0.69	−0.69	−0.82	0.00	0.13
nitrobenzene	tr	tr	−1.88	−1.88	−1.94	0.00	0.06
nitromethane	tr	tr	−0.26	−0.26	0.08	0.00	0.34
N-methylaniline	tr	tr	−1.28	−1.27	−1.00	0.01	0.28
N,*N*-dimethyl formamide	tr	tr	1.14	1.15	1.04	0.01	0.10
nonanal	tr	tr	−2.58	−2.58	−2.58	0.00	0.00
nonanoic acid	tr	tr	−2.87	−2.87	−2.79	0.00	0.08
n-pentylamine	tr	tr	1.06	1.06	−0.98	0.00	2.04
n-pentylcyclopentane	tr	tr	−6.08	−6.08	−5.44	0.00	0.64
n-propyl acetate	tr	tr	−0.73	−0.72	−0.47	0.01	0.26

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
octachloronaphthalene	tr	tr	−9.7	−9.70	−8.74	0.00	0.96
octanal	tr	tr	−2.36	−2.35	−2.57	0.01	0.21
octane	tr	tr	−5.11	−5.11	−4.67	0.00	0.44
octanoic acid	tr	tr	−2.26	−2.25	−2.09	0.01	0.17
p-chloronitrobenzene	tr	tr	−2.54	−2.53	−2.61	0.01	0.07
p-xylene	tr	tr	−2.82	−2.81	−3.02	0.01	0.20
pyridine	tr	tr	1.1	1.10	1.21	0.00	0.11
pyrrole	tr	tr	−0.15	−0.15	−0.21	0.00	0.06
pyrrolidine	tr	tr	1.15	1.15	1.24	0.00	0.09
quinoline	tr	tr	−1.33	−1.32	−1.55	0.01	0.22
sec-butylbenzene	tr	tr	−3.93	−3.93	−4.07	0.00	0.14
styrene	tr	tr	−2.53	−2.53	−2.38	0.00	0.15
tert-butyl methyl ether	tr	tr	−0.24	−0.23	−0.43	0.01	0.19
tert-butylbenzene	tr	tr	−3.66	−3.66	−3.67	0.00	0.01
tetrabromomethane	tr	tr	−3.14	−3.13	−2.96	0.01	0.18
tetrachloroethene	tr	tr	−2.54	−2.53	−2.79	0.01	0.25
tetrachloromethane	tr	tr	−2.31	−2.31	−2.32	0.00	0.01
tetrafluoroethene	tr	tr	−1.6	−1.59	−1.49	0.01	0.11
thiophene	tr	tr	−1.44	−1.44	−1.63	0.00	0.19
toluene	tr	tr	−2.22	−2.22	−2.23	0.00	0.01
trans-2-pentene	tr	tr	−2.54	−2.53	−2.75	0.01	0.21
tribromomethane	tr	tr	−1.91	−1.91	−1.63	0.00	0.28
trichloroethene	tr	tr	−1.96	−1.96	−2.15	0.00	0.19
trichloromethane	tr	tr	−1.17	−1.17	−1.28	0.00	0.11
trichloronitromethane	tr	tr	−2.01	−2.01	−2.19	0.00	0.18
trifluorobenzene	tr	tr	−2.51	−2.51	−2.33	0.00	0.18
trimethylamine	tr	tr	1.18	1.18	1.11	0.00	0.07
triphenylene	tr	tr	−6.74	−6.74	−7.29	0.00	0.55
ethyl formate	tr	tr	0.08	0.08	−0.13	0.00	0.21
pentane	tr	tr	−3.28	−3.27	−3.21	0.01	0.07
pentanoic acid	tr	tr	−1.29	−1.29	−1.45	0.00	0.16
perylene	tr	tr	−8.8	−8.80	−7.95	0.00	0.85
p-ethyltoluene	tr	tr	−3.1	−3.10	−3.05	0.00	0.05
phenanthrene	tr	tr	−5.26	−5.26	−5.71	0.00	0.45
phenol	tr	tr	−0.06	−0.06	−0.14	0.00	0.08
phenylacetylene	tr	tr	−2.35	−2.35	−2.84	0.00	0.49
propane	tr	tr	−2.85	−2.85	−2.44	0.00	0.41
propionaldehyde	tr	tr	0.72	0.72	0.65	0.00	0.07
propionic acid	tr	tr	1.13	1.13	0.91	0.00	0.22
propyl formate	tr	tr	−0.6	−0.60	−0.46	0.00	0.14
propylene oxide	tr	tr	1.01	1.01	0.98	0.00	0.03
1-chloro-1,1-difluoroethane	tr	tr	−1.2	−1.20	−1.14	0.00	0.06
4-bromo-1,1′-biphenyl	tr	tr	−5.55	−5.55	−5.21	0.00	0.34
2,2′,5-trichloro-1,1′-biphenyl	tr	tr	−6.02	−6.02	−6.26	0.00	0.24
2,4′-dichloro-1,1′-biphenyl	tr	tr	−5.28	−5.28	−5.56	0.00	0.28
2,5-dichloro-1,1′-biphenyl	tr	tr	−5.36	−5.36	−5.22	0.00	0.14
2,6-dichloro-1,1′-biphenyl	tr	tr	−5.21	−5.21	−5.69	0.00	0.48
2,2′,3,3′-tetrachloro-1,1′-biphenyl	tr	tr	−7.27	−7.27	−7.30	0.00	0.03
3,4-dichloro-1,1-biphenyl	tr	tr	−6.39	−6.39	−5.70	0.00	0.69
3,3′,4,4′-tetrachloro-1,1′-biphenyl	tr	tr	−8.73	−8.73	−8.94	0.00	0.21
4,4′-bromo-1,1′-biphenyl	tr	tr	−7.74	−7.74	−7.50	0.00	0.24
4,4′-dichloro-1,1′-biphenyl	tr	tr	−6.56	−6.56	−6.62	0.00	0.06
4,4′-dimethyl-1,1′-biphenyl	tr	tr	−6.02	−6.02	−5.98	0.00	0.04
2,2′,3,3′,4,6′-hexachlorobiphenyl	tr	tr	−6.96	−6.96	−7.18	0.00	0.22
2,2′,5,5′-tetrabromo-1,1′-biphenyl	tr	te	−8.06	−8.06	−8.60	0.00	0.54
2,2′,5,6′-tetrachloro-1,1′-biphenyl	tr	tr	−6.79	−6.78	−7.02	0.01	0.23
2,2′,4,4′,6,6′-hexabromobiphenyl	tr	tr	−9	−9.00	−8.78	0.00	0.22
2,2′,6,6′-tetrachloro-1,1′-biphenyl	tr	tr	−7.39	−7.39	−7.85	0.00	0.46
trans-1,2-diphenylethene	tr	tr	−5.8	−5.80	−5.88	0.00	0.08
7,12-dimethyl-1,2-benzanthracene	tr	tr	−7.02	−7.02	−7.86	0.00	0.84
2-ethyl-1,3-hexandiol	tr	te	−2.81	0.51	−2.42	0.01	0.39
cis-1,3-dichloro-1-propene	tr	tr	−2.05	−2.05	−1.58	0.00	0.47
hexachloro-1,3-butadiene	tr	tr	−4.92	−4.91	−4.59	0.01	0.33
2,3-dimethyl-1,3-butadiene	tr	tr	−2.4	−2.39	−3.00	0.01	0.60
trans-1,4-dimethylcyclohexane	tr	tr	−4.47	−4.47	−4.35	0.00	0.12
2,5-cyclohexadiene-1−4-dione	tr	te	−0.99	−0.98	−1.06	0.01	0.07
1,1-dichlorobutane	tr	tr	−2.4	−1.29	−2.23	0.00	0.17
1,1-dichloroethane	tr	tr	−1.29	−2.39	−1.40	0.01	0.11
1,1-dichloroethene	tr	tr	−1.63	−1.63	−1.18	0.00	0.45
1,1-dichlorotetrafluoroethane	tr	tr	−2.93	−2.93	−2.64	0.00	0.29
1,1-difluoroethane	tr	tr	−0.57	−0.57	−0.10	0.00	0.47
1,1-difluoroethene	tr	tr	−2.55	−2.55	−2.18	0.00	0.37

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
1,1-biphenylethylene	tr	tr	−4.44	−4.44	−5.49	0.00	1.05
1,2-dichlorobenzene	tr	te	−2.97	−2.97	−2.28	0.00	0.69
1,2-benzofluorene	tr	tr	−6.68	−6.68	−6.87	0.00	0.19
1,2-dibromo-3-chloropropane	tr	tr	−2.38	−2.38	−2.60	0.00	0.22
1,2-dichloroethane	tr	tr	−1.06	−1.05	−0.97	0.01	0.09
1,2-dichoroethene (trans)	tr	tr	−1.44	−1.44	−1.57	0.00	0.13
1,2-diiodobenzene	tr	tr	−4.24	−4.23	−4.19	0.01	0.05
1,2-diiodoethene	tr	tr	−7.37	−7.37	−6.33	0.00	1.04
1,3-dichloropropane	tr	te	−1.62	−1.62	−1.89	0.00	0.27
1,3-butadiene	tr	tr	−2.13	−2.12	−2.06	0.01	0.07
1,3-difluorobenzene	tr	tr	−2	−1.99	−1.74	0.01	0.26
1,3-dimethylnaphthalene	tr	tr	−4.29	−4.28	−4.37	0.01	0.08
1,4-cyclohexadiene	tr	tr	−2.06	−2.05	−2.26	0.01	0.20
1,4-cyclopentadiene	tr	tr	−2.06	−2.05	−1.13	0.01	0.93
1,4-dichlorobenzene	tr	tr	−3.27	−3.27	−3.59	0.00	0.32
1,4-dichlorobutane	tr	tr	−0.92	−0.92	−1.92	0.00	1.00
1,4-diethylbenzene	tr	tr	−3.75	−3.74	−4.52	0.01	0.77
1,4-difluorobenzene	tr	tr	−1.97	−1.96	−2.24	0.01	0.27
1,4-pentadiene	tr	tr	−2.09	−2.09	−2.29	0.00	0.20
1,5-dichloropentane	tr	tr	−3.05	−3.05	−2.93	0.00	0.12
1,5-hexadiene	tr	tr	−2.69	−2.69	−2.63	0.00	0.06
1,6-heptadiene	tr	tr	−3.34	−3.33	−3.33	0.01	0.01
1,6-heptadiyne	tr	tr	−1.75	−1.74	−2.29	0.01	0.54
2,2-dimethy-3-hexyne	tr	tr	−3.03	−3.03	−3.15	0.00	0.12
2,2-dimethylpropane	tr	tr	−3.33	−3.33	−2.72	0.00	0.61
2,2-dimethylpropanoic acid	tr	tr	−0.67	−0.67	−0.43	0.00	0.24
2,3-benzofluorene	tr	tr	−8.04	−8.04	−7.88	0.00	0.16
2,3-dichloro-2-methylbutane	tr	tr	−2.69	−2.69	−3.38	0.00	0.69
2,3-dichlorobutane	tr	tr	−2.7	−2.69	−2.74	0.01	0.04
2,3-dichloropropene	tr	tr	−1.71	−1.70	−2.08	0.01	0.37
2,3-dimethyl-2-butanol	tr	tr	−0.37	−0.37	−0.58	0.00	0.21
2,3-dimethylbutane	tr	tr	−3.58	−3.58	−3.67	0.00	0.09
2,3-dimethylnaphthalene	tr	tr	−4.72	−4.72	−4.27	0.00	0.45
2,3-dimethylpentane	tr	tr	−4.28	−4.28	−3.98	0.00	0.30
2,4-dimethyl-3-pentanone	tr	te	−1.3	−1.29	−1.24	0.01	0.06
2,4-dimethylpyridine	tr	tr	0.51	0.51	−0.69	0.00	1.20
2,5-dimethylpyridine	tr	tr	0.4	0.40	−0.68	0.00	1.08
1,2:5,6-dibenzanthracene	tr	tr	−7.44	−7.44	−8.22	0.00	0.78
2,6,dimethylnaphthalene	tr	tr	−4.89	−4.89	−4.97	0.00	0.08
2,6-dichlorophenol	tr	tr	−1.79	−1.79	−1.54	0.00	0.25
2,6-diethylaniline	tr	tr	−2.35	−2.35	−2.85	0.00	0.50
2,6-dimethyl-4-heptanone	tr	tr	−1.73	−1.73	−2.08	0.00	0.35
1,2:7,8-dibenzanthracene	tr	te	−8.66	−8.66	−8.22	0.00	0.44
3,3-dimethylpentane	tr	tr	−4.23	−4.23	−3.99	0.00	0.24
3,3-dimethyl-2-butanone	tr	tr	−0.72	−0.72	−0.75	0.00	0.03
3,4-dichlorotoluene	tr	tr	−3.79	−3.79	−2.85	0.00	0.94
3,4-dimethylpyridine	tr	tr	0.36	0.36	−0.52	0.00	0.88
3,5-dimethylpyridine	tr	te	0.38	0.39	−0.62	0.01	1.00
9,10-dimethylanthracene	tr	tr	−6.57	−6.56	−6.79	0.01	0.22
1,1,1-trichloroethane	tr	te	−2	−1.99	−1.95	0.01	0.05
1,1,1-trichloropropane	tr	tr	−1.89	−1.88	−2.31	0.01	0.42
2-chloro-1,1,1-trifluoroethane	tr	tr	−1.15	−1.15	−1.23	0.00	0.08
1,1,1,2-tetrachlorodifluoroethane	tr	tr	−3.31	−3.30	−2.98	0.01	0.33
1,1,1,2-tetrachloroethane	tr	tr	−2.18	−2.18	−2.50	0.00	0.32
1,1,2-trichloroethane	tr	tr	−1.48	−1.47	−1.71	0.01	0.23
1,1,2,2-tetrabromoethane	tr	tr	−2.72	−2.72	−2.23	0.00	0.49
1,1,2,2-tetrachloroethane	tr	tr	−1.74	−1.74	−2.04	0.00	0.30
2,2,2-trifluoroethanol	tr	tr	−0.23	−0.23	−0.72	0.00	0.49
1,2,3-trichloropropane	tr	tr	−1.93	−1.93	−2.01	0.00	0.08
1,2,3-trimethylbenzene	tr	tr	−3.2	−3.20	−3.72	0.00	0.52
2,2,3-trimethylbutane	tr	tr	−4.36	−4.36	−3.87	0.00	0.49
1,2,4-trichlorobenzene	tr	tr	−3.59	−3.58	−3.44	0.01	0.15
1,2,3,4-tetrahydronaphthalene	tr	te	−4.37	−4.36	−4.58	0.01	0.21
2,3,4-trimethylpentane	tr	te	−4.7	−4.70	−4.25	0.00	0.45
2,4,4′-trichloro-1,1′-biphenyl	tr	te	−6.21	−6.21	−6.43	0.00	0.22
1,1,3,4,4-pentachloro-1,2-butadiene	tr	tr	−4.23	−4.23	−4.45	0.00	0.22
2,2′,3,4,4′,5′-hexachlorobiphenyl	tr	tr	−8.32	−8.32	−8.24	0.00	0.08
1,3,5-trichlorobenzene	tr	tr	−4.48	−4.47	−4.22	0.01	0.26
1,3,5-trimethylbenzene	tr	tr	−3.4	−3.40	−3.72	0.00	0.32
1,4,5-trimethylnaphthalene	tr	tr	−4.91	−4.91	−5.06	0.00	0.15
2,2,5-trimethylhexane	tr	tr	−5.05	−5.05	−4.60	0.00	0.45
1,2,4,5-tetrafluorobenzene	tr	tr	−2.38	−2.38	−2.36	0.00	0.02

