



## Prediction of autoignition temperatures of organic compounds by the structural group contribution approach

Chan-Cheng Chen\*, Horng-Jang Liaw, Yu-Yu Kuo

Department of Occupational Safety and Health, China Medical University, 91 Hsueh-Shih Road, Taichung 40402, Taiwan, ROC

### ARTICLE INFO

#### Article history:

Received 6 December 2007

Received in revised form 11 May 2008

Accepted 21 May 2008

Available online 10 July 2008

#### Keywords:

Autoignition temperature

Structural group contribution

### ABSTRACT

A model to predict the autoignition temperatures (AIT) of organic compounds is proposed based on the structural group contribution (SGC) approach. This model has been built up using a 400-compound training set; the fitting ability for these training data is 0.8474, with an average error of 32 K and an average error percentage of 4.9%. The predictive capability of the proposed model has been demonstrated on an 83-compound validation set; the predictive capability for these validation data is about 0.5361, with an average error of 70 K and an average error percentage of 11.0%. The proposed model is shown to be more accurate than those of other published works. This improvement is largely attributed to the modifications of the group definitions for estimating the AIT instead of the type of empirical model chosen. Through the  $Q^2$  value and hypothesis testing, it was found that the empirical model should be chosen as a polynomial of degree 3. As compared to the known errors in experimentally determining the AIT, the proposed method offers a reasonable estimate of the AIT for the organic compounds in the training set, and can also approximate the AIT for compounds whose AIT is as yet unknown or not readily available to within a reasonable accuracy.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

The autoignition temperature (AIT) is defined as the lowest temperature at which a substance will ignite in the absence of a spark or flame. Based on the thermal theory of ignition and on classical reaction-rate theory, the AIT can be regarded as the temperature to which a combustible mixture must be raised so that the rate of heat evolution from the exothermic oxidation reactions of the system will just overbalance the rate at which heat is lost to the surroundings. The ability of a substance to spontaneously ignite introduces potential fire hazards for all who must handle, transport, or store combustible materials. Thus, risk assessment methods such as API-581 usually take the AIT of a substance as an essential input parameter to define the possible consequences of a leakage of combustible liquids [1]. Autoignition is also studied in relation to the performance of combustion engines through the phenomenon of engine knock [2].

As the AIT is the temperature at which a material will spontaneously ignite when exposed to the atmosphere, it depends not only on the chemical and physical properties of the substance but also the method and the apparatus employed for its determina-

tion, such as the test pressure, oxygen concentration, and volume and material of the vessel used. Hence, it is very common that the AIT of a specific compound is reported differently in different literatures; these differences may be as much as 300 K. For example, as shown in Table 1, the AIT of acetaldehyde is reported to range from 413 to 758.15 K in five different authoritative sources [3–7]. Table 1 lists some compounds for which the difference in AIT is more than 100 K across separate sources. One of the important reasons that results in this uncertainty is the value of AIT reported in different databases may be conducted by different experimental method. Even in the same database, different compounds may be conducted in different experimental methods [2]. However, all databases usually reported the AIT value without the information of experimental method employed. For example, there is no way to trace back the data reported in the hazardous chemical database or SAX'S dangerous properties of industrial materials. Even in the famous DIPPR® project, the data quality of AIT is still flagged as “unevaluated”. Since AIT is usually reported without the experimental method employed, it is not possible to account this bias by including experimental method as an additional explanatory variable or to group them by different experimental methods. In addition, because visual inspection is chosen to detect the sudden appearance of a flame inside the autoignition vessel, determining AIT is greatly limited by human capabilities [8]. Usually, the average error in experimentally determined AIT values is deemed to be about  $\pm 30$  K in the literature [9]. Besides this problem, the determination

\* Corresponding author. Tel.: +886 4 22076435; fax: +886 4 22076435.  
E-mail address: [ChanCheng.Chen@mail.cmu.edu.tw](mailto:ChanCheng.Chen@mail.cmu.edu.tw) (C.-C. Chen).

**Table 1**  
Experimental AITs for selected compounds from different sources<sup>a</sup>

Compound name	Ref. [3]	Ref. [4]	Ref. [5]	Ref. [6]	Ref. [7]
2-Butanone	677	788.71	–	788.7	778
2,4-Dimethylphenol	753	872	–	–	872
Hexadecanoic acid	513	650	–	–	–
Piperazidine	593	728.15	–	–	593
1,3-Diisopropylbenzene	722	349.82	–	–	–
Benzoyl chloride	873	358.15	–	–	470.2
Methylhexanone	728	464.15	–	–	–
2-Methylnitrobenzene	693	578.15	–	–	693
2,4-Dihydroxy-2-methylpentane	698	579	–	–	579
1-Methyl-2-pyrrolidinone	518	619.15	–	–	543
2-Heptanone	805	666.15	–	805.93	666
Crotonic acid	773	669.26	–	–	669
1,4-Benzenedicarboxylic acid	951	769	–	–	769
2,4-Dimethylpentane	598	708	–	–	–
1,3-Benzenedicarboxylic acid	973	769	–	–	–
Phenol	878	988.15	988	–	988
Isobutyl formate	698	593.15	593	–	–
Acetaldehyde	413	758.15	448	510.93	458

<sup>a</sup>The AIT values are in K.

of the AIT by experiment is very laborious and is not always feasible [8]. In this light, the ability to estimate AIT values by mathematical modeling will be a cost-efficient and critical aid to this discipline.

Multivariate statistical methods such as multiple linear regression and principal component regression are important approaches to predict the AITs of organic compounds. Several studies have considered using the physical properties of compounds, such as the critical pressure, parachor, and molecular weight, as descriptors to predict the AIT [2,9,10–13]. Such approaches are usually known as quantitative structure–property relation (QSPR) approaches in which the molecular structures are characterized by the various physical properties of the compound. The underlying assumption in the QSPR approach is that there is some sort of relationship between the properties of interest and the molecular descriptors (i.e., the measurable physical properties). Thus, the QSPR approach involves multivariate analysis using several measurements (or descriptors) to deduce the desired property. Although the QSPR approach has been shown to estimate the properties of compounds with certain degree of success, it is applicable, however, only if the measurements of the molecular descriptors are available. On the other hand, the structure group contribution (SGC) method directly uses the information of the molecular structure instead of the physical properties to predict the AIT. Thus, the SGC method may be a more attractive alternative as it is still applicable when the physical properties of the target compound are unavailable and even if the target compound is an unknown compound.

The SGC method has recently found wide commercial application in the form of computer programs that estimate the properties of pure substances from their chemical structures, for example, the ASTM CHETAH [14]. While the SGC approach has been successfully applied to predict many of the properties of compounds, very little literature is available on predicting the AIT through the SGC approach. Albahri [15] proposed a SGC method in which a polynomial of degree 4 is suggested as an empirical model for predicting the AITs of 137 pure hydrocarbons with an average error percentage of 4.2% and a maximum error percentage of 31%. To consider organic compounds other than pure hydrocarbons, Albahri and George [16] developed a predictive model based on a 490-compound database in which the organic compounds include hetero-atoms such as oxygen, and nitrogen. In their work, they first chose a polynomial of degree 4 as the empirical model, but found that the average error percentage and maximum error percentage for such a model were 9.2% and 125%, respectively. Because of the limited success, they proposed a three-layer artificial neural network (ANN) structure

to improve the predictive performance. This ANN-SGC approach seemed to offer a significant improvement in performance, with average error percentage of 2.8% and a maximum error percentage of 20%; however, although this ANN-SGC approach exhibited better performance than that of the classical SGC approach, it usually included too many weighting parameters and was also inconvenient for desk calculation. Thus, there is still a demand for a more accurate method to predict the AIT by the classical SGC approach.

This article is organized as follows: First, the AIT database and group definitions for this work are discussed in Section 2. In Section 3, the mathematics for developing an empirical model is discussed. A brief discussion on choosing and evaluating the empirical model is provided in this section. The results of this work and some discussion are provided in Section 4. Finally, the conclusions are presented in Section 5.

## 2. Database and group definitions

In this work, the prediction model was developed and validated from a 483-compound database in which some of the organic compounds contained hydrogen, carbon, oxygen, nitrogen, or halogen atoms. In this database, there were 150 pure hydrocarbons, and the other 333 compounds included hetero-atoms. These 483 compounds were randomly distributed into a training set with 400 compounds and a validation set with 83 compounds. As mentioned earlier, different AIT values were reported in different literature sources, thus the value in AIChE-DIPPR was adopted as the standard in this work. However, in the case that the AIT of a compound was not available in DIPPR or the AITs reported in all the other literature sources were consistent with each other but differed from that of DIPPR, the non-DIPPR value was adopted. The corresponding AITs for all 483 compounds in this work are listed in Tables A1 and A2. In these 483 data, 300 data were taken from AIChE-DIPPR and the other 183 data were taken from non-DIPPR sources.

