I&EC research
Industrial & Engineering Chemistry Research

ARTICLE

pubs.acs.org/IECR

Group Contribution-Based Method for Determination of Solubility Parameter of Nonelectrolyte Organic Compounds

Farhad Gharagheizi,[†] Ali Eslamimanesh,[‡] Amir H. Mohammadi,[*,‡,§] and Dominique Richon[‡]

[†]Saman Energy Giti Co., 3331619636 Tehran, Iran
[‡]MINES ParisTech, CEP/TEP—Centre Énergetique et Procédés, 35 Rue Saint Honoré, 77305 Fontainebleau, France
[§]Thermodynamics Research Unit, School of Chemical Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa

Supporting Information

ABSTRACT: The determination of the solubility parameter of organic compounds has been of much significance in the chemical industry. In this study, we propose a predictive method based on the combination of the Group Contribution strategy with the Artificial Neural Network to calculate/estimate the solubility parameter values of about 1620 nonelectrolyte organic compounds at 298.15 K and atmospheric pressure. The chemical functional groups are obtained for various compounds categorized in 81 different chemical families. The final results indicate the following statistical parameters of the presented method: average relative deviation (ARD %) of the determined properties from existing experimental values of 1.5% and a squared correlation coefficient of 0.985. It is finally inferred that the developed model is more accurate and predictive than our previously proposed models based on the Quantitative Structure−Property Relationship algorithm, which yielded 4.6, 3.4, and 3.1 ARD % from experimental values.

1. INTRODUCTION

Application of the solubility parameter concept in different chemical processes has been investigated since 1930s, when Scatchard[1] defined a physicochemical parameter standing for a solvent's affinity to dissolve a particular solute.[2] Hildebrand and Scott,[3] and Hansen[4,5] were the next researchers who improved the definition of the solubility parameter along with its various applications. So far, it has been demonstrated that the solubility parameter can be employed as a prominent property of a compound not only in coating and paint technologies, complex extraction operations, and polymer processes, and so forth[2,6] but also in many of the developed thermodynamic models for prediction of the amounts/conditions of precipitations/depositions of heavy petroleum fractions such as asphaltene and wax. Many of these models have been generally developed based on the regular solution theory,[7] which is based on the difference between the solubility parameter of the solute (asphaltene/wax) and related solvent (maltene/oil).[8−17]

Generally, the solubility parameter can be evaluated by the following equation:[1−5]

$$\delta = \left(\frac{\Delta E_v}{v}\right)^{1/2} = \left(\frac{\Delta U_{\mathrm{vap}}}{v}\right)^{1/2} = \left(\frac{\Delta H_{\mathrm{vap}} - RT}{v}\right)^{1/2} \quad (1)$$

where δ stands for the Hildebrand one-component solubility parameter, ΔE_v represents the cohesive energy, which is introduced as the energy required for separating a molecule from its surrounded neighbors,[1−5] v is the molar volume, ΔU_{vap} denotes the energy change upon isothermal vaporization of the saturated liquid to the ideal-gas state (energy of a complete vaporization),[17] and ΔH_{vap} is the enthalpy of vaporization. Internal pressure can also be applied for defining the physical meaning of the solubility parameter as follows:[1,18]

$$P_i = T\left(\frac{\partial P}{\partial T}\right)_{\mathrm{v}} - P = \delta^2 \quad (2)$$

where P_i stands for the internal pressure and T is temperature.

The interactions between the solvents and the solutes, originated from their electron pairs, donor−acceptors, and hydrogen bonding interactions, are not considered in the preceding equations. In other words, the aforementioned Hildebrand's parameter does not account for these interactions and considers only one part of the molecular forces (dispersion).[1,6] Consequently, the concept of the Hildebrand solubility parameter is normally applied for systems including weakly interacting species. As a result, Hildebrand's theory was modified by several researchers, from various fields, to define the two-component solubility parameter as follows:[19−25]

$$\delta = ({\delta_\lambda}^2 + {\delta_\tau}^2)^{1/2} \quad (3)$$

where subscripts λ and τ denote nonpolar and polar solubility parameters, respectively.

As a very fruitful modification, Hansen[5] proposed the three-component (Hansen) solubility parameter considering the effects of all of the cohesive bonds including the atomic dispersion forces, the molecular permanent dipole-permanent dipole forces,

Received: May 9, 2011
Accepted: August 2, 2011
Revised: July 28, 2011
Published: August 02, 2011

dx.doi.org/10.1021/ie201002e | *Ind. Eng. Chem. Res.* 2011, 50, 10344–10349

and the molecular hydrogen-bonding on the solubility parameter value as follows:[1,2,17]

$$\delta_{\mathrm{HSP}} = ({\delta_{\mathrm{D}}}^2 + {\delta_{\mathrm{P}}}^2 + {\delta_{\mathrm{H}}}^2)^{1/2} \tag{4}$$

where the subscripts D, P, and H denote the dispersion, polar, and hydrogen-bonding effects, respectively, and the subscript HSP denotes the total Hansen solubility parameter. The values of one-component and the total three-component solubility parameters would be almost the same for the substances with nonpolar, and non-hydrogen-bonding effects such as the light hydrocarbons.