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
1,2,4,5-tetramethylbenzene	tr	tr	−3.84	−3.84	−4.31	0.00	0.47
2,2,5,5-tetramethyl-3-hexyne	tr	tr	−3.69	−3.69	−4.50	0.00	0.81
2,2′,4,5,5′-pentachlorobiphenyl	tr	tr	−7.44	−7.44	−7.94	0.00	0.50
2,3,6-trichloronaphthalene	tr	tr	−7.14	−7.14	−6.80	0.00	0.34
2,4,6-tribromobiphenyl	tr	tr	−7.39	−7.39	−7.94	0.00	0.55
4-chloro-2-nitrophenol	val	val	−3.09	−2.51	−2.96	0.58	0.13
1,1,2,2,-tetrachlorodifluoroethane	val	val	−3.19	−3.30	−3.17	0.11	0.02
1,1,2-trichloropropane	val	val	−1.89	−1.88	−2.03	0.01	0.14
1,1,3-trimethylcyclopentane	val	val	−4.48	−4.47	−3.86	0.01	0.62
1,2,3,5-tetrafluorobenzene	val	val	−2.31	−2.38	−1.62	0.07	0.69
1,2,3,5-tetramethylbenzene	val	val	−3.76	−3.84	−3.71	0.08	0.05
1,2,4-trimethylbenzene	val	val	−3.32	−3.20	−3.38	0.12	0.06
1,2-dichloropropane	val	val	−1.61	−1.70	−1.50	0.09	0.11
1,2-dichlorotetrafluoroethane	val	val	−2.74	−2.93	−2.63	0.19	0.11
1,2-diethylbenzene	val	val	−3.28	−3.93	−3.53	0.65	0.25
1,2-difluorobenzene	val	val	−2	−1.99	−1.88	0.01	0.12
1,3-dichlorobenzene	val	val	−3.07	−3.13	−3.39	0.06	0.32
1,4-dimethylnaphthalene	val	val	−4.14	−4.89	−4.98	0.75	0.84
1,5-dimethylnaphthalene	val	val	−4.74	−4.89	−4.97	0.15	0.23
10-ethylbenz[*a*]anthracene	val	val	−6.78	−7.02	−7.90	0.24	1.12
10-methylbenz[*a*]anthracene	val	val	−6.64	−6.56	−7.57	0.08	0.93
1-decene	val	val	−5.51	−6.05	−4.99	0.54	0.52
1-nonanol	val	val	−3.13	−2.87	−2.96	0.26	0.17
2,2′,4,4′,6,6′-hexachlorobiphenyl	val	val	−8.71	−8.32	−8.55	0.39	0.16
2,2,4-trimethylpentane	val	val	−4.7	−4.70	−3.60	0.00	1.10
2,2-dimethylbutane	val	val	−3.56	−3.58	−3.35	0.02	0.21
2,2-dimethylpentane	val	val	−4.36	−4.60	−4.51	0.24	0.15
2,3,5,6-tetrachloro-1,1′biphenyl	val	val	−7.32	−7.39	−7.95	0.07	0.63
2,3-dimethylpyridine	val	val	0.38	0.40	−0.29	0.02	0.67
2,4,6-trichloro-1,1′-biphenyl	val	val	−6.14	−6.21	−6.38	0.07	0.24
2,4-dichloro-1,1′-biphenyl	val	val	−5.25	−5.21	−5.83	0.04	0.58
2,4-dimethylpentane	val	val	−4.26	−4.60	−3.98	0.34	0.28
2-bromonaphthalene	val	val	−4.4	−4.35	−4.04	0.05	0.36
2-chloro-1-nitrobenzene	val	val	−2.55	−2.53	−2.62	0.02	0.07
2-chloronaphthalene	val	val	−4.14	−3.93	−3.38	0.21	0.76
2-chloropentane	val	val	−2.63	−2.72	−2.52	0.09	0.11
2-ethylnaphthalene	val	val	−4.29	−4.17	−4.59	0.12	0.30
2-hexene	val	val	−3.1	−3.23	−3.28	0.13	0.18
2-methyl-1,3-butadiene	val	val	−2.03	−2.09	−1.93	0.06	0.10
2-methylnaphthalene	val	val	−3.76	−3.69	−3.97	0.07	0.21
2-methylpyridine	val	val	−1.03	−1.03	−1.16	0.00	0.13
2-octanone	val	val	−2.05	−2.35	−1.86	0.30	0.19
3,3′,5,5′-tetrachloro-1,1′-biphenyl	val	val	−8.37	−8.73	−7.83	0.36	0.54
3,3′,5-trichloro-1,1′-biphenyl	val	val	−6.01	−6.02	−6.55	0.01	0.54
3-chloro-1-propene	val	val	−1.28	−1.09	−0.95	0.19	0.33
3-methyl-3-pentanol	val	val	−0.38	−0.37	−0.53	0.01	0.15
3-methylhexane	val	val	−4.58	−4.60	−4.81	0.02	0.23
3-methylpentane	val	val	−3.84	−3.58	−3.55	0.26	0.29
4-methyl-1,1′-biphenyl	val	val	−4.62	−4.62	−4.94	0.00	0.32
4-methyl-2-pentanone	val	val	−0.72	−0.80	−0.83	0.08	0.11
5,6-dimethylchrysene	val	val	−7.01	−6.80	−7.84	0.21	0.83
5-methyl-2-hexanone	val	val	−1.33	−1.45	−1.27	0.12	0.06
6-fluoro-7-methylbenz[*a*]anthracene	val	val	−7.72	−7.72	−7.63	0.00	0.09
6-methylbenzo[*a*]pyrene	val	val	−8.53	−8.52	−8.13	0.01	0.40
6-methylchrysene	val	val	−6.57	−6.68	−6.93	0.11	0.36
amyl propionate (pentyl)	val	val	−2.25	−2.45	−2.14	0.20	0.11
benzo[*b*]fluoranthene	val	val	−8.23	−8.49	−7.88	0.26	0.35
benzo[*e*]pyrene	val	val	−7.6	−7.60	−7.90	0.00	0.30
butylbenzene	val	val	−4.06	−3.96	−4.00	0.10	0.06
cis-2-pentene	val	val	−2.54	−2.53	−2.75	0.01	0.21
cis-1,2-dichloroethene	val	val	−1.44	−1.44	−1.56	0.00	0.12
cis-1,2-dimethylcyclohexane	val	val	−4.44	−4.47	−4.37	0.03	0.07
cumene	val	val	−3.29	−3.10	−3.47	0.19	0.18
cycloheptene	val	val	−3.16	−3.27	−3.19	0.11	0.03
cyclohexene	val	val	−2.59	−2.22	−2.79	0.37	0.20
decane	val	val	−6.44	−6.08	−6.58	0.36	0.14
diethylamine	val	val	1.14	1.15	1.26	0.01	0.12
isobutene	val	val	−2.33	−2.39	−2.20	0.06	0.13
isobutyl formate	val	val	−1	−1.13	−0.86	0.13	0.14
isobutyl methyl ether	val	val	−0.9	−0.98	−0.87	0.08	0.03
isopropylamine	val	val	1.23	1.18	1.10	0.05	0.13
m-cresol (3-hydroxytoluene)	val	val	−0.68	−0.69	−0.95	0.01	0.27

Table 2. (Continued)

compound	data sets[a]		logS [mol/L]			absolute logS error	
	FAM	Bk-Pro	reported	FAM	Bk-Pro	FAM	Bk-Pro
methycyclopentane	val	val	−3.3	−3.18	−3.20	0.12	0.10
m-xylene	val	val	−2.8	−2.81	−2.79	0.01	0.01
n-amyl acetate	val	val	−1.88	−1.88	−1.41	0.00	0.47
n-butane	val	val	−2.98	−3.08	−2.82	0.10	0.16
N,*N*-dimethylaniline	val	val	−1.79	−1.70	−1.33	0.09	0.46
nonane	val	val	−5.88	−6.05	−5.73	0.17	0.15
n-pentylbenzene	val	val	−4.59	−4.15	−4.59	0.44	0.00
sec-butyl acetate	val	val	−1.27	−1.27	−1.19	0.00	0.08
tert-butyl amine	val	val	1.14	1.13	1.10	0.01	0.04
trichlorofluoromethane	val	val	−2.1	−1.88	−1.88	0.22	0.22
triiodomethane	val	val	−3.55	−3.13	−2.57	0.42	0.98

[a] te = test set, tr = training set, val = validation set, FAM = fuzzy ARTMAP/QSPR, Bk-Pro = back-propagation/QSPR. [b] *otl* = outlier identified and removed from back-propagation data sets.