To elucidate the classification of group definitions for estimating the AITs of organic compounds, a brief review of the mechanism of the autoignition process is provided below. Swarts and Orchin [17] reported that the autoignition mechanism proceeds by a free radical reaction and the stability of free-radical intermediates determines the ease of oxidation. Thus, the AIT of a compound is highly affected by its molecular structure because the stability of the free radicals that are formed is related directly to the molecular structure of the compound. For hydrocarbons, decreasing the chain length, addition of methyl groups, unsaturation, branching, and cyclic and aromatic

**Table 2**  
Group contribution for estimation of the AIT

Group no.	Group	Remark	MLR	Degree 2	Degree 3	Degree 4
1	—CH <sub>3</sub>		-22.6813	-29.9804	-22.8857	-23.1736
2	>CH <sub>2</sub>		-21.3527	-39.6875	-28.5961	-30.5288
3	>CH—		4.0509	0.5544	1.3340	0.5424
4	>C<		56.6464	54.9627	49.7423	48.2437
5	=CH <sub>2</sub>		-17.7405	-31.9088	-21.6668	-22.9287
6	=CH—		-39.7751	-66.2563	-46.3286	-49.5017
7	>C=		-31.1140	-36.4970	-32.7605	-33.1968
8	≡CH		-71.9616	-117.6474	-81.0169	-85.7134
9	≡C—		-55.6234	-94.6821	-64.4957	-68.9986
10	>CH <sub>2</sub>		-24.6453	-40.1768	-28.4401	-30.1512
11	>CH—		-4.0091	-8.4713	-6.7179	-7.3545
12	>C<		-15.3282	-62.9661	-21.7342	-26.1364
13	=CH—		20.5679	17.5968	19.5293	20.2876
14	>C=		-48.4525	-43.7461	-49.0687	-50.6287
15	=CH—		6.1265	2.5173	6.2350	6.1348
16	>C	(Fused)	10.0809	7.8480	5.8332	4.9528
17	>C	(Nonfused)	22.0000	24.2898	15.9976	14.9799
18	—CH <sub>3</sub>	(Attached to at least	120.6545	82.3095	103.2738	114.6114
19	>CH <sub>2</sub>	one halogen atom)	-3.0301	-17.5437	-9.8344	-11.3142
20	>CH—		-23.3669	-37.6850	-24.2759	-26.4969
21	>C<		242.6675	247.3442	293.5064	287.6001
22	—F	(Nonring)	-42.1638	-65.1907	-45.0477	-48.2274
23	—Cl	(Nonring)	37.8512	34.3163	33.9332	33.9492
24	—Br	(Nonring)	-26.7706	-38.4893	-27.8628	-29.3678
25	—F	(Ring)	49.5353	49.8512	88.8289	90.5979
26	—Cl	(Ring)	27.1279	10.1433	79.4122	73.0277
27	—Br	(Ring)	54.0271	26.6567	53.4870	50.0560
28	—OH	(Alcohol)	-10.2828	-15.2847	-8.9378	-9.0980
29	—OH	(Phenol)	43.0417	19.8772	134.3524	135.0214
30	—O—	(Nonring)	-54.7172	-102.1454	-70.0383	-74.8046
31	—O—	(Ring)	-27.3097	-40.1570	-28.4801	-29.4928
32	>C=O	(Nonring)	10.3136	4.4457	8.1173	7.4529
33	>C=O	(Ring)	46.6977	37.5549	57.5044	58.9186
34	O=CH—	(Aldehyde)	-120.8556	-208.1346	-138.3186	-147.3836
35	—COOH	(Acid)	7.4758	-5.7007	4.0037	3.5463
36	—COO—	(Ester)	35.2785	34.2704	35.2011	35.1839
37	—NH <sub>2</sub>		-17.8076	-33.4276	-17.7579	-18.9367
38	>NH	(Nonring)	-0.7601	1.5113	-1.8223	-1.5263
39	>NH	(Ring)	22.7367	22.3273	24.5474	25.8972
40	>N—	(Nonring)	-1.1683	-20.3543	-4.7926	-8.1033
41	>N—	(Ring)	-42.3160	-47.2060	-49.3834	-51.0105
42	—N=	(Nonring)	-35.3780	-75.6105	-41.9897	-47.4827
43	—N=	(Ring)	33.2611	3.5793	31.8743	24.8482
44	—CN		82.0245	86.5584	80.5038	81.3197
45	—NO <sub>2</sub>		-46.9847	-65.2997	-52.7670	-53.2661

structures will elevate the AIT [18]. The relationship between the AITs of lower alkanes and their corresponding alcohols and aldehydes is in the following order: alkane > alcohol > aldehyde [17]. It has also been reported that there is no distinction in the AITs for the *cis* and *trans* structure orientations in olefins and cyclic compounds. It was found to be unnecessary to account for the location of the alkyl substitutions on the benzene ring in the para, meta, and ortho positions in aromatics; the location of the alkyl branches along the chain for *iso*-paraffins and *iso*-olefins; the location of the double bond along the chain in olefins; and the locations of the alkyl substitutions and the ring size for naphthenes [15]. Albahri and George [16] also indicated that using two sets of structural groups, one for the aromatic ring in aromatics and the other for the cyclic ring in naphthenes, did not result in a significant improvement in the prediction performance.

Table 2 summarizes all the group types used for estimating the AIT in this work. The group definitions in this work basically follow those in Albahri and George's work [16]; however, some of them have been modified as follows. After carefully examining the AITs of organic compounds in our database, it was found that the condition of the carbon to which the halogen group is directly attached affects the AIT of the compound. This effect may be understood from the type of chemical bonding. A carbon that bears one or

more halogen groups forms a polar covalent bond with the halogen group instead of the pure covalent bonds that exist between the carbon atoms; this polar covalent bond will change the ability of the carbon atom to form free radicals. In Table 2, groups 18–21 are introduced to elucidate the effect of the addition of halogen atoms to paraffins; however, the effect of adding halogen groups to organic compounds was found to be different for compounds with nonring and ring structures. It was found that adding halogen atoms increased the AIT for nonring hydrocarbons, but decreased the AIT for ring-structure compounds. Thus, the halogen atoms were divided into nonring attachment groups (groups 22–24) and ring attachment groups (groups 25–27) in this study. Finally, to include organic compounds such as 1-hexyne and hexyl acetylene, group 9 is introduced in this work to explain their structure.

### 3. Developing the model

The simplest empirical model to predict the AIT of an organic compound is the multiple linear regression (MLR) model:

$$\text{AIT} = f_o + \sum_{i=1}^n v_i f_i \quad (1)$$

where  $n$  is the number of group types defined in Table 2;  $v_i$ , the number from group  $i$  in a molecule;  $f_i$  ( $i = 1, \dots, n$ ), the group contribution for the  $i$ th defined group; and  $f_0$ , the intercept of the fitting line. For a more accurate estimation, other nonlinear models could be adopted as the empirical model. In the literature, the nonlinear form of Eq. (2) has been announced to be the most suitable form for predicting AITs [15,16].

$$\text{AIT} = a' + b' \left( \sum_i v_i f_i \right) + c' \left( \sum_i v_i f_i \right)^2 + d' \left( \sum_i v_i f_i \right)^3 + e' \left( \sum_i v_i f_i \right)^4 \quad (2)$$

However, such a claim is debatable. First, the parameter  $b'$  in Eq. (2) is a redundant parameter. It is obvious that the following Eq. (3) has the same fitting and predictive abilities as Eq. (2), but the number of unknown parameters in Eq. (3) is one less than that in Eq. (2). It is well known that solving a nonlinear regression problem is usually time-consuming and can easily get bogged down at

local solutions; and a redundant parameter will aggravate these two problems. Second, it is obvious that Eq. (3) can be reasonably taken as a modified form from the conventional MLR model. So, the terms corresponding to parameters  $b$ ,  $c$ , and  $d$  in Eq. (3) could be considered as the correcting terms for the MLR model, and the results of the MLR model could be then taken as the start point (i.e., all  $f_i$ 's and  $a$  are taken from the MLR, and the others are set to be zero) to solve the nonlinear regression problem of Eq. (3). For a given iterative algorithm, a good start point always decreases its possibility of being bogged at local solutions and enhance its efficiency, thus Eq. (3) is superior to Eq. (2) for building a model in this respect.

$$\text{AIT} = a + \left( \sum_i v_i f_i \right) + b \left( \sum_i v_i f_i \right)^2 + c \left( \sum_i v_i f_i \right)^3 + d \left( \sum_i v_i f_i \right)^4 \quad (3)$$

To demonstrate the superiority of Eq. (3), data listed in Table A1 are used to build the models of Eqs. (2) and (3), respectively. Then, data listed in Table A2 are used to compare their predictive capability. The SGC groups' definitions listed in Table 2 and the linearized algorithm listed in Eq. (8) is employed to calculate the parameters