The determination of the solubility parameter has been therefore of critical importance for the industry.[1−39] A detailed review of the corresponding methods for its evaluation can be found elsewhere.[17] This work is a continuation of the series of our efforts to develop predictive tools for the determination of the physicochemical properties of nonelectrolyte organic compounds using different approaches. In a previous communication,[17] our group presented three reliable models based on Quantitative Structure−Property Relationship (QSPR) to represent/predict the one-component solubility parameter of nonelectrolyte organic compounds reported in the DIPPR 801 database.[40] The results showed that the later models were reliable and comprehensive although developing such molecular-based models may not contain easy computational procedure. In this work, a new approach based on the combination of the Group Contribution (GC) method with Artificial Neural Network (ANN) is presented for this purpose. One of the main characteristics of the GC method is that this algorithm divides a molecule into small parts (generally named as "segments"). Each of these segments is considered as a functional group and has a contribution to the physicochemical properties of the specified molecule. Finally, the value of the property is defined through calculating the summation of the contributions of all functional groups in a molecule.

Furthermore, the Artificial Neural Networks have been applied to various scientific and engineering applications,[41−82] for example, calculations/estimations of the physical properties of different pure compounds[41−45] and phase behavior predictions of complex semiclathrate hydrate systems.[60] The theoretical explanations about Neural Networks have been well-established elsewhere.[83] As a consequence, a combination of the GC and the ANN methods normally leads to obtain accurate predictive tools for the evaluation of the desired properties of organic compounds. However, ANN is a mathematical tool that users must be very careful to apply its consequent results within the frame of the hypotheses and within the field of the data that allowed determination of the parameters (any extrapolation may not be recommended).

2. EXPERIMENTAL DATA AND MATHEMATICAL METHODS

2.1. Experimental Data. In this study, we have used the DIPPR 801 database,[40] which is one of the most reliable sources of physical property data for pure compounds, based on more than 23000 scientific sources. The solubility parameter values of 1620 nonelectrolyte organic species from various chemical families (81 families) at 298.15 K and atmospheric pressure have been treated for the calculation procedure. All of the data points have been evaluated by the DIPPR 801 project[40] for the investigated compounds.

2.2. Determination of New Functional Groups. Having defined the database, the chemical structures of all of the studied compounds have been analyzed in great detail using an algorithm comparing the chemical groups to define the most efficient contributions for evaluation of the solubility parameter. As a result, a new collection of 176 functional groups have been found to be more efficient for the representation/prediction of the corresponding parameter. These functional groups are more general than those of first-order, second-order, or third-order groups used in conventional group contribution methods. The functional groups used in this study are presented as Supporting Information. Moreover, the table of their numbers of occurrences in the investigated compounds is presented as Supporting Information.

2.3. Optimization of Group Contributions. The first calculation step is to find a relationship between the chemical functional groups and the desired physical property.[41−82] The traditional and perhaps the easiest method for this purpose is the assumption of the existence of a multilinear relationship between these groups and the property (here the solubility parameter).[41] This technique is a similar method to that used in the most of classical group contribution methods.[41] Several calculations show that application of the aforementioned methodology for the current problem does not contribute to accurate results within the range of the deviations from experimental values[40] that we are interested in. Consequently, a nonlinear mathematical method such as ANN is preferred and investigated here. Using the Artificial Neural Network toolbox of the MATLAB software (Mathworks Inc.), a three layer Feed Forward Artificial Neural Network (FFANN) has been developed for the problem.

Because we face a wide range of solubility parameter values for different compounds, these values have been normalized between −1 and +1 to prevent truncation errors.[41] This can be performed using maximum and minimum numbers of each functional group in each compound for input data and using maximum and minimum values of solubility parameter for output parameters.[41] In addition, this procedure, which is done in the optimization process, is performed to obtain the parameters of the Neural Networks (weights and bias), and it has no effects on the model results. Later, these values are again changed to the original solubility parameter values, which are finally used as the inputs and reported as outputs of the developed model.[41] In the second step, the database is divided into three subdata sets including the "Training" set, the "Optimization" set, and the "Test" set. In this study, the "Training" set is used to generate the ANN structure, the "Optimization" set is applied for optimization of the model, and the "Test (prediction)" set is used to investigate the prediction capability and validity of the proposed model. The division of database into three subdata sets is normally performed randomly. For this purpose, about 80%, 10%, and 10% of the main data set are randomly selected for the "Training" set (about 1296 solubility parameter data), the "Optimization" set (162 solubility parameter data), and the "Test" set (162 solubility parameter data). The effect of the percent allocation of the three subdata sets from the database on the accuracy of the final model has been studied elsewhere.[76] In distribution of the data through the three subdata sets, we generally perform many distributions to avoid the local accumulations of the data in the feasible region of the problem. As a result, the acceptable distribution is the one with homogeneous accumulations of the data on the domain of the three subdata sets.