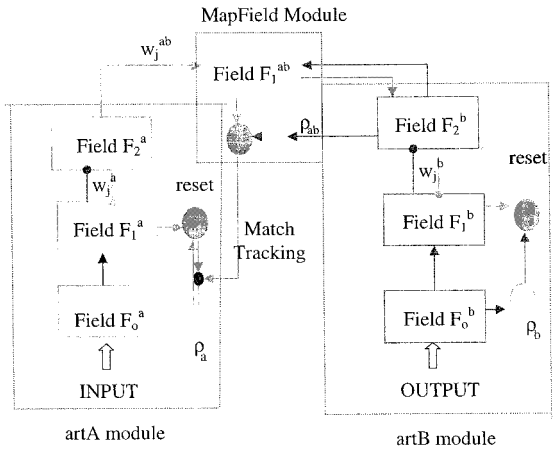

Figure 2. Block diagram of fuzzy ARTMAP neural network.

system. The compounds were then classified based on their molecular descriptors and the target physical property, which were all presented to the fuzzy ARTMAP as an input vector. Complement coding was used in fuzzy logic to represent the inverse relationship of the data (or attributes "not present"). The program then generated a complement code of the data set. Training of the fuzzy ARTMAP consisted of presenting the molecular descriptors and target property of the training set to modules *artA* and *artB* (see Figure 2), respectively, to establish input with the output categories and relate them (F_1ab). After training, the hypothesis components of the *artB* modules (F_o^b and F_1^b) were disconnected, and an output in its category layer F_2^b was implemented;[34−38] the intent is to provide, through the map field module F_1^{ab}, a prediction for the target physical property for any input of descriptors presented to module *artA*. The model was then validated using a validation set containing 78 compounds.

Back-Propagation Neural Network System. The normalized aqueous solubility data and molecular descriptors, scaled from 0 to 1, were divided into three data sets: training, test (or recall), and validation (or generalization). To compare the two neural network approaches for estimating aqueous solubility, for the same set of compounds, the validation set for the back-propagation model, containing 78 data points (15% of the complete data set), was taken to be identical to the validation set used for the fuzzy ARTMAP model. The remaining 437 compounds were randomly divided into training and test sets. To retain a sufficient number of compounds for training (∼72% of the data), while maintaining an adequate data set for testing, the test set included the training data (372) and nontraining data (65). During model development, five compounds from the training set (butyraldehyde, di-*n*-propylamine, difluoromethane, butylamine, and isobutanol) could not be handled by the back-propagation network, possibly due to lack of sufficient compounds in their respective classes. Although solubility estimation errors for these compounds were prohibitively large. they did not affect the quality of training. Therefore, the above four compounds were removed from the data set, and the back-propagation model was built from a training and test set of 367 and 432 compounds, respectively. It is emphasized that the above problem was not encountered with the fuzzy ARTMAP neural network model.

Model building with the back-propagation neural network proceeded with same 11 input descriptors, as used for the fuzzy ARTMAP based model, with aqueous solubility as output. The architecture of the neural network was developed using a cascade method of network construction, together with an adaptive gradient learning rule.[59] The hyperbolic tangent transfer functions were chosen to correlate weighted inputs and outputs of the hidden layer. The performance of the resulting neural network model was subsequently tested and validated using two separate data subsets as described above.

III. RESULTS AND DISCUSSION

The optimal fuzzy ARTMAP/QSPR model was obtained (i.e. training phase) for the following parameter settings: vigilance parameters $\rho_a = 0.0$, $\rho_b = 0.999$, $\rho_{ab} = 0.999$, learning rate parameters $\beta_a = 1$, $\beta_b = 1$, and choice parameter $\alpha = 0.1$. The optimal back-propagation neural network solubility QSPR had an architecture of 11−13−1. The estimated aqueous solubilities from the fuzzy ARTMAP and the optimal 11−13−1 back-propagation neural network-QSPR models are summarized in Table 2 and Figures 3−6. Table 2 also lists, for each model, the absolute error between predictions and the complete data set, training, and validation sets. A summary of the error analysis is provided in Table 3.

The fuzzy ARTMAP model, which was trained with 437 compounds selected with the fuzzy ART classifier, predicted the aqueous solubility for the complete data set of 515 compounds with average absolute and maximum errors and standard deviation of 0.02, 0.75, and 0.08 logS units, respectively (see Table 3). The fuzzy ARTMAP/QSPR,

which was validated with a vigilance parameter of $\rho_a = 0.9$, performed with an average absolute error, maximum absolute error, and standard deviation of 0.14, 0.75, and 0.16 logS units, respectively. Clearly, performance of the fuzzy ARTMAP was excellent for the relatively large heterogeneous compound data set. However, there were several compounds for which solubility errors were significantly higher than the average. Inspection of the fuzzy ARTMAP classification revealed that the errors were linked to the manner by which the network classified the input descriptors. Fuzzy ARTMAP generally tends to group together compounds with similar sets of descriptors. However, in certain cases, parameter sets can be similar for structural isomers, while there maybe significant differences in their aqueous solubilities. For example, 1,4-dimethylnaphthalene, with a logS = −4.14, was classified with other dimethylnaphthalene isomers, where the predicted aqueous solubility for the group was based on the aqueous solubility of 2,6-dimethylnaphthalene (logS = −4.89). As a consequence, this classification resulted in an absolute error of 0.75 logS units (for the validation set). Another example is 4-chloro-2-nitrophenol, the only chloronitrophenol in the data set, which could not be properly classified with other chloronitrophenols, thus the associated absolute error was relatively high (0.58 logS units). 4-Chloro-2-nitrophenol, which contains multiple functional groups (a halogen, a nitro group, and a phenol), was classified into a group containing other halogenated aromatic hydrocarbons, the best available category in terms of structural features. In retrospect, since the data set lacked other chloronitrophenols, it could have been more appropriate to place 4-chloro-2-nitrophenol into the training set. Another compound that exhibited a relatively high solubility estimation error was 1,2-diethylbenzene (0.65 logS units) that was classified with *sec*-butylbenzene, both compounds with a molecular formula of $C_{10}H_{14}$. If 1,2-diethylbenzene would have been classified with 1,4-diethylbenzene (logS = −3.75), the absolute error would decrease to −0.47 logS units. The above examples suggest that, although the overall performance of the fuzzy ARTMAP is as good or better than other previously published approaches, further improvements would necessitate a larger data set that contains multiple compounds per class. Moreover, it may prove useful to explore alternate sets of molecular descriptors for the fuzzy ARTMAP that would allow a greater ability to differentiate among complex or apparently very similar structures.

Performance comparison of the optimal 11−13−1 back-propagation and fuzzy ARTMAP based logS models was accomplished using the data set used for validation of the fuzzy ARTMAP based QSPR. The average absolute and maximum errors and standard deviation for the validation set of the back-propagation/QSPR model were 0.28, 1.1, and 0.26 logS units, respectively. The average absolute and maximum errors and standard deviation for both the training and test sets of the back-propagation model were 0.29, 2.0, and 0.28 logS units, respectively. Compounds from the training set which exhibited the largest average absolute errors included *n*-pentylamine (2.04 logS units), decachlorobiphenyl (1.8 logS units), diphenylmethane (1.58 logS units), 2-ethylpyridine (1.32 logS units), and 3-ethylphenol (1.26 logS units). For the validation set, the maximum absolute error of 1.12 logS units resulted from 10-ethylbenz[*a*]anthracene. 10-Methylbenz[*a*]anthracene, also present in the validation set, exhibited a large average absolute error of 0.93 logS units. Also, relatively high average absolute errors in the validation set were encountered for 2,2,4-trimethylpentane (1.1 logS units) and triiodomethane (0.97 logS units). We note that dimethylpyridine's (C_7H_9N), as a class, exhibited a high absolute error. However, although the average absolute error attributed to the five dimethylpyridine isomers and ethylpyridine in the data set, was 1.0 logS units, the average absolute error attributed to the three methylpyridine isomers was only 0.17 logS units. The relative performance of the fuzzy ARTMAP and back-propagation QSPRs for aqueous solubility are summarized in Table 3. Overall, the errors were higher for the back-propagation/QSPR model.

Table 3. QSPR Performance for Solubility ($logS_{[mol/L]}$) Predictions Using Fuzzy ARTMAP and Back-Propagation Neural Networks

		average absolute error		maximum absolute error		standard deviation	
data set	records	logS	%	logS	%	logS	%
Fuzzy ARTMAP							
all data	515	2.4×10^{-2}	0.94	0.75	20	0.08	2.6
train	437	2.8×10^{-3}	0.3	1.3×10^{-2}	14	4.5×10^{-3}	1.4
validation	78	0.14	4.5	0.75	20	0.16	4.8
Back-Propagation							
all data	510	0.29	21	2.0	408	0.28	45
train	367	0.29	23	2.0	408	0.28	48
test	432	0.29	23	2.0	408	0.28	48
validation	78	0.28	11	1.1	177	0.26	21

Additional insight into the performance of the present QSPR models is revealed by chemical group-specific errors as given in Table 4. In general, the quantum chemical descriptors were satisfactory for characterizing the structural information of about 23 different chemical groups in the data set. This was demonstrated by the ability of the fuzzy ARTMAP/QSPR to classify and predict solubility based on the input descriptors. Overall, the quantum chemical descriptors were more suitable for characterizing molecular structure for PCBs, aromatics, alkanes, and PAHs as judged by the more accurate solubility predictions for these compound groups. Average |logS| errors of less than 1% were predicted by the fuzzy ARTMAP/QSPR for PCBs, halogenated aromatic and aliphatic hydrocarbons, amines, PAHs, nitro compounds, and alkanes. For the fuzzy ARTMAP model, average |logS| errors were in the lower 1% range for some of the oxygen containing compounds such as ester, ethers, alcohols, and anilines. Ketones and the 15 heterocyclic ring nitrogen containing compounds (i.e., pyridines, pyrrole) exhibited average |logS| errors around 1.6%. Somewhat higher errors were encountered for phenols, where the average |logS| error was 2%. Similar group-specific analysis of the back-propagation based QSPR revealed a less accurate differentiation among the chemical classes with the same set of quantum chemical descriptors. Based on chemical classes, the back-propagation based QSPR was most accurate for PCBs, aromatics, alkanes, and PAHs with average |logS| errors of 5.3%, 6.53%, 7.45%, and 7.53%, respectively. Reasonable results were also obtained for the amides (carbonyl and nitrogen group), furans (heterocyclic oxygen ring), alkenes, and nitriles (nitrogen with triple bond), where average |logS| errors for the groups were about 9−10%. It is noted that although amides, furans, and nitriles represented less than 0.05% of the data set, the back-propagation logS

Table 4. Error Analysis Based on Functional Group

functional groups	count	average absolute logS error		percent absolute logS error	
		fuzzy ARTMAP	back-propagation	fuzzy ARTMAP	back-propagation
alkanes	42	0.04	0.33	0.93	7.45
alkenes	30	0.05	0.26	1.44	9.51
alkynes	11	0.01	0.28	0.31	11.63
aromatics	27	0.05	0.17	1.49	6.53
PAHs (with some halogenated PAHs)	71	0.04	0.45	0.73	7.53
PCBs, PBrBs	28	0.03	0.37	0.43	5.37
halogenated nonaromatic	88	0.02	0.31	0.81	16.65
halogenated aromatics	33	0.01	0.34	0.29	10.58
carboxcyclic acids	11	0.00	0.13	0.04	11.96
esters	27	0.01	0.21	1.00	34.79
aldehydes	11	0.00	0.19	1.19	31.71
ketones	25	0.02	0.19	1.64	30.84
ethers	11	0.01	0.17	1.37	64.31
alcohols	30	0.01	0.24	1.27	59.40
phenols (with halogens, nitro, sulfur)	12	0.05	0.28	2.07	43.44
furan (heterocyclic ring − oxygen)	1	0.01	0.08	1.20	9.51
mercaptans (sulfure containing)	7	0.00	0.25	0.31	15.34
amines	16	0.01	0.44	0.73	65.24
nitriles	2	0.01	0.10	0.50	9.97
nitro	8	0.00	0.16	0.22	25.14
anilines	6	0.02	0.25	1.08	15.88
amides (carbonyl CO + nitrogen)	2	0.01	0.10	0.90	9.39
heterocyclic ring-nitrogen (e.g. pyridines)	15	0.00	0.50	1.59	121.82
quinoline, pyrole, pyridine, methlypridines	9	0.00	0.15	1.98	41.51
dimethylpyridines, ethylpyridines	6	0.01	1.03	1.00	242.30