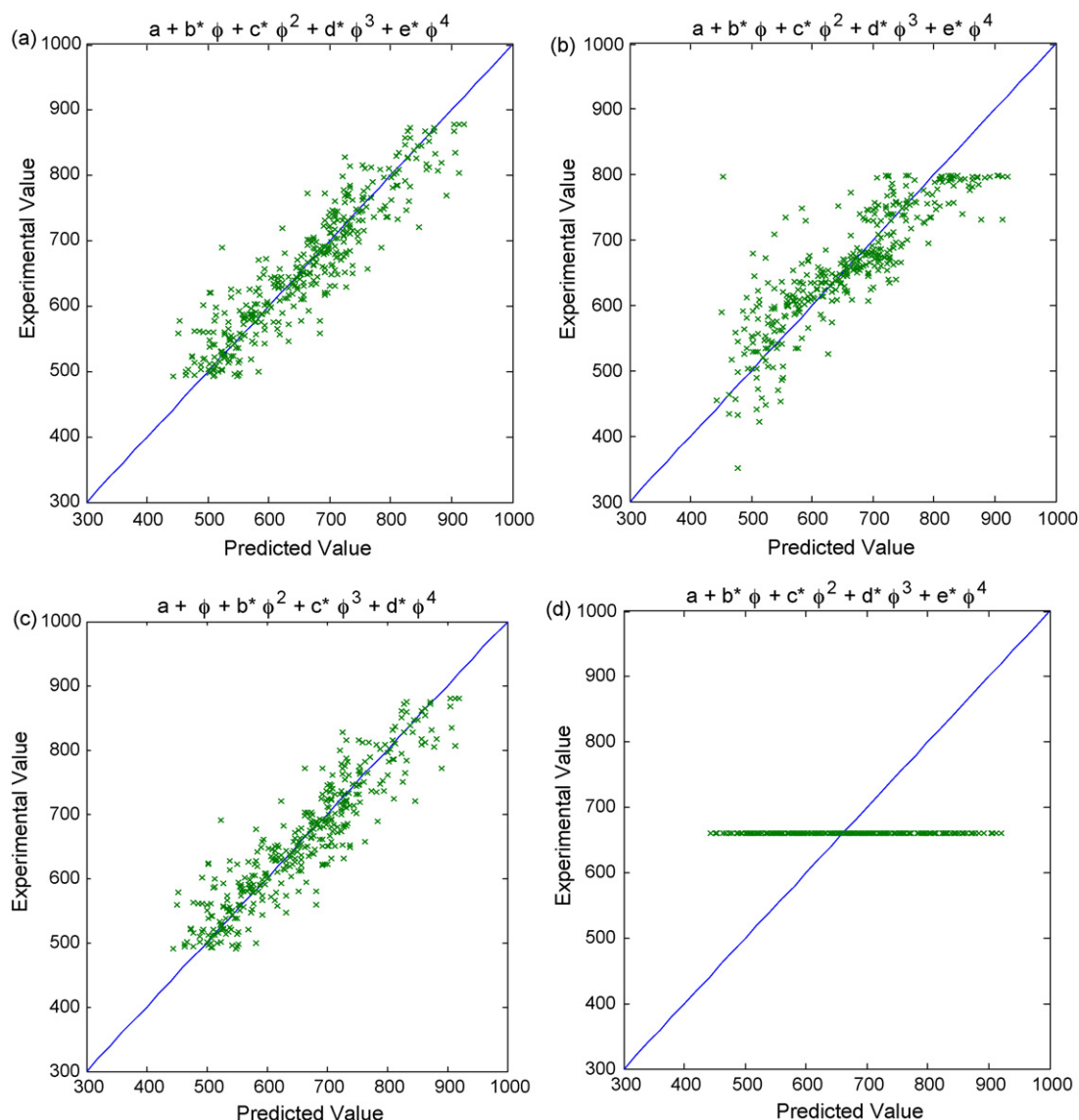


Fig. 1. Parity plot of Eq. (2) for four different initial guesses: (a) case 1; (b) case 2; (c) case 3; (d) case 4.

in these two equations, respectively. There are four different cases of initial conditions considered in present study, which includes: (1) case 1 – initial guesses of parameters  $a$  and  $f_i$ 's are set to be the values obtained from the results of a MLR model, and all other parameters are set to be zero; (2) case 2 – initial guesses of parameters  $a$  and  $f_i$ 's are set to be the values obtained from the results of a MLR model, and all other parameters are set to be one; (3) case 3 – initial guesses of parameters  $f_i$ 's are set to be the values obtained from the results of a MLR model, and all other parameters (including  $a$ ) are set to be zero; and (4) case 4 – initial guesses of all parameters are set to be zero.

Figs. 1–2 demonstrate the fitting abilities of the resulting models of these two equations for these four different cases of initial conditions, respectively. As shown in Fig. 1, the resulting models of Eq. (2) are very sensitive to the initial conditions employed, and it obviously get bogged down at the local solution in both cases 2 and 4. However, as shown in Fig. 2, the resulting models of Eq. (3) are almost the same in these four cases, which means that a model of this form is more robust to initial conditions than that of Eq. (2). Table 3 compares the fitting abilities and the predictive abilities of the resulting models of these two equations for these

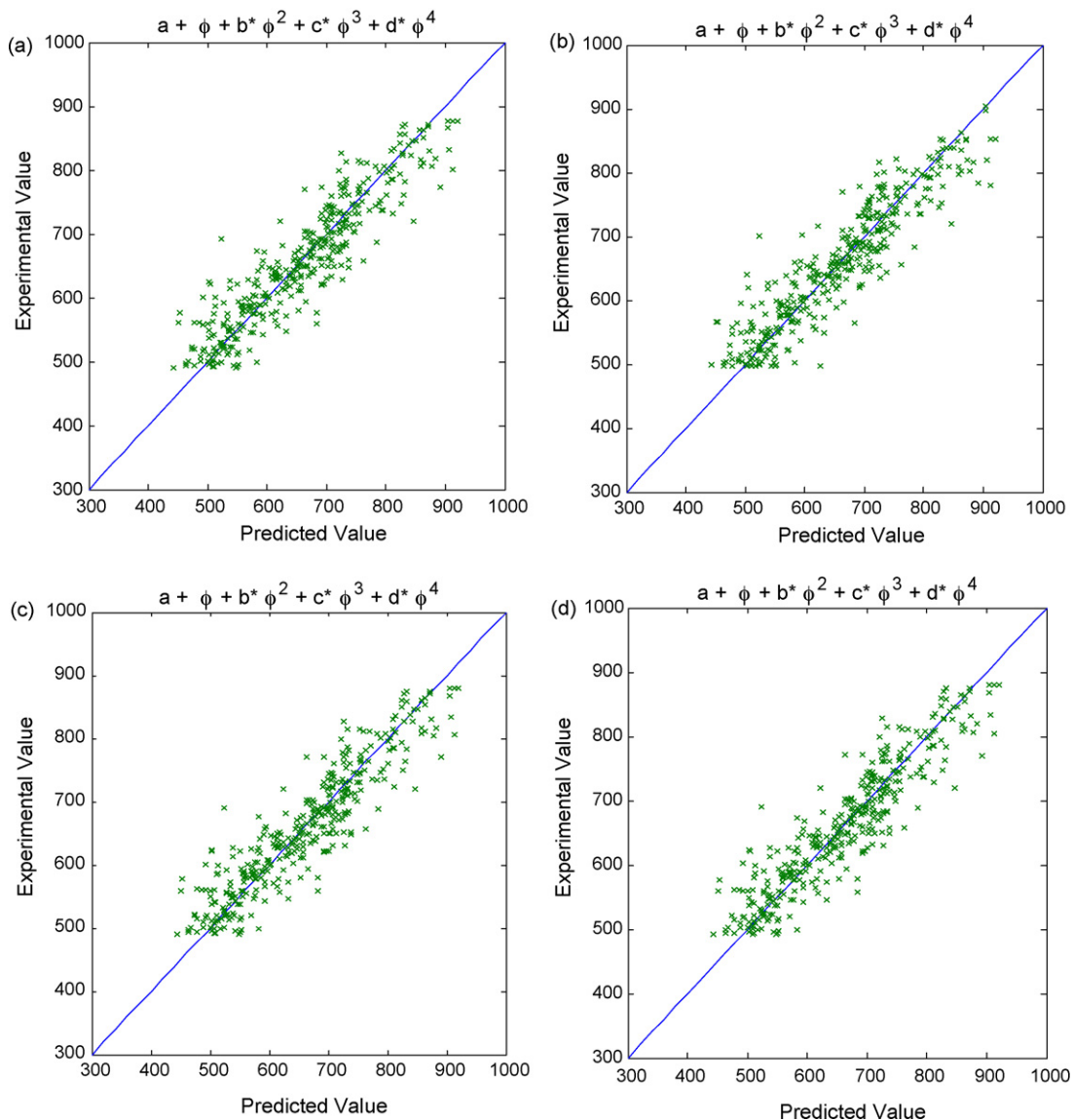
**Table 3**

Comparison of robustness to initial conditions between Eqs. (2) and (3)