There are generally two weight matrices and two bias vectors in a three layer FFANN: W_1 and W_2, b_1 and b_2. These parameters should be evaluated by minimization of an objective function. The objective function here is the summation of squares of errors between the outputs of the ANN (represented/predicted properties) and the target values (experimental solubility parameters). This minimization is performed by the Levenberg−Marquardt (LM)[83] optimization strategy. There are also more accurate optimization methods other than this algorithm; however, they need much more convergence time.[41−48] The more accurate optimization, the more time is needed for the algorithm to converge to the global optimum. The LM[83] is the most-widely used optimization method in these kinds of problems.[41]

In most cases, the number of neurons in the hidden layer (n) is fixed. Therefore, the objective is to produce the ANN model, which is capable of predicting the target values as accurately as possible. This step is repeated until the best ANN is obtained. Normally, in three-layer FFANNs, it is more efficient that the number of neurons in the hidden layer is optimized according to the accuracy of the produced FFANN.[29−61] Some factors should be taken into account in the determination of the optimum number of the neurons. By increasing the number of neurons, the accuracy of the model, that is, squared correlation coefficient (R^2), is increased on the "Training set". However, the accuracy of the model on the "Test set" is decreased gradually, and the model may become unstable. Consequently, the overall R^2, which depends on the three subdata sets, fluctuates during the changing of the numbers of neurons. The final (overall) R^2 value should be found through selecting the different number of neurons for a specified problem.[17]

3. RESULTS AND DISCUSSION

An optimized GC-ANN model has been obtained applying the preceding procedure for the determination of the desired parameter. For this purpose, several 3FFANNs modules were generated assuming numbers 1 through 50 for n (number of neurons in hidden layer). The most accurate results (no over-fitted and no under-fitted results) were observed at n = 10. It should be noted that this value is not a global one, because the optimization method used to train the ANN has great effects on the obtained value.[17,41] Therefore, the developed three-layer FFANN has the structure of 176-10-1 (176 chemical groups are regarded as the inputs of the algorithm).

A significant point that needs to be considered about the number of the model parameters is that, in each compound, only a few functional groups are present simultaneously (with the maximum number of 27 groups in pimaric acid and isopimaric acid). Therefore, for each compound, many of the model parameters, which in total are 10 × 176 = 1760, are zero and consequently, the developed model has few parameters for each compound, that is, between 0 to 270 parameters. The value of zero indicates that three of the investigated compounds do not contain the determined functional groups in their structures by the previous computational step. For these compounds, the model results in the solubility parameter value of 18.19 $(\mathrm{J/cm^3})^{0.5}$, which has been calculated by the intercept of the transfer function of the ANN algorithm (Refer to the Supporting Information file to observe the characteristics of all of the investigated organic compounds in this work). It should be pointed out that the number of the model parameters for the previous developed models[17] were 11, 131, and 13 for the linear-QSPR, ANN-QSPR, and LSSVM-QSPR models, respectively. However, the number of parameters for the two later nonlinear QSPR models[17] and the proposed GC-ANN model in this study may not be used as the only criterion for a comparison between the developed models because they are based on different network concepts with different mathematical characteristics, which have been well-established in our previous work.[17]

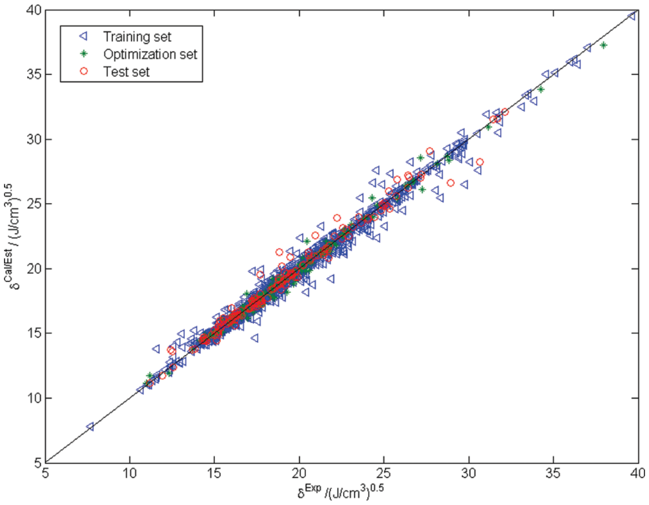

Figure 1. Comparison between the Cal (calculated)/Est (estimated) results using the developed model and experimental values[40] of solubility parameters.