QSPR predicted solubilities for the group within an average |logS| error of 10%. The average |logS| error for the halogenated aromatic hydrocarbons of 10.6% (0.34 logS units) was slightly lower than the average |logS| error of 16.6% (0.31 logS) for the halogenated aliphatic hydrocarbons. Halogenated aliphatic hydrocarbons, the dominant chemical group, represented 17% of the compounds in the data set. The larger errors in predicted solubilities for this group generally occurred for $-7.5 \leq \log S \leq -3.5$ and where the data were sparse. Absolute logS errors in the 11% range occurred for alkynes and carboxylic acids, while errors in the 15% range resulted for the mercaptans (sulfur containing compounds) and anilines (aromatic with a nitrogen). Relatively low average absolute solubility estimation errors of 0.19, 0.19, and 0.21 logS units resulted for ketones, aldehydes, and esters, respectively, although the average |logS| errors were in the range of 30%. Nitro containing compounds and ethers also exhibited relatively low average absolute errors of 0.16 and 0.17 logS units, respectively, but relatively high average |logS| errors of 25% and 64%, respectively. The solubility data for the group of ethers were poorly distributed, thus attributing to the higher error during network training. Errors for phenols and alcohols were similar with average absolute and percent logS errors of 0.28 (43%) and 0.24 (59%) logS units, respectively. We note that the solubility data for the amines were not adequately distributed within the solubility domain; thus, the back-propagation based QSPR was less optimal for this group. Amines, representing 3% of the data set, exhibited relatively high average and percent |logS| errors of 0.44 logS and 65%, respectively. Although the data set for the heterocyclic nitrogen compounds, which included pyrrolidine, dimethylpyridine isomers, and methylpyridine isomers, was relatively sparse, errors were only about 0.15 logS units (or 41%), with the highest errors for compounds with molecular formula C_7H_9N (i.e. dimethylpyridine). It is emphasized, that reasonable solubility correlations were achieved with the present back-propagation QSPR model. Many of the errors were attributed to those chemical groups whose data sets were either too sparse or too concentrated in a particular region (e.g., amines, ether). These observations reinforce the importance of using a uniformly distributed data set for training to improve the back-propagation network's ability to interpolate correctly.

It is important to recognize that there is significant uncertainty in reported aqueous solubility data. Uncertainties in published experimental aqueous solubility data have been reported to range from 0.01 up to 0.70 logS units.[10,12−16,19,23,40] For example, Sutter and Jurs[12] reported that variability of reported aqueous solubility data, compiled from several sources, was as high as 0.34 logS units (or mean absolute difference of 0.11 logS units).[12] Discrepancies in reported aqueous solubility measurements are typically due to large experimental error usually associated with compounds with low solubility. According to Katritzky et al.,[16] the average standard deviation of experimental aqueous solubility data (from various references) has been reported to be about 0.58 logS units.[16] In contrast, validation set estimates from the present fuzzy ARTMAP and back-propagation QSPRs had significantly lower standard deviations of 0.16 and 0.28 logS units, respectively.

The range of average estimation errors for published neural network and multiple linear regression/QSPR solubility models[12−16,19,21] are in the range of 0.168 to 0.573 logS units, relative to an average absolute error of 0.16 logS units for the present fuzzy ARTMAP model. Although the performance of the present back-propagation QSPR was inferior to the fuzzy ARTMAP QSPR, the absolute error and standard deviation of 0.28 and 0.26 logS units, respectively, were within the range or lower relative to other previously published models. For example, Mitchell and Jurs[13] reported a 9-6-1 architecture neural network based logS QSPR model which performed with rms error of 0.343 logS units, for a

Table 5. Comparison of Neural Network and Multilinear Regression Models for Predicting Aqueous Solubility of Selected Organic Compounds

	logS [mol/L]						
compound	reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
1,1,1,2-tetrachlorodifluoroethane	−3.31	−3.30	−2.98			−3.83	
1,1,1,2-tetrachloroethane	−2.18	−2.18	−2.50			−2.66	−2.62
1,1,1-trichloroethane	−2.00	−1.99	−1.95			−1.95	−1.97
1,1,1-trichloropropane	−1.89	−1.88	−2.31			−1.89	
1,1,2,2,-tetrachlorodifluoroethane	−3.19	−3.30	−3.17				−3.03
1,1,2,2-tetrabromoethane	−2.72	−2.72	−2.23			−3.26	−3.96
1,1,2,2-tetrachloroethane	−1.74	−1.74	−2.04	−1.78		−2.86	−2.45
1,1,2-trichloroethane	−1.48	−1.47	−1.71			−1.90	−1.67
1,1,2-trichloropropane	−1.89	−1.88	−2.03			−2.27	
1,1-dichlorobutane	−2.40	−2.39	−2.23			−2.43	−2.51
1,1-dichloroethane	−1.29	−1.29	−1.40			−1.52	−1.27
1,1-dichloroethene	−1.63	−1.63	−1.18			−1.65	
1,1-difluoroethene	−2.55	−2.55	−2.18			−0.81	
1,2,3,4-tetrahydronaphthalene	−4.37	−4.36	−4.58				−3.91
1,2,3-trichloropropane	−1.93	−1.93	−2.01			−2.23	
1,2,3-trimethylbenzene	−3.20	−3.20	−3.72	−3.50	−3.30	−3.69	−3.62
1,2,4,5-tetramethylbenzene	−3.84	−3.84	−4.31	−4.04			−4.27
1,2,4-trichlorobenzene	−3.59	−3.58	−3.44	−3.84	−3.89	−3.43	−3.79
1,2,4-trimethylbenzene	−3.32	−3.20	−3.38	−3.53		−3.69	−3.62
1,2:5,6-dibenzanthracene (DB[a,h])	−7.44	−7.44	−8.22	−8.86			
1,2-benzofluorene	−6.68	−6.68	−6.87	−7.47			
1,2-dibromo-3-chloropropane	−2.38	−2.38	−2.60			−2.01	
1,2-dichlorobenzene	−2.97	−2.97	−2.28	−2.98	−3.18	−2.93	−3.11
1,2-dichloroethane	−1.06	−1.05	−0.97			−1.30	−1.08
1,2-dichloropropane	−1.61	−1.70	−1.50			−2.28	−1.67
1,2-dichoroethene (trans)	−1.44	−1.44	−1.57				−1.36
1,2-diethylbenzene	−3.28	−3.93	−3.53			−4.11	
1,2-difluorobenzene	−2.00	−1.99	−1.88				−1.96
1,2-diiodobenzene	−4.24	−4.23	−4.19				−3.97
1,3,5-trichlorobenzene	−4.48	−4.47	−4.22	−3.76	−3.88		−3.78
1,3,5-trimethylbenzene	−3.40	−3.40	−3.72	−3.82		−3.70	−3.69
1,3-butadiene	−2.13	−2.12	−2.06		−1.98		
1,3-dichlorobenzene	−3.07	−3.13	−3.39	−3.10	−3.13	−2.91	−3.08
1,3-dichloropropane	−1.62	−1.62	−1.89			−1.73	−1.60
1,3-difluorobenzene	−2.00	−1.99	−1.74				−1.85
1,3-dimethylnaphthalene	−4.29	−4.28	−4.37	−4.47			−4.72
1,4,5-trimethylnaphthalene	−4.91	−4.91	−5.06	−4.91			−5.15
1,4-cyclohexadiene	−2.06	−2.05	−2.26		−2.24		−1.99
1,4-dichlorobenzene	−3.27	−3.27	−3.59	−3.16	−3.14	−2.91	−3.00
1,4-dichlorobutane	−0.92	−0.92	−1.92			−2.17	
1,4-difluorobenzene	−1.97	−1.96	−2.24				−1.87
1,4-dimethylnaphthalene	−4.14	−4.89	−4.98	−4.38			−4.66
1,4-pentadiene	−2.09	−2.09	−2.29		−1.96		−1.63
1,5-dichloropentane	−3.05	−3.05	−2.93			−2.66	
1,5-dimethylnaphthalene	−4.74	−4.89	−4.97	−4.38			−4.62
1,5-hexadiene	−2.69	−2.69	−2.63		−2.57		−2.33
1,6-heptadiene	−3.34	−3.33	−3.33		−3.30		−2.78
1,6-heptadiyne	−1.75	−1.74	−2.29		−1.76		
1-bromo-2-chloroethane	−1.32	−1.32	−1.86			−1.15	−1.33
1-bromobutane	−2.20	−2.20	−1.99		−2.40	−2.37	−2.61
1-bromoheptane	−4.42	−4.42	−3.78			−3.77	−4.53
1-bromohexane	−3.81	−3.81	−3.12		−3.78	−3.28	−3.52
1-bromonaphthalene	−4.35	−4.35	−4.05			−4.53	
1-bromopentane	−3.08	−3.08	−2.82		−3.07	−2.79	−3.25
1-bromopropane	−1.70	−1.70	−1.49			−1.95	−2.01
1-butene	−2.40	−2.39	−1.97		−2.31	−1.72	
1-chloro-1,1-difluoroethane	−1.20	−1.20	−1.14			−1.13	−0.53
1-chloro-2-fluoroethane	−0.51	−0.50	−0.60				−0.52
1-chloro-2-methylpropane	−2.00	−1.99	−1.93				−2.18
1-chlorobutane	−2.03	−2.03	−1.99		−2.20	−1.88	−2.16
1-chlorohexane	−3.12	−3.12	−2.77			−2.80	
1-chloronaphthalene	−3.93	−3.93	−3.45			−4.07	
1-chlorooctane	−4.48	−4.47	−4.62			−3.80	
1-chloropentane	−2.73	−2.72	−2.46			−2.32	−2.82
1-choropropane	−1.46	−1.45	−1.03			−1.44	−1.54
1-decanol	−3.63	−3.63	−3.74			−2.93	
1-decene	−5.51	−6.05	−4.99				−5.17
1-ethylnaphthalene	−4.17	−4.17	−4.65	−3.19			−4.68
1-heptanol	−1.95	−1.94	−1.59	−1.78	−1.72	−1.27	
1-heptyne	−3.01	−3.01	−2.91		−2.94		