Initial conditions	Eq. (2)		Eq. (3)	
	$R^2$	$Q^2$	$R^2$	$Q^2$
Case 1	0.8477	0.5345	0.8477	0.5362
Case 2	0.7367	-0.3191 <sup>a</sup>	0.8469	0.5336
Case 3	0.8478	0.5354	0.8472	0.5346
Case 4	0.0025	0.0018	0.8477	0.5347

<sup>a</sup> In this case, the predictive error of the seventh compound is found to be up to 1000 K.

four different initial conditions. As shown in Table 3, the resulting model of Eq. (2) for case 2 gives a very poor performance in prediction, although the fitting performance for this case seems to be of an acceptable value. In this case, it was found that the fitting error of the *N,N*-Dimethylbenzenamine in the testing set is more than 1000 K (the experimental value is 644.26 K, but the predictive value is -427.7 K), and the predictive error of this compound makes the  $Q^2$  to be an unreasonable negative value. This result shows that when redundant parameters are introduced into



**Fig. 2.** Parity plot of Eq. (3) for four different initial guesses: (a) case 1; (b) case 2; (c) case 3; (d) case 4.

an empirical model, there is an increased chance for the estimation process to draw noises and other spurious phenomena from the training data into the resulting model, which always decreases the predictive capability of the resulting model.

The sum of squared errors is the usual index to evaluate the feasibility of a given model. In this study, the quantitative measure of goodness of fit is given by the explained variation in the training set ( $R^2$ ), whereas the predictive capability, on the other hand, is given by the predicted variation in the validation set ( $Q^2$ ). These two indices are defined as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^K (y_i - \hat{y})^2}{\sum_{i=1}^K (y_i - \bar{y})^2}, \quad \text{for the training set}$$

$$Q^2 = 1 - \frac{\sum_{i=1}^K (y_i - \hat{y})^2}{\sum_{i=1}^K (y_i - \bar{y})^2}, \quad \text{for the validation set}$$

where  $y_i$  is the  $i$ th sample measurement;  $\hat{y}_i$ , the predicted value of the  $i$ th sample;  $\bar{y}$ , the average of all sample measurements. Usually, the  $R^2$  and  $Q^2$  vary differently with increasing model complexity (i.e., number of parameters in a model).  $R^2$  is inflationary and approaches unity as the model complexity increases. Hence, it is not sufficient to only have a high  $R^2$  for a practical model. The goodness of predictive capability  $Q^2$ , on the other hand, is not inflationary and will not automatically approach unity with increasing model complexity. Commencing with a very simple model,  $Q^2$  will increase with model complexity. However, at a certain degree of complexity,  $Q^2$  will reach a plateau and subsequently reduce. Usually, the point at which  $Q^2$  reaches a plateau is the trade-off point between the fitting and predictive capability. Besides aforementioned method, a common alternative to determine model complexity is to examine the hypothesis test for a given parameter in that model. In this study, both  $Q^2$ -statistics and hypothesis testing are used to determine whether the complexity of an empirical model is adequate or not.

The determination of the least-squares solution of Eq. (3) is a nonlinear regression problem. There are many different methods for solving this problem, and the solutions from different methods may differ slightly. In this study, the linearized algorithm and asymptotic confidence intervals are adopted [19]. The following paragraph briefly discusses this method.

Consider the following nonlinear empirical model:

$$y_i = f(x_i; \theta) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

where  $y_i$  is the AIT measurement of  $i$ th sample;  $n$ , the number of compounds in the training set;  $x_i \in R^K$ , the number of group  $i$  in a molecule;  $K$ , the number of function groups defined in Table 2;  $\theta \in R^p$ , the parameter vector (includes  $f_i$ ,  $a$ ,  $b$ ,  $c$ , and  $d$ ) in the empir-

ical model; and  $P$ , the number of parameters in the model.  $\varepsilon_i$  is the measurement error and is assumed to be i.i.d.  $N(0, \sigma^2)$ . Let us define

$$y = [y_1, y_2, \dots, y_n]^T \quad (4)$$

$$f_i(\theta) \equiv f(x_i; \theta) \quad i = 1, 2, \dots, n \quad (5)$$

$$f(\theta) \equiv [f_1(\theta), f_2(\theta), \dots, f_n(\theta)]^T \quad (6)$$

$$F(\theta) \equiv \begin{bmatrix} \frac{\partial f_1(\theta)}{\partial \theta_1} & \frac{\partial f_1(\theta)}{\partial \theta_2} & \dots & \frac{\partial f_1(\theta)}{\partial \theta_p} \\ \frac{\partial f_2(\theta)}{\partial \theta_1} & \frac{\partial f_2(\theta)}{\partial \theta_2} & \dots & \frac{\partial f_2(\theta)}{\partial \theta_p} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n(\theta)}{\partial \theta_1} & \frac{\partial f_n(\theta)}{\partial \theta_2} & \dots & \frac{\partial f_n(\theta)}{\partial \theta_p} \end{bmatrix} = \left[ \left( \frac{\partial f_i(\theta)}{\partial \theta_j} \right) \right] \quad (7)$$

Let  $\theta^*$  be the true value of  $\theta$ , and  $\hat{\theta}$ , which is the convergent solution of the following iterative Eq. (8), be the estimate of  $\theta$ .

$$\theta^{(a+1)} = \theta^{(a)} + (F^T(\theta^{(a)})F(\theta^{(a)}))^{-1} F^T(\theta^{(a)})[y - f(\theta^{(a)})] \quad (8)$$

It has been shown in the literature that an approximate, for large  $n$ ,  $100(1 - \alpha)\%$  confidence interval for  $\theta_r$  is as follows [19].

$$\hat{\theta}_r \pm t_{n-p}^{\alpha/2} s \sqrt{\hat{C}^r} \quad (9)$$

where  $\theta_r$  is the  $r$ th element of  $\theta$ ;  $t_{n-p}$  is the  $t$ -distribution with  $(n - p)$  degrees of freedom;  $s$  is the sample standard deviation; and  $[(\hat{C}^r)] = \hat{C}^{-1}$  with  $\hat{C} = \hat{F}^T(\hat{\theta})\hat{F}(\hat{\theta})$ .

#### 4. Results and discussion

For a given model, the group contributions ( $f_i$ ) of different molecular groups and the other parameters in the model were solved to minimize the squared error for the 400-compound training set. Different models, including MLR and polynomial models of degree 2–4, were considered in this study. The parameters in the MLR model were calculated by the classical least-squares method. For nonlinear models, Eqs. (8) and (9) were used to solve model parameters and estimate their corresponding 95% confidence intervals, respectively. Table 2 summarizes all the group contributions for the interested models in this work, and the other parameters (i.e.,  $a$ ,  $b$ ,  $c$ , and  $d$ ) for the corresponding models are listed in Table 4. Table 5 lists the 95% confidence intervals for the parameters  $a$ ,  $b$ ,  $c$ , and  $d$  in all the nonlinear models. The parameters in Tables 2 and 4 were then used to calculate the predicted AITs of the corresponding model for the 83 compounds in the validation set. The fitting abilities and predictive abilities of these models were then calculated according to the predicted AITs. The fitting abilities of the different models for estimating AIT are summarized in Table 6, and

**Table 4**  
Parameters for polynomial models of different degrees

Coefficients	MLR	Degree 2	Degree 3	Degree 4
$a$	731.4902	771.1828	750.3065	754.0344
$b$	–	8.5082E–04	–8.6444E–04	–7.5627E–04
$c$	–	–	–4.5604E–06	–5.0831E–06
$d$	–	–	–	–2.4496E–09

**Table 5**  
95% confidence intervals for parameters  $a$ ,  $b$ ,  $c$  and  $d$  in different models

Model degree	$a_-$	$a_+$	$b_-$	$b_+$	$c_-$	$c_+$	$d_-$	$d_+$
2	7.17984E+02	8.24381E+02	6.76307E–04	1.02534E–03	–	–	–	–
3	7.07108E+02	7.93505E+02	–1.68051E–03	–4.83678E–05	–6.15028E–06	–2.97048E–06	–	–
4	7.08928E+02	7.99141E+02	–1.67053E–03	1.57982E–04	–7.77587E–06	–2.39041E–06	–9.58518E–09	4.68588E–09

**Table 6**  
Fitting ability for different types of models

Model	$R^2$	Max err (K)	Avg err (K)	Max err (%)	Avg err (%)
MLR	0.8266	178	34	34	5.3
Degree 2	0.8361	163	33	31	5.1
Degree 3	0.8474	169	32	32	4.9
Degree 4	0.8478	168	31	32	4.9
Albahri <sup>a</sup>	0.8464	166	28	31	4.2
Albahri and George <sup>b</sup>	0.7900	–	58	125	9.2

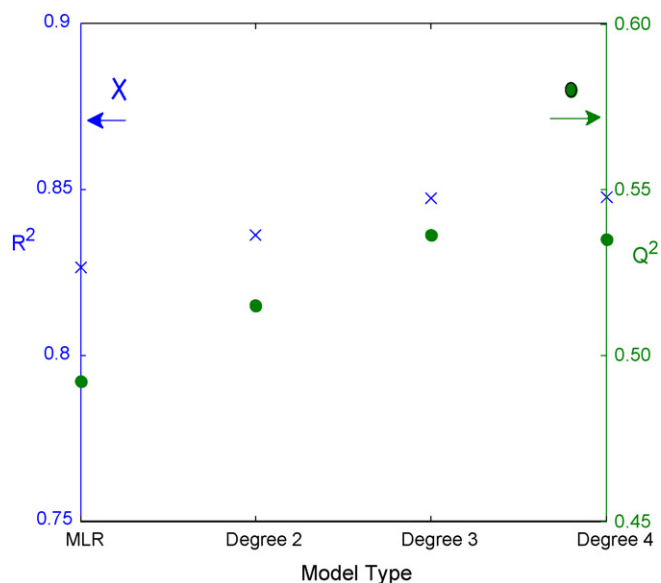
<sup>a,b</sup>These values are taken from the original papers.