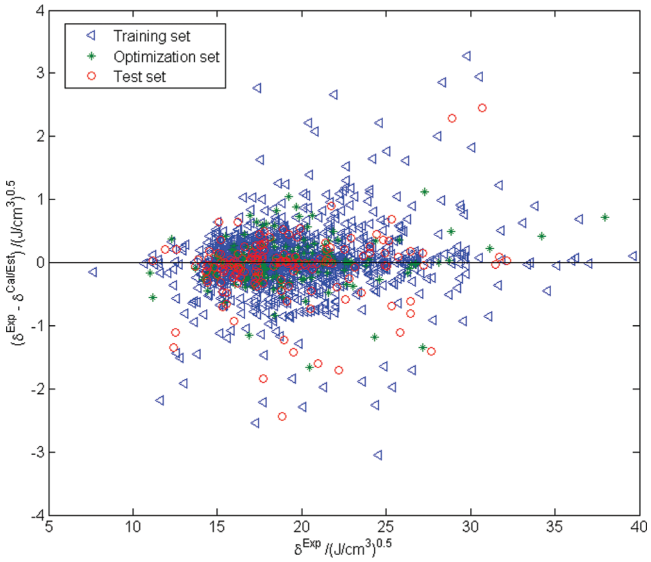

Figure 2. Deviations of the obtained results vs the corresponding experimental (Exp) solubility parameter values.[40]

The *mat* file (MATLAB file format) of the obtained GC-ANN containing all the parameters of the model (weight matrices and bias vectors) and the instructions for running the program are freely available upon request to the authors. Running the provided software, any researcher/engineer is able to determine the solubility parameter of a particular substance quickly. The employed functional groups to develop the model have been reported as Supporting Information. The calculated/estimated solubility parameter values are shown in Figure 1 in comparison with the experimental values.[40] Figure 2 indicates the deviations of the obtained values versus the experimental ones.[40]

The statistical results obtained by the GC-ANN model are reported in Table 1. Furthermore, the absolute average deviations

Table 1. Statistical Parameters of the Proposed GC-ANN Model

statistical parameter	value
Training Set	
R^2 [a]	0.985
average relative deviation [b]	1.5%
standard deviation error, [c] $(\mathrm{J\ cm^{-3}})^{0.5}$	3.86
root mean square error	0.48
N[d]	1296
Optimization Set	
R^2	0.991
average relative deviation	1.3%
standard deviation error, $(\mathrm{J\ cm^{-3}})^{0.5}$	3.95
root mean square error	0.38
N	162
Test Set	
R^2	0.991
average relative deviation	1.6%
standard deviation error, $(\mathrm{J\ cm^{-3}})^{0.5}$	4.06
root mean square error	0.55
N	162
Training + Optimization + Test Set	
R^2	0.985
average relative deviation	1.5%
standard deviation error, $(\mathrm{J\ cm^{-3}})^{0.5}$	3.89
root mean square error	0.48
N	1620

[a] R^2: Squared correlation coefficient, $R^2 = 1 - \{\sum_i^N [(\mathrm{Cal.}(i)/\mathrm{Est.}(i)) - \mathrm{Exp.}(i)]^2 / \sum_i^N [(\mathrm{Cal.}(i)/\mathrm{Est.}(i)) - \mathrm{average}(\mathrm{Exp.}(i))]^2\}$. [b] $\%\mathrm{ARD} = (100)/(N - n)\sum_i^N(|\mathrm{Cal.}(i)/\mathrm{Est.}(i) - \mathrm{Exp.}(i)|)/(\mathrm{Exp.}(i))$, where n is the number of the model parameters. [c] $\mathrm{std} = (1/N)\sum_i^N((\mathrm{Cal.}(i)/\mathrm{Est.}(i) - \mathrm{average}(\mathrm{Cal.}(i)/\mathrm{Est.}(i)))^2)^{1/2}$.

[d] Number of data points.

of the results from experimental values[40] for each 81 chemical families are reported as Supporting Information. The results imply that the obtained ANN-GC model is more accurate and predictive than our previously presented models[17] based on the QSPR strategy. For better illustration of the organic compounds investigated in this work, their chemical structures are sketched and presented as Supporting Information.

It should be noted that we have observed that there are 17 data points (from 1620 available data points) for which the presented model results lead to more than 10% average relative deviations from experimental values.[40] It seems that there is no relation between these compounds' structures to show some weaknesses in determining the solubility parameter values of related chemical families. Therefore, it is probable that the solubility parameter values for these compounds are not accurate or may be somehow erroneous (or with high uncertainty) because of the existing difficulties and possible errors in experimental measurements.[84−87] The proposed method has been developed based on all of the available data including the aforementioned 17 data points. The final significant point is that we do not have access to any other solubility parameter database to check the capability of the presented model for prediction of solubility parameters of completely nonsimilar compounds to those investigated. This fact again proves the imperative need for the development of new experimental techniques and apparatuses to generate more solubility parameter data for particular industrial and theoretical purposes.

4. CONCLUSION

In this work, a group contribution-based model was presented for the representation/prediction of the solubility parameter of nonelectrolyte organic compounds at 298.15 K and atmospheric pressure. The model is the result of a combination of Feed Forward Artificial Neural Networks and Group Contributions. The required parameters of the model are the numbers of occurrences of 176 functional groups in each investigated molecule. It should be noted that most of these functional groups are not simultaneously available in a particular molecule. Therefore, the computation of the required parameters from the chemical structure of any molecule is simple. For developing the model, the experimental solubility parameter values from the largest available data set[40] containing 1620 pure organic compounds from 81 different chemical families were applied. As a consequence, a reliable and predictive tool was developed to determine the solubility parameter values of many of organic compounds, which are especially applied in chemical and petroleum industry. However, one point should not be omitted from our conclusion: The model has a wide range of applicability but the prediction capability of the model is restricted to the compounds which are similar to the applied to develop the model. Application of the model for totally different compounds than the investigated ones is conservative although it may be used for a rough estimation of the solubility parameter of these kinds of compounds.