Table 5. (Continued)

compound	logS [mol/L] reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
1-hexene	−3.23	−3.23	−2.83		−3.28	−2.70	−2.66
1-hexyne	−2.36	−2.35	−2.20	−2.36	−2.28		
1-iodopropane	−2.29	−2.29	−1.92			−2.03	−2.51
1-methoxy-2-propanol	1.05	1.06	0.89			0.12	
1-methylcyclohexene	−3.27	−3.27	−3.19			−3.08	−3.09
1-methylnaphthalene	−3.70	−3.69	−4.01	−3.60			−4.12
1-nitropropane	−0.77	−0.77	−0.52			−1.19	
1-nonanol	−3.13	−2.87	−2.96		−3.21	−2.38	
1-nonene	−5.05	−5.05	−4.76		−5.24		−4.59
1-nonyne	−4.26	−4.25	−3.94		−4.41		
1-octanol	−2.38	−2.38	−2.24	−2.52		−1.81	
1-octene	−4.44	−4.44	−3.85		−4.59	−3.75	−3.89
1-octyne	−3.61	−3.60	−3.23	−3.93	−3.60		
1-pentanol	−0.60	−0.60	−0.59	−0.42	−0.70	−0.20	
1-pentene	−2.68	−2.67	−2.31		−2.71	−2.21	−2.08
1-pentyne	−1.64	−1.63	−1.70	−1.80	−1.74		
1-propanol	−1.22	−1.22	−1.08			0.90	
1-propene	−2.32	−2.32	−1.88	−1.66	−2.08		
2,2,2-trifluoroethanol	−0.23	−0.23	−0.72			−0.25	
2,2′,3,3′,4,6′-hexachloro-1,1′-biphenyl	−6.96	−6.96	−7.18				
2,2′,3,3′-tetrachloro-1,1′-biphenyl	−7.27	−7.27	−7.30				−6.97
2,2′,3,4,4′,5′-hexachloro-1,1′-biphenyl	−8.32	−8.32	−8.24				
2,2′,4,4′,6,6′-hexabromo-1,1′-biphenyl	−9.00	−9.00	−8.78	−8.11	−8.29		
2,2′,4,4′,6,6′-hexachloro-1,1′-biphenyl	−8.71	−8.32	−8.55	−7.38	−7.69		−8.32
2,2′,4,5,5′-pentachloro-1,1′-biphenyl	−7.44	−7.44	−7.94	−4.55	−4.69		−6.68
2,2,4-trimethylpentane	−4.70	−4.70	−3.60				−4.77
2,2′,5,5′-tetrabromo-1,1′-biphenyl	−8.06	−8.06	−8.60				
2,2,5,5-tetramethyl-3-hexyne	−3.69	−3.69	−4.50				
2,2′,5,6′-tetrachloro-1,1′-biphenyl	−6.79	−6.78	−7.02				−6.78
2,2′,5-trichloro-1,1′-biphenyl	−6.02	−6.02	−6.26	−5.28	−5.26		−6.15
2,2,5-trimethylhexane	−5.05	−5.05	−4.60				−5.38
2,2′,6,6′-tetrachloro-1,1′-biphenyl	−7.39	−7.39	−7.85				−6.88
2,2-dimethylbutane	−3.56	−3.58	−3.35	−3.59	−3.58	−3.33	−3.55
2,2-dimethylpentane	−4.36	−4.60	−4.51	−4.32		−3.86	−4.13
2,2-dimethylpropane	−3.33	−3.33	−2.72		−3.12	−2.80	
2,2-dimethylpropanoic acid	−0.67	−0.67	−0.43			−0.65	
2,3,4-trimethylpentane	−4.70	−4.70	−4.25	−4.18	−4.71	−4.42	−4.68
2,3,5,6-tetrachloro-1,1′biphenyl	−7.32	−7.39	−7.95				−6.94
2,3-benzofluorene	−8.04	−8.04	−7.88	−7.56			
2,3-dichlorobutane	−2.70	−2.69	−2.74			−2.17	
2,3-dichloropropene	−1.71	−1.70	−2.08			−1.41	
2,3-dimethyl-1,3-butadiene	−2.40	−2.39	−3.00				−2.31
2,3-dimethyl-2-butanol	−0.37	−0.37	−0.58			−0.74	
2,3-dimethylbutane	−3.58	−3.58	−3.67	−3.66	−0.36	−3.34	−3.48
2,3-dimethylnaphthalene	−4.72	−4.72	−4.27	−4.47			−4.70
2,3-dimethylpentane	−4.28	−4.28	−3.98	−4.17		−3.86	−4.17
2,3-dimethylpyridine	0.38	0.40	−0.29			−0.18	
2,4,4′-trichloro-1,1′-biphenyl	−6.21	−6.21	−6.43	−6.34			−6.01
2,4,6-trichloro-1,1′-biphenyl	−6.14	−6.21	−6.38	−5.99	−6.14		−6.01
2,4′-dichloro-1,1′-biphenyl	−5.28	−5.28	−5.56	−5.40			−5.45
2,4-dichloro-1,1′-biphenyl	−5.25	−5.21	−5.83	−5.46			−5.44
2,4-dimethyl-3-pentanone	−1.30	−1.29	−1.24			−1.85	
2,4-dimethylpentane	−4.26	−4.60	−3.98	−4.27	−4.21	−3.87	−4.18
2,4-dimethylpyridine	0.51	0.51	−0.69			−0.14	
2,5-cyclohexadiene-1,4-dione	−0.99	−0.98	−1.06			−0.58	
2,5-dichloro-1,1′-biphenyl	−5.36	−5.36	−5.22	−5.36	−5.47		−5.50
2,5-dimethylpyridine	0.40	0.40	−0.68			−0.17	
2,6,dimethylnaphthalene	−4.89	−4.89	−4.97	−4.65			−4.70
2,6-dichloro-1,1′-biphenyl	−5.21	−5.21	−5.69	−5.18	−5.36		−5.35
2,6-dichlorophenol	−1.79	−1.79	−1.54			−1.28	
2,6-diethylaniline	−2.35	−2.35	−2.85			−1.73	
2,6-dimethyl-4-heptanone	−1.73	−1.73	−2.08			−2.87	
2-bromonaphthalene	−4.40	−4.35	−4.04				
2-bromopropane	−1.59	−1.59	−1.44			−1.93	−1.93
2-butanol	0.39	0.39	−0.28			0.24	
2-butanone	0.49	0.49	−0.11	−0.55	−0.17	−0.05	
2-chloro-1-nitrobenzene	−2.55	−2.53	−2.62			−2.22	
2-chloro-2-methylpropane	−2.20	−2.20	−1.97				−2.20
2-chlorobutane	−1.96	−1.96	−1.87			−1.77	
2-chloroethanol	1.09	1.09	0.81			1.03	
2-chloronaphthalene	−4.14	−3.93	−3.38				

Table 5. (Continued)

compound	logS [mol/L] reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
2-chloropentane	−2.63	−2.72	−2.52			−2.23	−2.84
2-chloropropane	−1.41	−1.41	−1.25			−1.39	−1.54
2-chloropyridine	−0.75	−0.74	−0.89			−1.17	
2-ethyl-1,3-hexandiol	−2.81	−2.81	−2.42		−2.75		
2-ethylbutyric acid	−0.81	−0.80	−0.98			−1.17	
2-ethylnaphthalene	−4.29	−4.17	−4.59				−4.69
2-ethylpyridine	0.50	0.51	−0.82			−0.16	
2-fluorochlorobenzene	−2.42	−2.42	−2.66				−2.38
2-heptanone	−1.45	−1.45	−1.32		−1.32	−1.69	
2-heptene	−3.82	−3.81	−3.84		−4.08		−3.50
2-hexanol	−1.42	−1.41	−0.96			−0.70	
2-hexanone	−0.79	−0.79	−0.84			−1.12	
2-hexene	−3.10	−3.23	−3.28				−2.82
2-iodopropane	−2.09	−2.09	−1.84			−2.06	−2.47
2-methyl-1,3-butadiene	−2.03	−2.09	−1.93		−1.90		−1.68
2-methyl-1-butanol	−0.47	−0.47	−0.61	−0.05		−0.09	
2-methyl-1-pentene	−3.03	−3.03	−2.81		−3.16		−2.72
2-methyl-2-butanol	0.10	0.10	−0.31			−0.20	
2-methyl-2-butene	−2.56	−2.56	−2.21			−2.23	−2.21
2-methyl-2-pentanol	−0.50	−0.50	−0.82			−0.75	
2-methyl-3-pentanone	−0.81	−0.80	−0.84			−1.16	
2-methylaziridine	1.24	1.24	1.01			1.05	
2-methyldecalin	−6.57	−6.56	−6.16				−5.96
2-methylhexane	−4.60	−4.60	−3.94	−4.54		−3.86	−4.23
2-methylnaphthalene	−3.76	−3.69	−3.97	−3.75		−4.61	−4.22
2-methylpentane	−3.74	−3.74	−3.63	−3.88	−3.69	−3.32	−3.59
2-methylpyridine	−1.03	−1.03	−1.16			0.22	
2-nitropropane	−1.75	−1.74	−1.49			−1.20	
2-nonanone	−2.57	−2.56	−2.49		−2.50	−2.91	
2-octanone	−2.05	−2.35	−1.86			−2.23	
2-pentanol	−0.30	−0.30	−0.50			−0.20	
2-pentanone	−0.30	−0.30	−0.46			−0.77	
3,3′,4,4′-tetrachloro-1,1′-biphenyl	−8.73	−8.73	−8.94	−8.29			−6.83
3,3′,5,5′-tetrachloro-1,1′-biphenyl	−8.37	−8.73	−7.83				−6.68
3,3′,5-trichloro-1,1′-biphenyl	−6.01	−6.02	−6.55				−6.02
3,3-dimethylpentane	−4.23	−4.23	−3.99	−4.16		−3.87	−4.14
3,4-dichloro-1,1-biphenyl	−6.39	−6.39	−5.70				
3,4-dichlorotoluene	−3.79	−3.79	−2.85			−3.02	
3,4-dimethylpyridine	0.36	0.36	−0.52			−0.26	
3,5-dimethylpyridine	0.38	0.39	−0.62			−0.21	
3,3-dimethyl-2-butanone	−0.72	−0.72	−0.75	−0.81		−1.12	
3-bromo-1-propene	−1.50	−1.50	−1.21				−1.28
3-chloro-1-propene	−1.28	−1.09	−0.95				−2.23
3-chloroiodobenzene	−3.55	−3.55	−3.33				−3.41
3-chlorophenol	−0.69	−0.69	−0.82			−0.63	
3-heptanol	−1.46	−1.45	−1.51			−1.32	
3-heptanone	−1.42	−1.41	−1.34			−1.88	
3-hexanone	−0.83	−0.82	−0.87			−1.19	
3-methyl-1-butene	−2.73	−2.72	−2.39	−2.78			−2.11
3-methyl-2-butanol	−0.16	−0.15	−0.32			−2.21	
3-methyl-2-butanone	−0.15	−0.15	−0.38			−0.29	
3-methyl-3-pentanol	−0.38	−0.37	−0.53			−0.74	
3-methylcholanthrene	−7.92	−7.92	−8.08	−8.02			
3-methylheptane	−5.16	−5.16	−5.36	−5.15			−4.80
3-methylhexane	−4.58	−4.60	−4.81	−4.42	−4.30	−3.86	−4.16
3-methylpentane	−3.84	−3.58	−3.55	−3.79	−3.69	−3.83	−3.56
3-methylpyridine	−1.03	−1.03	−1.17			0.16	
3-methylthiophene	−2.39	−2.39	−1.95			−2.28	
3-pentanol	−0.23	−0.23	−0.33			−0.25	
3-pentanone	−0.28	−0.28	−0.45	−0.27	−0.63	−0.61	
4,4′-dichloro-1,1′-biphenyl	−6.56	−6.56	−6.62	−6.50			−5.44
4-bromo-1-butene	−2.25	−2.25	−1.64		−1.79		−1.76
4-bromoiodobenzene	−4.56	−4.56	−4.22				−3.62
4-chloro-2-nitrophenol	−3.09	−2.51	−2.96			−1.44	
4-chloroaniline	−1.51	−1.50	−1.46			−0.42	
4-chlorophenol	−0.73	−0.72	−1.04			−0.59	
4-ethenylcyclohexene	−3.34	−3.33	−3.47		−3.38		
4-methyl-1-pentene	−3.24	−3.23	−3.46		−3.22	−2.71	−2.71
4-methyl-2-pentanone	−0.72	−0.80	−0.83			−1.26	
4-methyl-2-pentylacetate	−2.05	−2.05	−1.90			−2.43	
4-methyloctane	−6.05	−6.05	−5.81	−5.79			−5.46