**Table 7**  
Predictive capability for different models

Model	$Q^2$	Max err (K)	Avg err (K)	Max err (%)	Avg err (%)
MLR	0.4921	184.1	75.2	47.4	11.9
Degree 2	0.5151	189.0	72.3	45.4	11.3
Degree 3	0.5361	179.6	69.8	45.9	11.0
Degree 4	0.5349	178.5	69.7	45.5	11.0

the predicting abilities of the corresponding models are listed in Table 7.

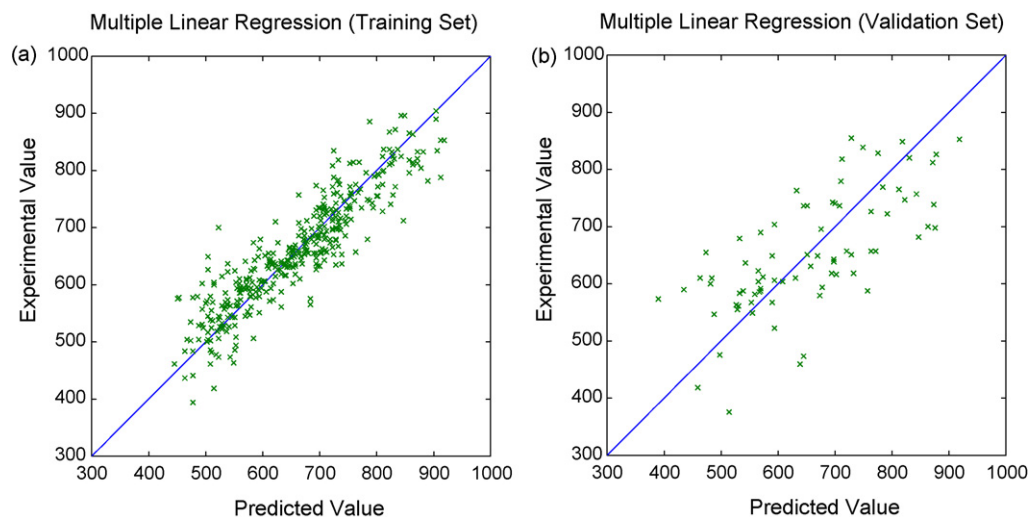
It is very possible that a model fits well for training data but gives poor predictive performance for testing data. Thus, when developing a suitable model for prediction, there is a need to balance the fitting ability and predictive capability of the model. The fitting ability measures the ability to mathematically reproduce the data of the training set, and the predictive capability gauges the reliability of the predicted outcomes of other experiments. Usually, this compromise is achieved through the model complexity. Both  $R^2$  and  $Q^2$  for different models are plotted against their model complexity in Fig. 3. It was found that the predictive capability ( $Q^2$ ) reached a maximum for a model with a polynomial of degree 3, and the fitting ability ( $R^2$ ) for a polynomial of degree 4 was only a little better than that for a polynomial of degree 3. Thus, the nonlinear model suggested by Albahri, a polynomial model of degree 4, might be an overfitting model for predicting the AITs of organic compounds. Moreover, as shown in Table 5, the 95% confidence interval for parameter  $d$  in the polynomial model of degree 4 contains the zero value. This implies that we cannot reject the hypothesis  $d=0$  with 95% confidence, and also indicates that the model of degree 4 is an overfitting model. For a model of degree 3, although the estimate of  $c$  is a very small number ( $-4.5604 \times 10^{-6}$ ), the 95% confidence interval of  $c$  does not include the zero value; and the hypothesis  $c=0$  will



**Fig. 3.** Adequacy of model complexity.

be rejected with 95% confidence in this case. Thus, from the method of hypothesis testing, it was also concluded that the most adequate model is a polynomial of degree 3. Figs. 4–7 show the experimental AITs vs. the predicted values for all the interested models in both the training and validation sets. It could be found in these figures that the drawing at lower experimental AIT part become more and more flat as the model complexity increases. This means that the compounds with large fitting errors will tend to concentrate on the compounds with the lower experimental AIT as the model complexity increases. Usually, the occurring of a specific pattern of fitting errors is an evidence to indicate that the corresponding model is possibly overfitting; thus, the empirical model of degree 4 is possibly an overfitting model. It should be noted that this phenomenon that the fitting errors tend to concentrate on a specific region was also found in Albahri and George's original work [16]. From all aforementioned facts, it is then concluded by us that a model of degree 3 is more adequate than a model of degree 4 for predicting the AIT of an organic compound.

A comparison of the fitting ability between this study and the other two studies in the literature is also listed in Table 6. In



**Fig. 4.** Parity plot of the MLR model: (a) training set; (b) validation set.

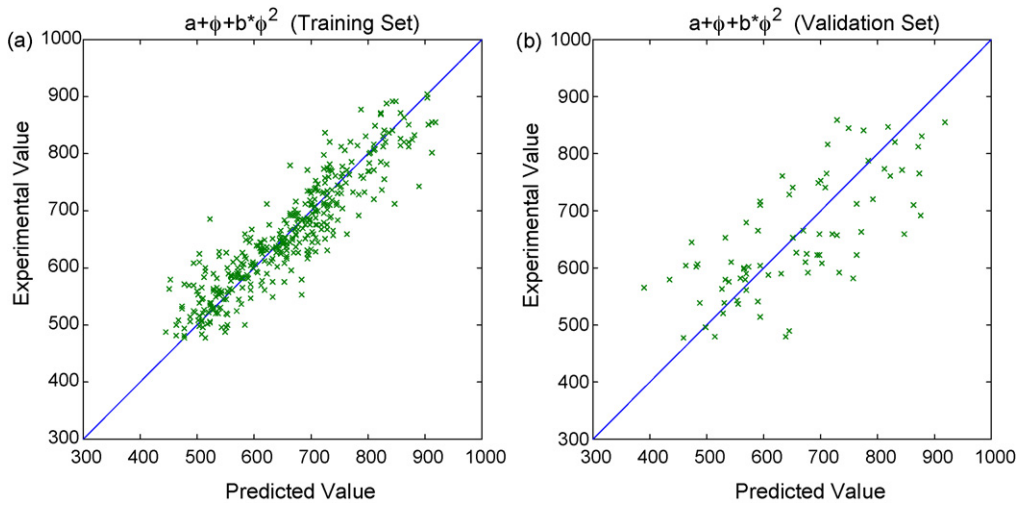


Fig. 5. Parity plot for the polynomial model of degree 2: (a) training set; (b) validation set.

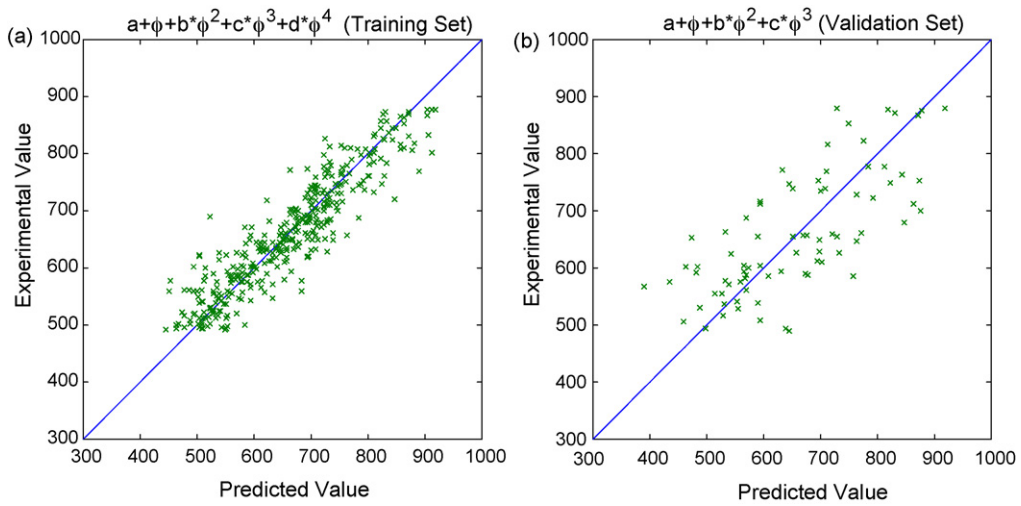


Fig. 6. Parity plot for the polynomial model of degree 3: (a) training set; (b) validation set.