■ ASSOCIATED CONTENT

Ⓢ **Supporting Information.** There are four files including the number of occurrences of the 176 functional groups in all of 1620 nonelectrolyte organic compounds in the main data set, the distributions of the data in three data sets, the obtained results in the first Excel file, the functional groups used to develop the model in the second Excel file, the average absolute deviations of the obtained results from experimental values of various chemical families in the third Excel file, and the schematic structures of all of the investigated compounds in the pdf format file. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: amir-hossein.mohammadi@mines-paristech.fr. Phone: + (33) 1 64 69 49 70. Fax: + (33) 1 64 69 49 68.

■ ACKNOWLEDGMENT

Ali Eslamimanesh wishes to thank MINES ParisTech for providing him a Ph.D. scholarship.

■ REFERENCES

(1) Scatchard, G. Equilibria in Non-electrolyte solutions in relation to the vapor pressures and densities of the components. *Chem. Rev.* **1931**, *8*, 321–333.

(2) Eslamimanesh, A.; Esmaeilzadeh, F. Estimation of solubility parameter by the modified ER equation of state. *Fluid Phase Equilib.* **2010**, *291*, 141–150.

(3) Hildebrand, J. H.; Scott, R. L. *The Solubility of Nonelectrolytes*; Rheinhold Publishing Corporation: New York, 1950.

(4) Hansen, C. M. The three dimensional solubility parameter - key to paint component affinities I. - Solvents, plasticizers, polymers, and resins. *J. Paint Technol.* **1967**, *39*, 104–117.

(5) Hansen, C. M. The Universality of the Solubility Parameter. *Ind. Eng. Chem. Pro. Res. Develop.* **1969**, *8*, 2–11.

(6) Gharagheizi, F.; Sattari, M.; Angaji, M. T. Effect of calculation method on values of Hansen solubility parameters of polymers. *Polym. Bull.* **2006**, *57*, 377–384.

(7) Hildebrand, J. H. Solubility: XII. Regular Solutions. *J. Am. Chem. Soc.* **1929**, *51*, 66–80.

(8) Mousavi-Dehghani, S. A.; Mirzayi, B.; Mousavi, S. M. H.; Fasih, M. An applied and efficient model for asphaltene precipitation in production and miscible gas injection. *Pet. Sci. Technol.* **2010**, *28*, 113–124.

(9) Novosad, Z.; Costain, T. G. Experimental and modeling studies of asphaltene equilibria for a reservoir under CO_2 injection. In *Proceedings of the 65th Annual Technical Conference and Exhibition of the SPE*, SPE 20530, New Orleans, Los Angeles, United States, 1990.

(10) Nor-Azian, N.; Adewumi, M. A. Development of asphaltene phase equilibrium predictive model. In *Proceedings of the Eastern Regional Conference and Exhibition of the SPE*, SPE 26905, Nov., Richardson, Texas, 1993.

(11) MacMillan, D. J.; Tackett, J. E.; Jessee, M. A.; Monger-McClure, T. G. A unified approach to asphaltene precipitation: Laboratory measurement and modeling. In *Proceedings of the SPE International Symposium on Oilfield Chemistry*, San Antonio, Texas, Feb., SPE 28990, Richardson, Texas, 1995.

(12) Yang, Z.; Ma, C. F.; Lin, S. X.; Yang, J. T.; Guo, T. M. Experimental and modeling studies on the asphaltene precipitation in degassed and gas-injected reservoir oils. *Fluid Phase Equilib.* **1999**, *157*, 143–158.

(13) Alboudwarej, H.; Akbarzadeh, K.; Beck, J.; Svercek, W. Y.; Yarranton, H. W. Regular solution model of asphaltene precipitation from bitumen. *AIChE J.* **2003**, *11*, 2948–2956.

(14) Akbarzadeh, K.; Alboudwarej, H.; Svercek, H. Y.; Yarranton, H. W. A generalized regular solution model for the prediction of asphaltene precipitation from n-alkane diluted heavy oils and bitumens. *Fluid Phase Equilib.* **2005**, *232*, 159–170.

(15) Yarranton, H. W.; Masliyah, J. H. Molar mass distribution and solubility modeling of asphaltenes. *AIChE J.* **1996**, *42*, 3533–3543.

(16) Mohammadi, A. H.; Richon, D. The Scott-Magat polymer theory for determining onset of precipitation of dissolved asphaltene in the solvent + precipitant solution. *Open Thermodyn. J.* **2008**, *2*, 13–16.

(17) Gharagheizi, F.; Eslamimanesh, A.; Farjood, F.; Mohammadi, A. H.; Richon, D. Solubility parameter of non-electrolyte organic compounds: Determination using quantitative structure−property relationship strategy. *Ind. Eng. Chem. Res.* **2011**, accepted manuscript, DOI: 10.1021/ie200962w.