Table 5. (Continued)

	logS [mol/L]						
compound	reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
4-methylpyridine	−0.10	−0.09	−0.35			0.19	
5-methyl-2-hexanone	−1.33	−1.45	−1.27			−1.68	
6-chloro-10-methylbenz[*a*]anthracene	−8.70	−8.70	−7.95				
9,10-dimethylanthracene	−6.57	−6.56	−6.79	−5.98			
a,a,a-trifluorotoluene	−2.51	−2.51	−2.09				−1.92
acenaphthene	−4.63	−4.62	−4.24	−4.78			
acetaldehyde	1.36	1.36	1.39			0.75	
acetamide	1.08	1.09	1.19			2.08	
acetic acid	−1.22	−1.22	−1.41			1.21	
acetonitrile	1.39	1.39	1.37			2.26	
acetophenone	−1.29	−1.29	−1.13	−1.64		−1.63	
a-chlorotoluene	−2.43	−2.43	−1.83				−2.54
allyl alcohol (2-propen-1-ol)	0.74	0.74	0.80			1.32	
amyl propionate (pentyl)	−2.25	−2.45	−2.14			−2.45	
aniline	−0.41	−0.41	−0.35	−0.59		−0.09	
anthracene	−6.35	−6.35	−5.88	−5.95			−5.50
benz[*a*]anthracene	−7.21	−7.21	−7.35	−7.57			−6.96
benzaldehyde	−3.24	−3.23	−2.70	−1.12		−1.57	
benzene	−1.64	−1.63	−1.69	−1.85	−1.79	−2.91	−2.15
benzo[*a*]pyrene	−7.80	−7.80	−7.96	−8.26			−7.74
benzo[*b*]fluoranthene	−8.23	−8.49	−7.88				
benzo[*e*]pyrene	−7.60	−7.60	−7.90				
benzo[ghi]perylene	−9.03	−9.03	−8.29	−8.92			
benzo[*j*]fluoranthene	−8.00	−8.00	−7.81				
benzo[*k*]fluoranthene	−8.49	−8.49	−7.83				
benzoic acid	−1.81	−1.81	−1.78	−1.07		−1.85	
benzonitrile	−1.01	−1.00	−0.82			−0.49	
benzyl acetate	−1.69	−1.69	−1.65			−5.55	
biphenyl	−4.31	−4.31	−4.90		−3.96	−2.97	−4.59
bromochloromethane	−0.89	−0.89	−1.07		−1.05		−0.91
bromoethane	−1.09	−1.09	−1.24			−1.61	−1.39
bromomethane	−0.80	−0.80	−1.15	−1.02		−1.50	
butanol	−0.07	−0.06	−0.22	0.11	−0.69		
butyl acetate	−1.27	−1.27	−0.90	−0.74		−1.35	
butylamine	1.14	1.15	−1.27			0.41	
butylbenzene	−4.06	−3.96	−4.00	−3.82	−3.85	−4.09	−4.13
butyraldehyde	−0.01	−0.01	−0.14			−0.32	
chlorobenzene	−2.44	−2.43	−2.33		−2.41	−2.53	−2.44
chlorodifluoromethane	−1.47	−1.47	−0.92			−0.98	
chloroethane	−1.06	−1.05	−0.95			−1.05	
chloromethane	−0.98	−0.98	−0.76			−0.87	
chrysene	−8.06	−8.06	−7.32	−7.73			
cis-1−2-dichloroethene	−1.44	−1.44	−1.56				−1.35
cis-1−2-dimethylcyclohexane	−4.44	−4.47	−4.37				−4.40
coronene	−9.33	−9.33	−8.73	−9.32			
cumene	−3.29	−3.10	−3.47	−3.24		−3.66	
cycloheptane	−3.51	−3.51	−3.98		−3.63	−3.73	−3.83
cycloheptene	−3.16	−3.27	−3.19		−3.25		−3.11
cyclohexane	−3.18	−3.18	−3.48	−3.03	−3.10	−3.19	−3.25
cyclohexanol	−0.43	−0.43	−0.64			−0.56	
cyclohexanone	−0.63	−0.62	−0.98			−1.02	
cyclohexene	−2.59	−2.22	−2.79	−2.48	−2.67	−2.57	−2.45
cyclooctane	−4.15	−4.15	−4.55		−4.15		−4.50
cyclopentane	−2.59	−2.58	−2.91	−2.69	−2.48	−2.66	−2.60
cyclopentanone	−0.96	−0.96	−0.80			−0.51	
cyclopentene	−2.10	−2.09	−2.23	−2.02	−2.01	−2.07	−1.83
cyclopropane	−2.04	−2.03	−1.87			−1.66	
decachlorobiphenyl	−11.62	−11.62	−9.82	−10.60			
decalin (decahydronaphthalene)	−5.19	−5.18	−5.66			−5.19	−5.31
decanal	−3.41	−3.40	−3.28			−3.50	
decane	−6.44	−6.08	−6.58	−6.65		−5.49	−6.12
di(*n*-butyl) ether	−2.64	−2.64	−2.55	−2.01		−2.27	
di(*n*-propyl) ether	−1.32	−1.32	−1.30			−1.20	
dibromomethane	−1.16	−1.15	−1.26			−1.65	−1.67
dichlorodifluoromethane	−2.64	−2.64	−2.49	−2.18		−2.30	
dichlorofluoromethane	−1.03	−1.03	−1.13			−0.87	
dichloromethane	−0.82	−0.82	−1.02	−0.75		−1.20	−0.62
diethyl sulfide	−1.25	−1.24	−1.34			−1.54	
diethylamine	1.14	1.15	1.26			0.28	
diethyl ether	−0.09	−0.09	−0.26	−0.34		−0.15	
diisobutylamine	−1.77	−1.77	−2.04			−2.29	

Table 5. (Continued)

compound	logS [mol/L] reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
diisopropylamine	0.04	0.04	0.01			−0.91	
diisopropyl ether	−1.06	−1.05	−0.93			−1.15	
dimethylamine	1.56	1.56	1.38			1.53	
dimethyl ether	0.55	0.55	0.44			0.70	
dimethyl sulfide	−0.74	−0.74	−0.83			−0.49	
di-*n*-butylamine	−1.57	−1.57	−1.65			−1.97	
di-*n*-propylamine	−0.24	−0.23	−1.35			−0.86	
diphenylmethane	−4.08	−4.62	−5.66	−4.87			−5.24
dodecane	−7.67	−7.67	−6.94				−6.74
ethane	−2.70	−2.69	−2.82	−2.82	−2.75	−1.27	
ethanol	1.34	1.34	1.35			1.36	
ethene	−2.33	−2.32	−2.23		−2.51	−1.27	
ethyl acetate	−0.04	−0.04	−0.10	0.04	−0.34	−0.30	
ethyl benzene	−1.36	−1.35	−1.84	−2.54	−2.67	−3.21	
ethyl benzoate	−2.32	−2.32	−2.13			−2.12	
ethyl formate	0.08	0.08	−0.13			0.11	
ethyl mercaptan	−0.73	−0.72	−0.67			−0.54	
ethylbutyrate	−1.37	−1.37	−1.56			−1.36	
ethylcyclohexane	−4.25	−4.25	−4.34			−4.21	
fluorene	−5.00	−5.00	−4.85	−5.37			−4.82
formaldehyde	1.12	1.13	0.93			1.12	
formic acid	−1.34	−1.34	−1.31			1.52	
furan	−0.83	−0.82	−0.75			−0.45	
heptanal	−1.96	−1.96	−1.73			−1.88	
heptanoic acid	−1.66	−1.66	−1.71			−1.55	
hexachlorobenzene	−6.78	−6.78	−6.88		−6.44		−6.27
hexachlorocyclopentadiene	−5.18	−5.18	−5.33			−5.43	
hexachloroethane	−4.49	−4.49	−4.54				−4.77
hexadecane	−8.40	−8.40	−7.67				−9.95
hexamethylbenzene	−5.23	−5.23	−5.27				−5.48
hexanal	−1.22	−1.22	−0.89			−1.36	
hexane	−3.96	−3.96	−3.72	−4.06	−3.84		−3.63
hexanoic acid	−1.05	−1.05	−0.95			−0.93	
indan	−3.13	−3.13	−2.95				−3.31
iodobenzene	−2.95	−2.95	−2.84	−2.57	−2.87	−3.25	−2.99
iodoethane	−1.60	−1.59	−1.68			−1.52	−1.90
iodomethane	−1.00	−1.00	−1.14			−1.16	−1.10
isoamyl acetate	−1.81	−1.81	−1.73			−1.89	
isoamyl formate	−1.52	−1.52	−1.38			−1.43	
isobutane	−3.08	−3.08	−2.82	−2.78			
isobutanol	0.01	0.01	−0.18	0.47		0.46	
isobutene	−2.33	−2.39	−2.20			−1.75	
isobutyl acetate	−1.24	−1.24	−0.95	−1.04			
isobutylamine	1.14	1.15	1.24			0.47	
isobutyl formate	−1.00	−1.13	−0.86			−0.86	
isobutylbenzene	−4.12	−4.12	−3.87			−4.10	
isobutyraldehyde	0.09	0.10	−0.11			−0.31	
isophorone	−0.55	−0.55	−0.82			−2.34	
isopropyl alcohol	1.22	1.23	1.16			0.86	
isopropyl formate	−0.63	−0.62	−0.47			−0.36	
isopropylamine	1.23	1.18	1.10			0.92	
m-bromotoluene	−3.52	−3.51	−3.27			−3.05	
m-chlorotoluene	−3.52	−3.51	−3.70			−2.60	
m-cresol (3-hydroxytoluene)	−0.68	−0.69	−0.95	−0.70		−0.44	
methanol	1.49	1.49	1.61			1.85	
methycyclopentane	−3.30	−3.18	−3.20		−3.14		−3.18
methyl acetate	0.52	0.52	0.26			0.13	
methylamine	1.54	1.54	1.60			2.04	
methyl butyl ether	−0.99	−0.98	−0.65			−0.72	
methyl butyrate	−0.83	−0.82	−0.72			−0.87	
methyl formate	0.58	0.58	0.32			0.52	
methyl hexanoate	−1.99	−1.99	−1.74			−1.92	
methyl isopropyl ether	−0.06	−0.06	−0.18			−0.18	
methyl mercaptane	−0.49	−0.49	−0.55			−0.17	
methyl propanoate	−0.15	−0.15	−0.07			−0.36	
methyl propyl ether	−0.39	−0.39	0.22			−0.21	
methylcyclohexane	−3.85	−3.84	−3.83	−3.59	−3.77		−3.85
m-xylene	−2.80	−2.81	−2.79	−2.99	−2.74	−3.27	
n-amyl acetate	−1.88	−1.88	−1.41			−1.89	
naphthalene	−3.60	−3.60	−4.11	−3.48		−4.64	−3.74
n-butane	−2.98	−3.08	−2.82	−2.65	−2.85	−2.27	

Table 5.