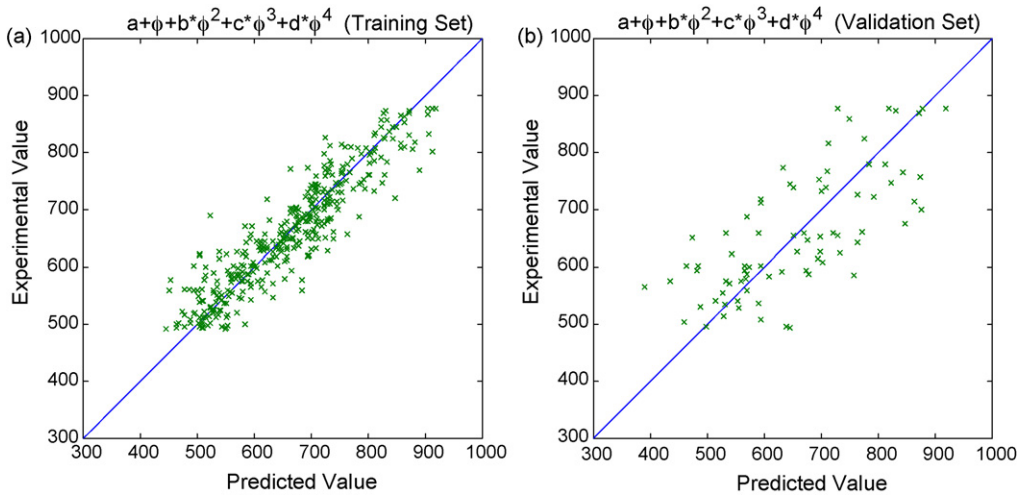


Fig. 7. Parity plot for the polynomial model of degree 4: (a) training set; (b) validation set.



Albahri's work, the  $R^2$  of their empirical model (of degree 4) is 0.846 [15]. However, that empirical model was conducted from 137 pure hydrocarbons, and hence, such a model cannot be applied to the case of organic compounds containing hetero-atoms. Albahri and George's work explored 490 organic compounds containing hetero-atoms and the  $R^2$  of the corresponding empirical model (of degree 4) is 0.790 [16]. In present work, the results are drawn from 483 organic compounds containing hetero-atoms, and the  $R^2$  of the proposed method is 0.847—which is better than that of Albahri and George's work and is comparable with that of the Albahri's work for only pure hydrocarbons. Moreover, the average error and maximum error of the proposed model are 4.9% and 32%, respectively; and these two values are 9.2% and 125% for the model proposed by Albahri and George. Since AIT is a safety related parameter, the improvement in maximum error should be notified.

As it was shown in Table 6, the  $R^2$  of the MLR model in present work is 0.827, which is also better than that of Albahri and George's work; thus, it was deemed by us that much of the improvement in this work is attributed to the revised group definitions instead of the type of empirical model chosen. In fact, we also attempted to use many other types of empirical models to improve the prediction performance, but the improvement was limited. However, as the database in present study is different from that in Albahri and George's work, aforementioned conclusion that much of the improvement in this work is attributed to the revised group definition still needs more evidences to support it. To make a more objective comparison, the group definitions and predictive equation of Albahri and George's work have been applied to our database to obtain the corresponding predictive AIT value for all compounds. Because the predictive equation proposed by Albahri and George is obtained from their own database, some pretreatments in our database are needed to avoid this possible bias. First, the compounds that could not be decomposed according to their group definitions are excluded out; and thus, there are each three compounds dropped from the training set and validation set, respectively. Second, as Albahri and George announced that the maximum error is 125% and the average error is 9.2% in their training set, the compounds of which the absolute predictive error are larger than 150 K by their predicted equation are also dropped from the database to let the overall performance of the predicted results meets aforementioned two requirements. Thus, we delete, according to this criterion, 10 compounds from

our training set and 9 compounds from our validation set, respectively. With these pretreatments, the average errors of Albahri and George's model for the trimmed training set and the trimmed validation set are 7.20% and 9.40%, respectively; and the maximum errors are 26.95% and 26.73%, respectively. Fig. 8 shows the fitting performance of Albahri and George's model for both the trimmed training set and the trimmed validation set. It could be found that the drawings are of two flat parts at the lower experimental AIT zone and higher experimental AIT zone. As this phenomenon was also found in their original work, the fitting results for these two trimmed sets are of similar characteristics with their original work [16]. With above efforts, it is then assumed by us that the effects of the bias of establishing the model parameters from different database have been moderated. With aforementioned pretreatment, the fitting performance, i.e., the  $R^2$  value, of their models are 0.7094 and 0.6609 for the trimmed training set and the trimmed validation set, respectively. It is obvious that the resulting fitting performance in both trimmed sets is still inferior to that of the MLR case in present study. Thus, our previous conclusion that much of the improvement in present study is attributed to the revised group definitions is supported by these results.

To assess whether a model is capable of predicting AITs or not, the number of compounds in the validation set is very important. In many commercial softwares, it is a common practice that the numbers of observations in the validation set should be at least one-fifth the numbers of observations in the training set to avoid underestimating the predictive error [20]. However, this ratio is only 20/470 in Albahri and George's work, which is deemed to be too small to derive a reasonable conclusion about the predictive performance of their model. To show this point, we take 2-hydroxy-1-ethylaziridine (CC1CN1O) as an example to explain the case of underestimating the predictive error for their model. By using their group definitions and predictive model, the estimated AIT for 2-hydroxy-1-ethylaziridine is about 1630 K, but the reported experimental value is only 607 K. Thus, the prediction error is more than 1000 K in this example, which means that if this compound is included in their validation set, the predictive capability of their model will drastically decrease. So, it is obvious that the predictive error is possibly underestimated in their work.

As shown in Table 7, the  $Q^2$ , average error and maximum error for the present work are 0.5361, 11.0% and 45.9%, respectively. A rule of thumb for developing a practical model is: the difference

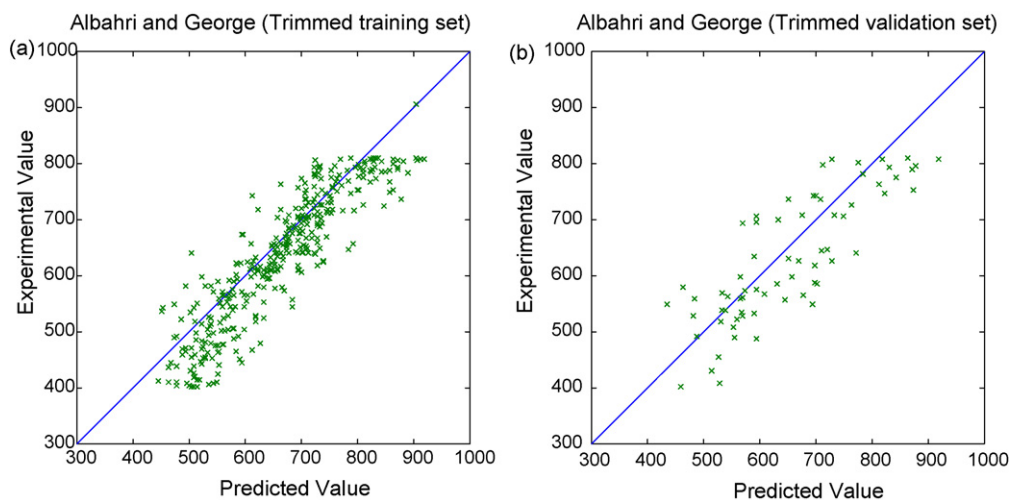


Fig. 8. Parity plot for Albahri and George's model: (a) trimmed training set; (b) trimmed validation set.

between  $R^2$  and  $Q^2$  must not be too large and preferably not exceeding 0.2–0.3. Moreover, a  $Q^2 > 0.5$  is regarded as good and a  $Q^2 > 0.9$  as excellent [21]. In this study, the  $Q^2$  value for the 83-compound validation set was 0.5361 and the  $R^2$  value for the 400-compound training set was 0.8474; thus, the proposed method coincides with this rule of thumb and can be taken as a reasonable model for estimating the AIT of an unavailable or unknown compound in practical applications.