(18) Van Arkel, A. E. Mutual solubility of liquids. *Faraday Trans.* **1946**, *42B*, 81–84.

(19) Small, P. A. Some factors affecting the solubility of polymers. *J. Appl. Chem.* **1953**, *3*, 71–80.

(20) Anderson, R. Polar organic solvents and aromatic hydrocarbons. Ph.D. Thesis, Dept. of Chemical Engineering, University of California, Berkeley, CA, 1961.

(21) Blanks, R. F.; Prausnitz, J. M. Thermodynamics of polymer solubility in polar and nonpolar systems. *Ind. Eng. Chem. Fundam.* **1964**, *3*, 1–8.

(22) Pausnitz, J. M.; Anderson, R. Thermodynamics of solvent selectivity in extractive distillation of hydrocarbons. *AIChE J.* **1961**, *7*, 96–101.

(23) Weimer, R. F.; Prausnitz, J. M. Complex formation between carbon tetrachloride and aromatic hydrocarbons. *J. Chem. Phys.* **1965**, *42*, 3643–3644.

(24) Prausnitz, J. M.; Shair, F. H. A thermodynamic correlation of gas solubilities. *AIChE J.* **1961**, *7*, 682–687.

(25) Lyckman, E. W.; Eckert, C. A.; Prausnitz, J. M. Generalized liquid volumes and solubility parameters for regular solution application. *Chem. Eng. Sci.* **1965**, *20*, 703–706.

(26) Fedors, R. F. A method for estimating both the solubility parameter and molar volumes of liquids. *Polym. Eng. Sci.* **1974**, *14*, 147–154.

(27) Lozada, C. M.; Del Rio, F. Calculation of solubility parameters from an equation of state. *Chem. Phys. Lett.* **1977**, *45*, 130–1130.

(28) Hansen, C. M.; Beerbower, A. Solubility parameters. In *Kirk-Othmer Encyclopedia of Chemical Technology*, Suppl. Vol., 2nd ed.; Standen, A., Ed.; Interscience: New York, 1971.

(29) Beerbower, A. Environmental Capability of Liquids, in Interdisciplinary Approach to Liquid Lubricant Technology. NASA Publication SP-318; NASA: Washington, DC, 1973.

(30) Van Krevelen, D. W.; Hoftyzer, P. J. *Properties of Polymers: Their Estimation and Correlation with Chemical Structure*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 1976.

(31) Allada, S. R. Solubility parameters of supercritical fluids. *Ind. Eng. Chem. Prod. Res. Dev.* **1984**, *23*, 344–348.

(32) Panayiotou, C. Solubility parameter revisited: An equation-of-state approach for its estimation. *Fluid Phase Equilib.* **1997**, *131*, 21–35.

(33) Williams, L. L.; Rubin, J. B.; Edwards, H. W. Calculation of Hansen solubility parameter values for a range of pressure and temperature conditions, including the supercritical fluid region. *Ind. Eng. Chem. Res.* **2004**, *43*, 4967–4972.

(34) Bozdogan, A. E. A method for determination of thermodynamic and solubility parameters of polymers from temperature and molecular weight dependence of intrinsic viscosity. *Polymer* **2004**, *45*, 6415–6424.

(35) Utracki, L. A.; Simha, R. Statistical thermodynamics predictions of the solubility parameter. *Polym. Int.* **2004**, *53*, 279–286.

(36) Stefanis, E.; Tsivintzelis, I.; Panayiotou, C. The partial solubility parameters: An equation-of-state approach. *Fluid Phase Equilib.* **2006**, *240*, 144–154.

(37) Zeng, Z.-Y.; Xu, Y.-Y.; Li, Y.-W. Calculation of solubility parameter using Perturbed-Chain SAFT and Cubic-Plus-Association equations of state. *Ind. Eng. Chem. Res.* **2008**, *47*, 9663–9669.

(38) Code, J. E.; Holder, A. J.; Eick, J. D. Direct and indirect quantum mechanical QSPR Hildebrand solubility parameter models. *QSAR Comb. Sci.* **2008**, *27*, 841–849.

(39) Vargas, F. M.; Chapman, W. G. Application of the One-Third rule in hydrocarbon and crude oil systems. *Fluid Phase Equilib.* **2010**, *290*, 103–108.

(40) Project 801, Evaluated Process Design Data, Public Release Documentation, Design Institute for Physical Properties (DIPPR), American Institute of Chemical Engineers (AIChE), 2006.

(41) Gharagheizi, F.; Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. Use of artificial neural network-group contribution method to determine surface tension of pure compounds. *J. Chem. Eng. Data* **2011**, *56*, 2587−2601.

(42) Gharagheizi, F.; Sattari, M. Prediction of triple-point temperature of pure components using their chemical structures. *Ind. Eng. Chem. Res.* **2010**, *49*, 929–932.

(43) Gharagheizi, F.; Abbasi, R.; Tirandazi, B. Prediction of Henry's law constant of organic compounds in water from a new group-contribution-based model. *Ind. Eng. Chem. Res.* **2010**, *49*, 10149–10152.