	logS [mol/L]						
compound	reported	fuzzy ARTMAP	back-propagation	Mitchell & Jurs (1998)	Sutter & Jurs (1996)	Katritzky et al. (1998)	Huibers & Katritzky (1998)
n-butyl propionate	−1.94	−1.94	−1.83			−1.91	
n-butyl acetate	−1.27	−1.27	−0.90		−1.14	−1.35	
n-butylformate	−1.13	−1.13	−0.83			−0.90	
n-ethylaniline	−1.70	−1.70	−1.57			−0.82	
n-heptane	−4.53	−4.53	−4.20	−4.70	−4.44	−3.87	−4.16
n-hexyl acetate	−2.45	−2.45	−2.00			−2.42	
n-hexylamine	−0.69	−0.69	−0.82			−0.70	
n-hexylbenzene	−5.21	−5.21	−5.18		−5.20		−5.48
nitrobenzene	−1.88	−1.88	−1.94	−1.66		−2.18	
nitromethane	−0.26	−0.26	0.08	−0.78		−0.33	
N-methylaniline	−1.28	−1.27	−1.00			−0.31	
N,*N*-dimethyl formamide	1.14	1.15	1.04			1.39	
N,*N*-dimethylaniline	−1.79	−1.70	−1.33			−0.64	
nonane	−5.88	−6.05	−5.73	−5.98		−4.95	−5.49
n-pentylamine	1.06	1.06	−0.98			−0.13	
n-pentylbenzene	−4.59	−4.15	−4.59	−4.52	−4.54		−4.85
n-pentylcyclopentane	−6.08	−6.08	−5.44	−6.26			−5.74
n-propyl acetate	−0.73	−0.72	−0.47		−0.69	−0.84	
n-propylbenzene	−3.36	−3.36	−3.36		−3.25	−3.63	
n-propylcyclopentane	−7.74	−7.74	−7.23	−4.43			−4.44
o-cresol (2-hydroxytoluene)	−0.62	−0.62	−0.64	−0.65		−0.63	
octanal	−2.36	−2.35	−2.57			−2.40	
octane	−5.11	−5.11	−4.67	−5.39	−5.05	−4.40	−4.85
octanoic acid	−2.26	−2.25	−2.09			−2.07	
p-chloronitrobenzene	−2.54	−2.53	−2.61			−2.21	
pentachlorobenzene	−5.65	−5.65	−5.38	−5.87	−5.51		−5.35
pentane	−3.28	−3.27	−3.21	−3.52	−3.26	−2.79	−2.91
pentanoic acid	−1.29	−1.29	−1.45	−0.01		−0.52	
perylene	−8.80	−8.80	−7.95	−8.60			
p-ethyltoluene	−3.10	−3.10	−3.05			−3.65	
phenanthrene	−5.26	−5.26	−5.71	−5.79			−5.27
phenol	−0.06	−0.06	−0.14	−0.47		−0.28	
propane	−2.85	−2.85	−2.44	−2.65	−2.68	−1.77	
propionaldehyde	0.72	0.72	0.65			0.24	
propionic acid	1.13	1.13	0.91			0.36	
propyl formate	−0.60	−0.60	−0.46			−0.39	
propylene oxide	1.01	1.01	0.98			0.10	
propyne	−1.04	−1.03	−1.17		−1.06		
p-xylene	−2.82	−2.81	−3.02	−2.85	−2.78	−3.26	
pyridine	1.10	1.10	1.21			−0.43	
pyrrole	−0.15	−0.15	−0.21			1.05	
pyrrolidine	1.15	1.15	1.24			0.56	
quinoline	−1.33	−1.32	−1.55			−2.12	
sec-butyl acetate	−1.27	−1.27	−1.19			−1.37	
sec-butyl-benzene	−3.93	−3.93	−4.07	−3.70			−4.20
styrene	−2.53	−2.53	−2.38	−2.14		−3.75	−2.89
tert-butyl alcohol	1.13	1.13	−0.06			0.30	
tert-butyl amine	1.14	1.13	1.10			0.32	
tert-butyl-benzene	−3.66	−3.66	−3.67	−3.71		−4.13	−4.13
tetrachloroethene	−2.54	−2.53	−2.79	−2.81		−3.54	−3.31
tetrachloromethane	−2.31	−2.31	−2.32	−2.24			−2.93
tetrafluoroethene	−1.60	−1.59	−1.49				−1.27
tetrafluoromethane	−3.67	−3.66	−3.96			−3.16	
thiophene	−1.44	−1.44	−1.63	−1.42		−2.18	
thiophenol	−2.12	−2.12	−2.29			−2.14	
toluene	−2.22	−2.22	−2.23	−2.20	−2.19	−2.90	−2.89
trans-1,4-dimethylcyclohexane	−4.47	−4.47	−4.35				−4.46
trans-2-pentene	−2.54	−2.53	−2.75		−2.90	−2.21	−2.22
tribromomethane	−1.91	−1.91	−1.63				−2.61
trichloroethene	−1.96	−1.96	−2.15	−1.61	−2.12	−2.16	−2.32
trichlorofluoromethane	−2.10	−1.88	−1.88			−2.32	
trichloromethane	−1.17	−1.17	−1.28	−1.20		−2.15	−1.51
trichloronitromethane	−2.01	−2.01	−2.19			−2.56	
trimethylamine	1.18	1.18	1.11			1.00	
count		418	418	129	95	287	172
average absolute error, logS	0.03	0.27	0.38	0.23	0.47	0.34	
maximum absolute error, logS	0.75	2.41	5.18	3.22	3.86	3.30	

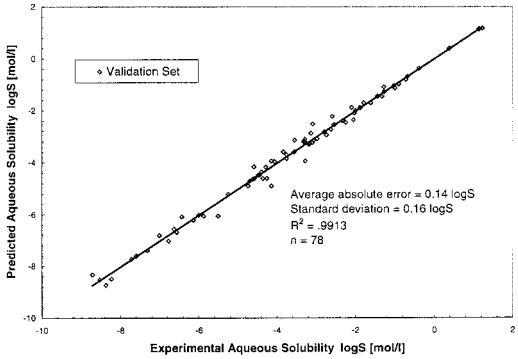

Figure 3. Aqueous solubility results for test set for fuzzy ARTMAP.

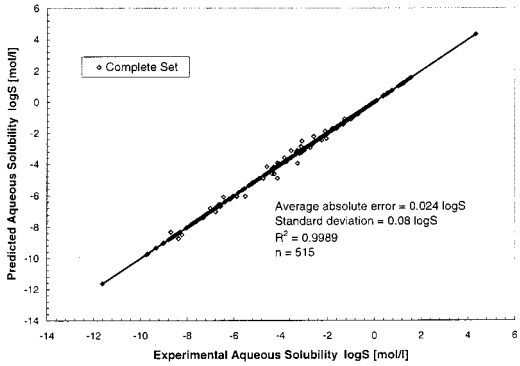

Figure 4. Aqueous solubility results for complete set for fuzzy ARTMAP.

diverse set of 332 compounds ($-12 \leq \log S \leq 2$). A multilinear regression model with six molecular descriptors developed by Katritzky et al.,[16] based on a diverse set of 411 compounds ($-6.44 \leq \log S \leq 1.57$) performed with standard errors of 0.573 logS units. It is noted that training errors are expected to be lower for models that are based on a more homogeneous set of compounds (i.e., containing similar classes). For example, Sutter and Jurs[12] reported an absolute error of 0.222 logS units, based on a 9−3−1 neural network based logS model developed using a less diverse set of 140 compounds ($-10.83 \leq \log S \leq 0.28$), with an even lower error of 0.197 logS units when polychlorinated biphenyls (PCB) were removed from the data set.[12]

Specific performance comparisons of the present models to previously published neural network-based QSPRs for aqueous solubility at 25 °C,[12,13,15,16] based on a sets of compounds common with the present study, are provided in Table 5 and Figure 7. For a set of 95 compounds common with the model of Sutter and Jurs,[12] the present back-propagation model yielded average absolute and maximum solubility estimation errors of 0.24 and 0.69 logS units, respectively, relative to corresponding errors of 0.23 and 3.2 logS units for the Sutter and Jurs[12] model. The present fuzzy ARTMAP QSPR on the other hand, for the same 95 compounds, was superior with a significantly lower average absolute and maximum aqueous solubility estimation errors of 0.035 and 0.64 logS units, respectively. Application of the present back-propagation model to a set of 129 compounds common with the set of the Mitchell and Jurs[13] network/QSPR model, yielded average and maximum absolute errors of 0.28 and 1.8 logS units, respectively, relative to higher corresponding errors of 0.34 and 3.3 logS units, respectively, for the Mitchell and Jurs[13] model. Performance

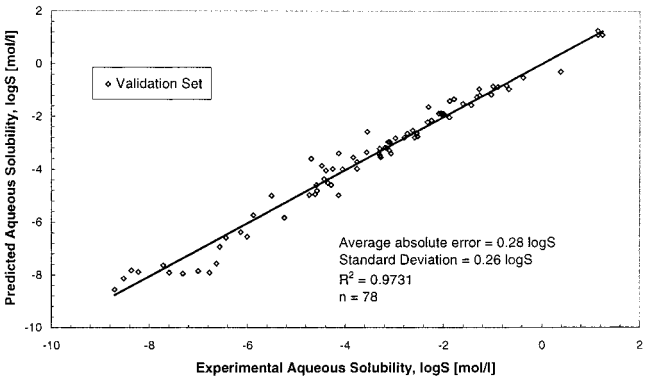

Figure 5. Aqueous solubility results for validation set for back-propagation.

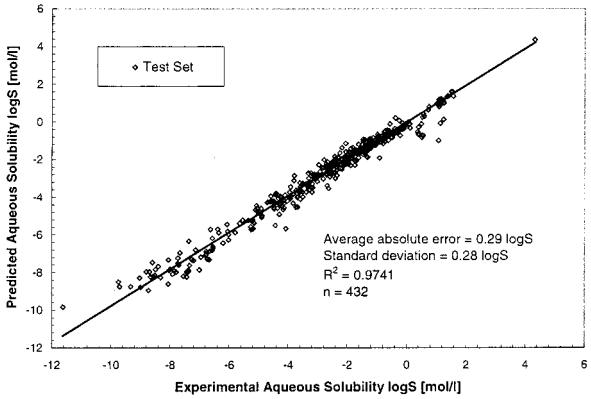

Figure 6. Aqueous solubility results for test set for back-propagation.

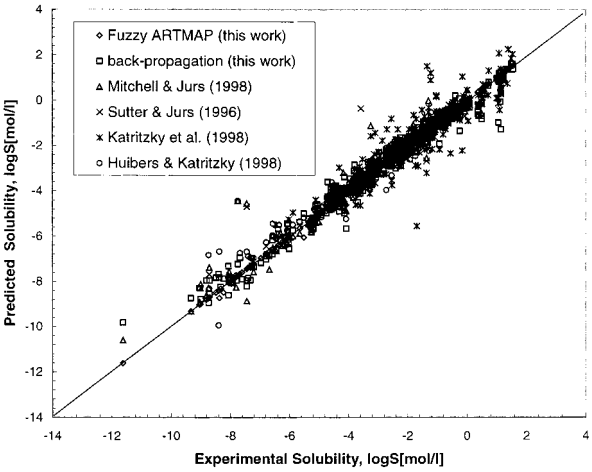

Figure 7. Comparison of aqueous solubility QSPR prediction methods for organics.

of the fuzzy ARTMAP, for the same 129 compounds, was superior, with average and maximum absolute errors of 0.037 and 0.75 logS units, respectively.

The present models were also compared to recent multilinear regression QSPRs developed by Katritzky et al.[16] and Huibers and Katritzky.[15] For the 287 compounds common with the study of Katritzky et al.,[16] the present back-propagation QSPR model performed with average and maximum absolute errors of 0.25 and 2.4 logS units, respectively, relative to the corresponding errors of 0.47 and 3.8 logS units for the Katritzky et al.[16] model. The present fuzzy ARTMAP was a significant improvement with average

absolute and maximum absolute errors of 0.027 and 0.67 logS units, respectively, for the same 287 compound data set. Finally comparison with the multilinear regression QSPR of Huibers and Katritzky[15] (develop based on a data set of 241 organics), for a set of 172 compounds common with the present study, revealed average and maximum absolute errors of 0.27 and 1.6 and 0.039 and 0.75 logS units, respectively, for the present back-propagation and fuzzy ARTMAP models, respectively, relative to the corresponding errors of 0.34 and 3.3 logS units for the Huibers and Katritzky[15] model. It must be emphasized that model performance is closely linked to the range of chemical classes and abundance of chemicals, within a class, which are presented to the system during training. Therefore, one must view performance comparisons with caution as models developed with heterogeneous data sets, especially where the number of compounds within a class is limited, are not likely to perform with a lower overall accuracy (when compared with a common data set), relative to models developed with a narrower distribution of chemical classes.

In closure, the performance of the fuzzy ARTMAP/QSPR was superior relative to the present back-propagation/QSPR, other published neural network based models, and multilinear regression QSPRs when compared for the same compound classes included in the development of the present models. The overall performance of the presented models suggests that the present set of quantum chemical parameters were satisfactory in characterizing molecular structure and differentiating between chemical groups. Notwithstanding, the presence of some outliers suggests that refinement of the set of chemical descriptors is warranted to improve the accuracy and applicability of the approach.

CONCLUSIONS

The applicability of fuzzy ARTMAP network for developing a QSPR for aqueous solubility was demonstrated using a set of chemical descriptors that included parameters obtained from PM3 semiempirical MO-theory calculations. The set of descriptors obtained from PM3 semiempirical MO-theory calculations represented different forms of three-dimensional information for characterizing the various atoms and chemical groups for a set of heterogeneous organic compounds. The fuzzy ARTMAP based model, and for comparison, a companion back-propagation neural network model were developed based on a heterogeneous set of more than 500 organic compounds. For the fuzzy ARTMAP-based QSPR, average absolute errors in aqueous solubility estimations for the overall and validation sets were 0.024 and 0.14 logS units, respectively. In contrast, the 11−13−1 back-propagation/QSPR model performed with average absolute errors for the overall and validation sets of 0.29 and 0.28 logS units, respectively. The present fuzzy ARTMAP neural network-based QSPR model was also shown to be superior to other neural network/QSPR and multilinear regression/QSPR aqueous solubility models reported in the literature.