## 5. Conclusions

A predictive model for AITs was proposed based on the SGC approach. The proposed equation to predict AITs may be expressed as Eq. (10). This model includes 45 molecular function groups and is a polynomial of degree 3. This model was deduced from a 400-compound training set. The fitting ability of the proposed model is about 0.8474, with an average error of 32 K and an average error percentage of 4.9%. The predictive capability of the model was demonstrated on 83 compounds which were not included in the original training set. The predictive capability of the proposed model is about 0.5361, with average error 70 K and an average error percentage of 11.0%

$$\text{AIT} = 750.3065 + \left( \sum_i v_i f_i \right) - 8.644 \times 10^{-4} \times \left( \sum_i v_i f_i \right)^2 - 4.5604 \times 10^{-6} \times \left( \sum_i v_i f_i \right)^3 \quad (10)$$

As compared to Albahri and George's work, the proposed model exhibits better performance in terms of  $R^2$ . It was also found that much of the improvement may be attributed to the modification of the group definitions and not the type of empirical model chosen. As mentioned earlier, the addition of halogen atoms to nonring hydrocarbons and ring-structure compounds has different effects on their AITs. Thus, in this study, 14 new groups were introduced to discriminate this effect for halogen compounds.

In this work, the average fitting error for the 400-compound training set was 32 K, and the average prediction error was 70 K for the 83-compound validation set. Because the average experimental error in measuring the AIT is deemed to be greater than 30 K in the literature, the proposed method could offer a reasonable estimate of the AIT value for the organic compounds in the training set and could also approximate the AITs of compounds that were unknown or whose AITs were not readily available to within a reasonable degree.

## Acknowledgements

The authors would like to thank both the National Science Council and China Medical University of the ROC for supporting this study financially under grant #NSC 96-2221-E-039-001 and #CMU 96-152, respectively.

## Appendix A

See Tables A1 and A2 .

**Table A1**  
Experimental values and predicted values of the compounds in the training set

	Compound name	Exp. value	Reference	MLR	Degree 3
1	Butane	645	[11]	643.42	643.16
2	Pentane	538	[11]	622.07	614.17
3	Hexane	513	[11]	600.72	586.71
4	Heptane	486	[11]	579.36	561.42
5	2-Methylpropane	733.15	[3]	667.50	680.46
6	2-Methylbutane	693.15	[3]	646.14	650.46
7	3-Methylpentane	551.15	[3]	624.79	621.19
8	2,2-Dimethylpropane	723.15	[3]	697.41	707.33
9	2,2-Dimethylbutane	678	[4]	676.06	677.22
10	2,3-Dimethylpentane	610.37	[3]	627.51	628.30
11	2,2,3-Trimethylbutane	685	[11]	678.78	684.64
12	1-Pentene	571	[11]	608.59	598.09
13	1-Heptene	536	[11]	565.88	548.06
14	1-Octene	523	[4]	544.53	527.50
15	1-Decene	508.15	[3]	501.82	498.46
16	1,3-Pentadiene	613	[11]	571.74	565.84
17	2-Methyl-1-pentene	579	[11]	594.57	589.28
18	2,4,4-Trimethyl-1-pentene	693	[3]	627.20	620.88
19	Cyclopentane	593	[11]	608.26	603.74
20	Methylcyclopentane	602.04	[3]	606.22	602.61
21	Ethylcyclohexane	535.37	[3]	560.22	551.89
22	<i>n</i> -Propylcyclohexane	521.15	[3]	538.87	530.74
23	<i>trans</i> -1,2-Dimethylcyclohexane	577.15	[3]	579.53	575.08
24	Dicyclohexyl	518.15	[3]	477.02	496.27
25	Decalin	541	[11]	526.31	522.96
26	Hydrindane	569	[11]	641.38	629.21
27	Cyclopentene	668.15	[3]	698.69	702.65
28	Cyclohexene	583.15	[3]	674.04	672.68
29	Benzene	771	[4]	768.25	786.27
30	Toluene	755	[11]	761.44	774.02
31	Ethylbenzene	705.37	[3]	740.09	745.98
32	<i>n</i> -Propylbenzene	729.15	[3]	718.74	716.63
33	<i>n</i> -Butylbenzene	685.37	[3]	697.38	686.60
34	1,3-Dimethylbenzene	800.93	[3]	754.63	761.36
35	1,4-Diethylbenzene	703.15	[3]	711.93	702.89
36	Biphenyl	813.15	[3]	836.76	833.13
37	Naphthalene	813	[11]	800.66	807.52

Table A1 (Continued)

	Compound name	Exp. value	Reference	MLR	Degree 3
38	1-Methylnaphthalene	802.04	[3]	793.86	796.18
39	Anthracene	828	[11]	833.08	826.77
40	Ethanol	673	[11]	677.17	687.74
41	1-Propanol	644.26	[3]	655.82	657.66
42	1-Butanol	616	[3]	634.47	628.16
43	2-Butanol	663	[4]	658.54	665.04
44	<i>tert</i> -Butanol	733	[11]	709.81	721.88
45	Cyclohexanol	573.15	[3]	593.97	588.85
46	Benzyl alcohol	709.26	[3]	752.49	759.86
47	1-Hexanol	558	[3]	591.76	573.45
48	Allyl alcohol	643	[11]	642.34	640.51
49	Dimethyl ether	623.15	[3]	631.41	629.99
50	Dibutyl ether	467.59	[3]	503.29	499.77
51	Methyl vinyl ether	560.15	[3]	596.58	586.01
52	Diphenyl ether	891.15	[3]	782.04	774.04
53	Propylene oxide	738.15	[3]	652.84	660.27
54	Propionaldehyde	500	[11]	566.60	560.55
55	Butyraldehyde	503.15	[3]	545.25	538.18
56	Acetophenone	833	[11]	771.76	781.65
57	2-Butanone	677	[11]	675.09	681.59
58	2-Pentanone	725.15	[3]	653.74	651.57
59	Cyclohexanone	693.15	[3]	654.96	662.18
60	Acetic acid	737	[11]	716.28	731.15
61	Butyric acid	718	[11]	673.58	671.24
62	Pentanoic acid	673.15	[3]	652.23	641.39
63	Acrylic acid	711.15	[3]	681.45	683.97
64	Dipropylamine	572.15	[3]	599.96	585.03
65	Diphenylamine	907.04	[3]	836.00	831.82
66	2-Aminoethanol	673	[11]	660.69	663.03
67	1-Chlorobutane	523	[11]	700.92	692.42
68	Acetyl chloride	663.15	[3]	756.97	769.12
69	Chlorobenzene	863	[4]	811.25	853.79
70	1-Bromobutane	538	[4]	636.30	627.99
71	Bromobenzene	838.15	[3]	838.15	837.56
72	Ethyl formate	708	[11]	722.73	733.82
73	Ethyl acetate	700	[3]	700.05	710.09
74	Propyl acetate	708	[11]	678.70	679.99
75	Butyl acetate	653	[11]	657.35	650.00
76	Isobutyl acetate	696	[3]	681.42	687.42
77	Methyl propionate	728	[11]	700.05	710.09
78	Ethyl propionate	718	[11]	678.70	679.99
79	Methyl butyrate	728	[11]	678.70	679.99
80	Methyl benzoate	783	[4]	796.72	805.78
81	Ethyl benzoate	763.15	[3]	775.37	780.24
82	Butyl benzoate	708	[11]	732.66	723.49
83	Ethyl acrylate	655.93	[3]	665.22	662.62
84	2-Methylpentane	579.26	[3]	624.79	621.19
85	2,2,4-Trimethylbutane	680	[11]	654.71	647.27
86	<i>trans</i> -2-Hexene	528	[11]	563.87	555.74
87	<i>trans</i> -2-Pentene	558	[3]	585.22	580.42
88	1,3-Hexadiene	593	[11]	550.39	542.81
89	1,5-Hexadiene	618	[11]	573.75	557.74
90	2-Methylpropene	738.15	[3]	637.27	646.02
91	3-Methyl-1-butene	638.15	[3]	632.66	633.43
92	4-Methyl-1-pentene	577	[11]	611.31	604.88
93	2-Ethyl-1-butene	597	[11]	594.57	589.28
94	2,3-Dimethyl-1-butene	633.15	[3]	618.64	623.97
95	2,3,3-Trimethyl-1-butene	656	[11]	648.56	650.14
96	2,4,4-Trimethyl-2-pentene	581	[3]	603.84	602.22
97	Ethylcyclopentane	533.5	[3]	584.87	575.97
98	Propylcyclopentane	542.15	[3]	563.51	551.77
99	<i>n</i> -Hexylcyclopentane	501	[11]	499.46	500.15
100	Isopropylcyclohexane	556	[4]	562.94	557.60
101	Butylcyclohexane	519.15	[3]	517.52	513.30
102	Isobutylcyclohexane	547	[11]	541.59	535.63
103	<i>sec</i> -Butylcyclohexane	550	[11]	541.59	535.63
104	<i>tert</i> -Butylcyclohexane	615	[11]	592.86	579.68
105	<i>trans</i> -1,3-Dimethylcyclohexane	579	[3]	579.53	575.08
106	<i>trans</i> -1,4-Dimethylcyclohexane	577	[3]	579.53	575.08
107	1,3,5-Trimethylcyclohexane	587	[11]	577.48	574.04
108	4-Isopropyl-1-methylcyclohexane	579	[11]	560.90	556.64
109	Cyclodecane	508	[11]	485.04	500.89
110	Isobutylbenzene	700.93	[3]	721.46	723.95
111	<i>sec</i> -Butylbenzene	690.93	[3]	721.46	723.95
112	<i>tert</i> -Butylbenzene	723.15	[3]	772.73	777.77