(44) Eslamimanesh, A.; Gharagheizi, F.; Mohammadi, A. H.; Richon, D. Artificial neural network modeling of solubility of supercritical carbon dioxide in 24 commonly used ionic liquids. *Chem. Eng. Sci.* **2011**, *66*, 3039–3044.

(45) Gharagheizi, F.; Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. Artificial neural network modeling of solubilities of 21 mostly-used industrial solid compounds in supercritical carbon dioxide. *Ind. Eng. Chem. Res.* **2011**, *50*, 221–226.

(46) Gharagheizi, F.; Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. Representation/prediction of solubilities of pure compounds in water using artificial neural network−group contribution method. *J. Chem. Eng. Data* **2011**, *56*, 720–726.

(47) Gharagheizi, F.; Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. QSPR approach for determination of parachor of non-electrolyte organic compounds. *Chem. Eng. Sci.* **2011**, *66*, 2959–2967.

(48) Gharagheizi, F.; Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. Determination of parachor of various compounds using an artificial neural network - group contribution method. *Ind. Eng. Chem. Res.* **2011**, *50*, 5815−5823.

(49) Chouai, A.; Laugier, S.; Richon, D. Modeling of thermodynamic properties using neural networks: Application to refrigerants. *Fluid Phase Equilib.* **2002**, *199*, 53–62.

(50) Piazza, L.; Scalabrin, G.; Marchi, P.; Richon, D. Enhancement of the extended corresponding states techniques for thermodynamic modelling. I. Pure fluids. *Int. J. Refrig.* **2006**, *29*, 1182–1194.

(51) Scalabrin, G.; Marchi, P.; Bettio, L.; Richon, D. Enhancement of the extended corresponding states techniques for thermodynamic modelling. II. Mixtures. *Int. J. Refrig.* **2006**, *29*, 1195–1207.

(52) Chapoy, A.; Mohammadi, A. H.; Richon, D. Predicting the hydrate stability zones of natural gases using Artificial Neural Networks. *Oil Gas Sci. Technol. − Rev. IFP* **2007**, *62*, 701–706.

(53) Mohammadi, A. H.; Richon, D. Hydrate phase equilibria for hydrogen + water and hydrogen + tetrahydrofuran + water systems: Predictions of dissociation conditions using an artificial neural network algorithm. *Chem. Eng. Sci.* **2010**, *65*, 3352–3355.

(54) Mohammadi, A. H.; Richon, D. Estimating sulfur content of hydrogen sulfide at elevated temperatures and pressures using an artificial neural network algorithm. *Ind. Eng. Chem. Res.* **2008**, *47*, 8499–8504.

(55) Mohammadi, A. H.; Richon, D. A mathematical model based on artificial neural network technique for estimating liquid water − hydrate equilibrium of water − hydrocarbon System. *Ind. Eng. Chem. Res.* **2008**, *47*, 4966–4970.

(56) Mohammadi, A. H.; Afzal, W.; Richon, D. Determination of critical properties and acentric factors of petroleum fractions using artificial neural networks. *Ind. Eng. Chem. Res.* **2008**, *47*, 3225–3232.

(57) Mohammadi, A. H.; Richon, D. Use of artificial neural networks for estimating water content of natural gases. *Ind. Eng. Chem. Res.* **2007**, *46*, 1431–1438.

(58) Mohammadi, A. H.; Martínez-López, J. F.; Richon, D. Determining phase diagrams of tetrahydrofuran+methane, carbon dioxide or nitrogen clathrate hydrates using an artificial neural network algorithm. *Chem. Eng. Sci.* **2010**, *65*, 6059–6063.

(59) Mehrpooya, M.; Mohammadi, A. H.; Richon, D. Extension of an artificial neural network algorithm for estimating sulfur content of sour gases at elevated temperatures and pressures. *Ind. Eng. Chem. Res.* **2010**, *49*, 439–442.

(60) Mohammadi, A. H.; Belandria, V.; Richon, D. Use of an artificial neural network algorithm to predict hydrate dissociation conditions for hydrogen + water and hydrogen + tetra-*n*-butyl ammonium bromide + water systems. *Chem. Eng. Sci.* **2010**, *65*, 4302–4305.

(61) Gharagheizi, F. A new group contribution-based method for estimation of lower flammability limit of pure compounds. *J. Hazard. Mater.* **2009**, *170*, 595–604.

(62) Gharagheizi, F. New Neural Network Group Contribution model for estimation of lower flammability limit temperature of pure compounds. *Ind. Eng. Chem. Res.* **2009**, *48*, 7406–7416.

(63) Gharagheizi, F.; Sattari, M. Estimation of molecular diffusivity of pure chemicals in water: A quantitative structure-property relationship study. *SAR QSAR Environ. Res.* **2009**, *20*, 267–285.

(64) Gharagheizi, F. Prediction of standard enthalpy of formation of pure compounds using molecular structure. *Aust. J. Chem.* **2009**, *62*, 376–381.