The results obtained with the fuzzy ART classifier and the fuzzy ARTMAP cognitive system demonstrate that it is possible to establish a reasonably accurate quantitative structure−property relationships for heterogeneous compounds based on a set of descriptors calculated from quantum mechanics and graph theory. Further optimization and expansion of the set of chemical descriptors is currently underway to better classify isomers and other complex chemical structures.

REFERENCES AND NOTES

(1) Reinhard, M.; Drefahl, A. *Handbook for Estimating Physicochemical Properties of Organic Compounds*; John Wiley & Sons. Inc: New York, 1999.
(2) Wakita, K.; Yoshimota; Miyamoto; Watanable A Method for Calculation of the Aqueous Solubility of Organic Compounds at 25 °C. *Chem. Pharm, Bull*. **1986**, *34*, 4663−4681.
(3) Suzuki, T. Development of an Automatic Estimation System for Both the Partition Coefficient and Aqueous Solubility. *J Comput.-Aid. Mol. Design* **1991**, *5*, 149−166.
(4) Klopman, G.; Wang, S,; Balthasar, D. M. Estimation of Aqueous Solubility of Organic Molecules by the Group Contribution Approach: Application to the Study Biodegradation *J. Chem. Inf. Comput. Sci*. **1992**, *32*, 474−4.
(5) Kühne, R.; Ebert, R. U.; Kleint, F.; Schmidt, G.; Schüürmann, G. Group Contribution Methods to Estimate Water Solubility of Organic Chemicals. *Chemosphere* **1995**, *30*, 11, 2061−2077.
(6) Yalkowsky, S. H.; Valvani, S. C. Solubility of Nonelectrolytes in Water. *J. Pharm. Sci.* **1980**, *69*, 8, 912−922.
(7) Isnard, P.; Lambert, S. Aqueous Solubility and *n*-Octanol/Water Partition Coefficient Correlations. *Chemosphere* **1989**, *18*, 1837−1853.
(8) Warne, M.; Connell, D. W.; Hawker, D. W,; Schhhrmann, G. Prediction of Aqueous Solubility and the Octanol−Water Partition Coefficient for Lipophilic Organic Compounds Using Molecular Descriptors and Physicochemical Properties. *Chemosphere* **1990**, *21*, 7, 877−888.
(9) Miller, M. M.; Ghodbane, S.; Waski, S. P.; Tewari, T. B.; Nartire, D. E. Aqueous Solubilities, Octanol/Water Partition Coefficients, and Entropies of Melting of Chlorinated Benzenes and Biphenyls. *J. Chem. Eng. Data* **1984**, *29*, 184−190.
(10) Yaws, C. L.; Pan, X.; Lin, X. Water Solubility Data for 151 Hydrocarbons *Chem Eng*. **1993**, Feb 108−111.
(11) Ruelle, P.; Kesselring, U. Aqueous Solubility Prediction of Environmentally Important Chemical from the Mobile Order Thermodynamics. *Chemosphere* **1997**, *2*, 275−298.
(12) Sutter, J. M.; Jurs, P. C. Prediction of Aqueous Solubility for a Diverse Set of Heteroatom-Containing Organic Compounds Using a Quantitative Structure−Property Relationship. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 100−107.
(13) Mitchell, B., E.; Jurs, P. C. Prediction of Aqueous Solubility of Organic Compounds from Molecular Structure. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 489−496.
(14) Makino, M Prediction of Aqueous Solubility Coefficient of Polychlorinated Biphenyls by use of Computer-Calculated Molecular Properties. *Environ. Intl.* **1998**, *24*, 5/6 653−663.
(15) Huibers, P. D. T.; Katritzky A. R. Correlation of the Aqueous Solubility of Hydrocarbons and Halogenated Hydrocarbons with Molecular Structure. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 283−292.
(16) Katritzky, A. R.; Wang, Y.; Sild, S.; Tamm, T.; Karelson, M. QSPR Studies on Vapor Pressure, Aqueous Solubility, and the Prediction of Water−Air Partition Coefficients. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 720−725.
(17) Medir, M.; Giralt F. Correlation of Activity Coefficients of Hydrocarbons in Water at Infinity Dilution with Molecular Parameters. *AIChE J.* **1982**, *28*, 341.
(18) Nirmalakhandan, N. N.; Speece, R. E. Prediction of Aqueous Solubility of Organic Chemicals Based on Molecular Structure. *Environ Sci. Technol.* **1988**, *22*, 328−338.
(19) Nirmalakhandan, N. N.; Speece, R. E. Prediction of Aqueous Solubility of Organic Chemicals Based on Molecular Structure. 2. Application to PNAs, PCBs, PCDDs, etc. *Environ. Sci. Technol.* **1989**, *23*, 708−713.
(20) Patil, G. S. Correlation of Aqueous Solubility and the Octanol−Water Partition Coefficient Based on Molecular Structure. *Chemosphere* **1991**, *22*, 8, 723−738.
(21) Huuskonen, J.; Salo, M.; Taskinen. Aqueous Solubility Prediction of Drugs Based on Molecular Topology and Neural Network Modeling. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 450−456.
(22) Stanton, D. T.; J. M.; Jurs, P. C. Development and Use of Charged Partial Surface Area Structural Descriptors in Computer-Assisted Quantitative Structure−Property Relationship Studies. *Anal. Chem.* **1990**, *62*, 2323−2329.
(23) Huuskonen, J. J.; Livingston, D. J.; Tetko, I. V. Neural Network Modeling for Estimation of Partition Coefficients Based on Atom-

Type Electrotopological State Indices. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 947−955.
(24) Breindl, A.; Beck, B.; Clark, T. Prediction of the *n*-Octanol/Water Partition Coefficient, logP, using a combination of Semiempirical MO−Calculations and Neural Network. *J. Molecular Modeling* **1997**, *3*, 142−155.
(25) Beck, B.; Breindl, A.; Clark, T. QM/NN QSPR Models with Error Estimations: Vapor Pressure and LogP. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 1046−1051.
(26) Liang, C.; Gallagher, D. A. QSPR Prediction of Vapor Pressure from Solely Theoretically-Derived Descriptors. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 321−324.
(27) McClelland, H. E.; Jurs, P. C. Quantitative Structure−Property Relationships for the Prediction of Vapor Pressures of Organic Compounds from Molecular Structures. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 967−975.
(28) Espinosa, G.; Yaffe, D.; Cohen, Y.; Arenas, A.; Giralt, F.; Neural Network Based Quantitative Structural Property Relations (QSPRs) for Predicting Boiling Points of Aliphatic Hydrocarbons. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 859−879.
(29) Randi×a6, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(30) Raymond, J.; Rogers, T. Molecular Structure Disassembly Program (MOSDAP): A Chemical Information to Automate Structure-Based Physical Property Estimation. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 463−474.
(31) Wikel, H. J. The Use of Neural Networks for Variable Selection in QSAR. *Bioorg. Med. Chem. Lett*. **1993**, *3*, 645−651.
(32) Randic, M.; Dobrowolski, J. Optimal Molecular Connectivity Descriptors for Nitrogen-Containing Molecules. *Intl. J. Quant. Chem.* **1998**, *70*, 1209−1215.
(33) Basak, S. C.; Gute, B. D.; Grunwald, G. D. Use of Topostructural, Topochemical, and Geometric Parameters in the Prediction of Vapor Pressure: A Hierarchical QSAR Approach. *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 651−655.
(34) Carpenter, A.; Grossberg, S. A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics, Image Process.* **1987**, *37*, 54−115.
(35) Carpenter, A.; Grossberg, S. The ART of Adaptive Pattern Recognition by a Self-organizing Neural Network. *Computer* **1988**, 77−88.
(36) Carpenter, G. A.; Grossberg, S.; Marcuzon, N.; Reynolds, J. H.; Rosen, D. B. Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analogue Multidimensional Maps. *IEEE Trans. Neural Networks* **1992**, *3*, 698−713.
(37) Carpenter, G. A.; Grossberg, S.; Marcuzon, N.; Rosen, D. B. Fuzzy ART: Fast Stable Learning and Categorization of Analogue Patterns by an Adaptive Resonance System. *Neural Networks* **1991**, *4*, 759−771.
(38) Carpenter, G.; Grossberg, S. A Self-Organizing Neural Network for Supervised Learning, Recognition, and Prediction. *IEEE Commun. Mag.* **1992**, 38−49.
(39) Giralt, F.; Arenas, A.; Ferre-Giné, J.; Rallo R. The Simulation and Interpretation of Turbulence with a Cognitive Neural System. *Phys. Fluids* **2000**, *12*, 1826.
(40) Gao, G.; Govind, R.; Tabak, H. Predicting Soil Sorption Coefficients of Organic Chemical using a Neural Network Model. *Environ. Toxicol. Chem.* **1996**, *15*, 7, 1089−1096.
(41) *Molecular Modeling Pro.− Revision 3.14*; ChemSM Inc.: 1998.
(42) *CAChe Version 3.2*; CAChe Chemistry Products, Oxford Molecular Ltd.
(43) Stewart, J. J. P.; MOPAC 6.0., Quantum Chemistry Program Exchange No. 455, Bloomington, IN, 1989.
(44) Stewart, J. J. P. Optimization of Parameter for Semiempirical Methods I Method *J. Comput Chem*. **1989**, *10*, 209−220.
(45) McWeeny, R. *Methods of Molecular Quantum Mechanics*; Academic Press: New York, 1989.
(46) Karelson, M.; Lobanov, V. S.; Katritzky, A. R. Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chem Rev*. **1996**, *96*, 102−1043.
(47) Gruber, C.; Buss, V. *Chemosphere* **1989**, *19*, 1595.
(48) Bodor, N.; Gabanyi, Z.; Wong, C. K. A New Method for the Estimation of Partition Coefficients. *J. Am. Chem. Soc*. **1989**, *111*, 3783−3786.
(49) Sklenar, H.; Jäger, J. Molecular Structure-Biological Activity Relationships on the Basis of Quantum-Chemical Calculations. *Intl. J. Quant. Chem.* **1979**, *16*, 467−484.
(50) Kier, L. B.; Hall, L. H.. *Molecular Connectivity in Chemistry and Drug Research;* Academic Press: New York, 1976.
(51) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis;* John Wiley & Sons Inc.: New York, 1985.
(52) Kier, L. B. A Shape Index from Molecular Graphs. *Quant. Struct-Act. Relat.* **1985**, *4*, 109−116.
(53) Ferre-Gine, J.; Rallo, R.; Arenas, A.; Giralt, F. Extraction of structures embedded in the velocity field of a turbulent wake. In *Solving Engineering Problems with Neural Networks, Proceedings of the International Conference on Engineering Applications of Neural Networks (EANN'96)*; Bulsari, A. B., Kallio, S., Tsaptsinos, D., Turku, Eds.; 1996; Vol. 1, pp 17−20.
(54) Bartfai, G. An Adaptive Resonance Theory-based Neural Network capable of Learning via Representational Redescription. *In Proceedings of IEEE International Joint Conference of Neural Networks*; 1998; p 1137−1149.
(55) Bartfai, G. Hierarchical Clustering with ART Neural Netwoks. In *Proceedings of IEEE International Conference of Neural Networks*; 1994; Vol. 2, p 940−944.
(56) Bartfai, G. On the Match Tracking Anomaly of ARTMAP Neural Network. *Neural Networks* **1996**, *9*, 295−308.
(57) Bartfai, G. *An Improved Learning Algorithm for the fuzzy ARTMAP Neural Network*; Technical Report CS-TR-95/10, 1995, 1−12.
(58) Bartfai, G. An ART-based Modular Architecture for Learning Hierarchical Clustering. *Neurocomputing* **1996**, *13*, 31−45.
(59) *NeuralSim Software*; Aspen Technology, Inc.: 1999.

CI010323U