(65) Gharagheizi, F.; Tirandazi, B.; Barzin, R. Estimation of aniline point temperature of pure hydrocarbons: A Quantitative Structure-Property Relationship approach. *Ind. Eng. Chem. Res.* **2009**, *48*, 1678–1682.

(66) Gharagheizi, F.; Mehrpooya, M. Prediction of some important physical properties of sulfur compounds using QSPR models. *Mol. Diversity* **2008**, *12*, 143–155.

(67) Sattari, M.; Gharagheizi, F. Prediction of molecular diffusivity of pure components into air: A QSPR approach. *Chemosphere* **2008**, *72*, 1298–1302.

(68) Gharagheizi, F.; Alamdari, R. F.; Angaji, M. T. A new neural network-group contribution method for estimation of flash point. *Energy Fuels* **2008**, *22*, 1628–1635.

(69) Gharagheizi, F.; Fazeli, A. Prediction of Watson characterization factor of hydrocarbon compounds from their molecular properties. *QSAR Comb. Sci.* **2008**, *27*, 758–767.

(70) Gharagheizi, F.; Alamdari, R. F. A molecular-based model for prediction of solubility of c60 fullerene in various solvents. *Fullerenes, Nanotubes, Carbon, Nanostruct.* **2008**, *16*, 40–57.

(71) Gharagheizi, F. A new neural network Quantitative Structure-Property Relationship for prediction of θ (Lower Critical Solution Temperature) of polymer solutions. *e-Polym.* **2007**, *114*, 1–5.

(72) Gharagheizi, F. QSPR studies for solubility parameter by means of genetic algorithm-based multivariate linear regression and generalized regression neural network. *QSAR Comb. Sci.* **2008**, *27*, 165–170.

(73) Gharagheizi, F. A chemical structure-based model for estimation of upper flammability limit of pure compounds. *Energy Fuels* **2010**, *27*, 3867–3871.

(74) Vatani, A.; Mehrpooya, M.; Gharagheizi, F. Prediction of standard enthalpy of formation by a QSPR Model. *Int. J. Mol. Sci.* **2007**, *8*, 407–432.

(75) Mehrpooya, M.; Gharagheizi, F. A Molecular approach for prediction of sulfur compounds solubility parameters, phosphorus sulfur and silicon and related elements. *Phosphorus, Sulfur Silicon Relat. Elem.* **2010**, *185*, 204–210.

(76) Gharagheizi, F. QSPR analysis for intrinsic viscosity of polymer solutions by means of GA-MLR and RBFNN. *Comput. Mater. Sci.* **2007**, *40*, 159.

(77) Gharagheizi, F. An accurate model for prediction of autoignition temperature of pure compounds. *J. Hazard. Mater.* **2011**, *189*, 211–221.

(78) Gharagheizi, F.; Sattari, M.; Tirandazi, B. Prediction of lattice crystal energy using enthalpy of sublimation: A group contribution-based model. *Ind. Eng. Chem. Res.* **2011**, *50*, 2482–2486.

(79) Gharagheizi, F.; Eslamimanesh, E.; Mohammadi, A. H.; Richon, D. Representation/prediction of molecular diffusivity of pure compounds in water at infinite dilution using artificial neural network-group contribution method. *J. Chem. Eng. Data* **2011**, *56*, 1741−1750.

(80) Gharagheizi, F.; Babaie, O.; Mazdeyasna, S. Prediction of vaporization enthalpy of pure compounds using a group contribution-based method. *Ind. Eng. Chem. Res.* **2011**, *50*, 6503−6507.

(81) Gharagheizi, F.; Mirkhani, S. A.; Mahyari, A. R. T. Prediction of standard enthalpy of combustion of pure compounds using a very accurate group contribution-based method. *Energy Fuels* **2011**, *25*, 2651−2654.

(82) Gharagheizi, F.; Salehi, G. R. Prediction of enthalpy of fusion of pure compounds using an artificial neural network-group contribution method *Thermochim. Acta* **2011**, *521*, 37−40.

(83) Hagan, M.; Demuth, H. B.; Beale, M. H. *Neural Network Design*; International Thomson Publishing Co.: Salt Lake City, UT, 2002.

(84) Mohammadi, A. H.; Eslamimanesh, A.; Richon, D. Wax solubility in gaseous system: Thermodynamic consistency test of experimental data. *Ind. Eng. Chem. Res.* **2011**, *50*, 4731–4740.

(85) Eslamimanesh, A.; Yazdizadeh, M.; Mohammadi, A. H.; Richon, D. Experimental data assessment test for diamondoids solubility in gaseous system. *J. Chem. Eng. Data* **2011**, *56*, 2655−2659.

(86) Eslamimanesh, A.; Mohammadi, A. H.; Richon, D. Thermodynamic consistency test for experimental data of water content of methane. *AIChE J.* **2011**, *57*, 2566−2573.

(87) Coquelet, C.; Galicia-Luna, L. A.; Mohammadi, A. H.; Richon, D. The essential importance of experimental research and the use of experimental thermodynamics to the benefit of industry. *Fluid Phase Equilib.* **2010**, *296*, 2–3.