Table A1 (Continued)

	Compound name	Exp. value	Reference	MLR	Degree 3
113	1,2-Dimethylbenzene	737.04	[3]	754.63	761.36
114	1,4-Dimethylbenzene	802.04	[3]	754.63	761.36
115	1,2,3-Trimethylbenzene	743.15	[3]	747.83	748.34
116	1,2,4-Trimethylbenzene	788.15	[3]	747.83	748.34
117	1-Methyl-2-ethylbenzene	721	[11]	733.28	732.64
118	1-Methyl-3-ethylbenzene	753.15	[3]	733.28	732.64
119	1-Methyl-4-ethylbenzene	748.15	[3]	733.28	732.64
120	1,2-Diethylbenzene	677	[11]	711.93	702.89
121	1,3-Diethylbenzene	723.15	[3]	711.93	702.89
122	1-Methyl-3,5-diethylbenzene	734	[11]	705.12	689.07
123	2-Ethylbiphenyl	722	[11]	808.60	799.87
124	2-Propylbiphenyl	725	[11]	787.24	773.77
125	2-Butylbiphenyl	706	[11]	765.89	745.72
126	Diphenylmethane	759	[11]	815.40	811.02
127	1-Ethyl-naphthalene	754	[11]	772.50	769.76
128	Tetralin	657.04	[3]	618.28	611.97
129	Methanol	728	[4]	698.53	717.75
130	3-Pentanol	638	[11]	637.19	635.34
131	2-Methyl-1-butanol	658.15	[3]	637.19	635.34
132	2-Propanol	672.04	[3]	679.90	695.16
133	2-Methyl-1-propanol	678	[11]	658.54	665.04
134	3-Methyl-1-butanol	623.15	[3]	637.19	635.34
135	2-Pentanol	616.48	[3]	637.19	635.34
136	2-Methyl-2-butanol	708	[11]	688.46	691.92
137	2,2-Dimethyl-1-propanol	693	[11]	688.46	691.92
138	4-Methyl-2-pentanol	613	[11]	639.91	642.58
139	1-Heptanol	555	[3]	570.41	549.52
140	4-Heptanol	568	[11]	594.48	579.75
141	2-Octanol	538	[11]	573.13	555.15
142	2-Ethyl-1-hexanol	560.93	[3]	573.13	555.15
143	1-Nonanol	533	[11]	527.70	511.71
144	1-Decanol	523	[11]	506.35	499.12
145	Ethylene glycol	673.15	[3]	668.22	672.30
146	1,2-Propanediol	694.26	[3]	670.94	679.72
147	Glycerol	673	[11]	661.99	664.30
148	2-Ethyl-1,3-hexanediol	633	[11]	588.25	572.95
149	2,2-Dimethyl-1,3-propanediol	672	[3]	679.50	676.48
150	3,5-Dimethylphenol	828	[11]	813.55	867.67
151	2,4-Dimethylphenol	872	[3]	813.55	867.67
152	2,4-Dimethyl-3-pentanol	668	[11]	642.64	649.88
153	Methoxybenzene	748	[4]	706.72	703.18
154	Dipentyl ether	444	[4]	460.59	489.67
155	Butyl vinyl ether	528	[11]	532.52	519.46
156	Ethylene oxide	702.04	[3]	654.89	661.48
157	1,2-Epoxyethylbenzene	811	[11]	728.16	733.63
158	Isobutyraldehyde	534	[11]	569.32	566.52
159	2-Propenal	573	[11]	553.12	547.25
160	Crotonaldehyde	553	[11]	508.40	515.34
161	2-Ethylcrotonaldehyde	523	[11]	473.03	498.22
162	3-Pentanone	725.37	[3]	653.74	651.57
163	Propionic acid	713	[11]	694.93	701.37
164	Isobutyric acid	733	[11]	697.65	708.76
165	Isopentanoic acid	689.15	[3]	676.30	678.66
166	Hexanoic acid	653.15	[3]	630.87	612.48
167	2-Methylpentanoic acid	651	[11]	654.95	648.68
168	Heptanoic acid	571	[3]	609.52	585.14
169	Decanoic acid	570	[3]	545.46	518.90
170	Dodecanoic acid	503	[11]	502.76	494.32
171	Tetradecanoic acid	508	[11]	460.05	491.37
172	Hexadecanoic acid	513	[11]	417.35	515.16
173	<i>o</i> -Phthalic acid	863	[11]	814.95	810.35
174	2,2-Dimethylpropionic acid	723	[11]	727.57	735.22
175	2-Ethylbutyric acid	663	[11]	654.95	648.68
176	2-Aminobiphenyl	725	[11]	834.82	827.27
177	1,2-Propanediamine	689	[11]	655.89	661.17
178	<i>DL</i> -1-Amino-2-propanol	647.04	[3]	663.42	670.43
179	Diisopropanolamine	647	[11]	630.20	625.75
180	Triisopropanolamine	593	[11]	579.52	567.12
181	2-Diethylaminoethanol	593	[11]	589.27	574.26
182	Benzyl chloride	858.15	[3]	818.94	815.54
183	1,1,1-Trichloroethane	810.15	[3]	799.00	801.62
184	Trichloroethylene	693	[11]	774.15	772.52
185	Bis(2-ethoxyethyl)ether	478	[11]	393.86	521.09
186	<i>n</i> -Hexyl Cellosolve	553	[11]	494.34	495.07
187	Methyl formate	729.26	[3]	744.09	762.48

Table A1 (Continued)

	Compound name	Exp. value	Reference	MLR	Degree 3
188	Propyl formate	708	[11]	701.38	704.10
189	<i>n</i> -Butyl formate	595.37	[3]	680.03	673.98
190	Isopropyl formate	713	[11]	725.46	741.00
191	Methyl acetate	748	[11]	721.41	739.65
192	Isopropyl acetate	698	[11]	702.78	717.44
193	Pentyl acetate	633.15	[3]	636.00	620.75
194	Isopentyl acetate	653	[11]	660.07	657.35
195	Hexyl acetate	528	[11]	614.64	592.88
196	<i>tert</i> -Butyl acetate	708	[11]	732.69	743.67
197	<i>sec</i> -Butyl acetate	683	[11]	681.42	687.42
198	<i>n</i> -Decyl acetate	488	[11]	529.23	508.03
199	Vinyl acetate	698	[11]	686.57	692.73
200	Allyl acetate	647.04	[3]	665.22	662.62
201	Phenyl acetate	858	[11]	796.72	805.78
202	<i>n</i> -Propyl propionate	703	[11]	657.35	650.00
203	Isopropyl propionate	698	[11]	681.42	687.42
204	Butyl propionate	658	[11]	636.00	620.75
205	Isobutyl propionate	708	[11]	660.07	657.35
206	Ethyl butyrate	713	[11]	657.35	650.00
207	Propyl butyrate	693	[11]	636.00	620.75
208	1-Hexyne	536	[4]	517.17	515.16
209	3-Hexyn-2,5-diol	553	[4]	562.42	560.40
210	Ethyne	578.15	[3]	587.57	584.98
211	1,2,4-Triethenyl-Cyclohexane	543	[4]	472.98	493.21
212	4-Fluorobenzyl chloride	863	[4]	863.00	861.47
213	1,1-Difluoro-1-chloroethane	905	[3]	905.00	880.03
214	Fluoroethene	658.15	[3]	631.81	632.80
215	Amyl nitrite	478	[4]	576.41	560.35
216	Tetrahydropyrrole	618	[4]	655.65	657.45
217	1-Octanamine	538	[4]	541.53	523.05
218	<i>N</i> -Ethyl- <i>N,N</i> -diisopropylamine	513	[4]	603.66	600.89
219	2-Amino-2-ethylhexane	538	[4]	565.61	548.13
220	<i>N</i> -Butyl-1-butanamine	533	[4]	557.25	537.62
221	<i>N,N</i> -Dimethylacetamide	627.15	[3]	672.59	682.56
222	2-Butoxime	588	[4]	588.00	588.66
223	2-Hydroxy-1-ethylaziridine	607	[4]	586.90	575.59
224	1-Benzazine	753.15	[3]	811.82	814.86
225	Nitrocarbol	652.15	[3]	661.82	671.68
226	Aminomethane	703.15	[3]	691.00	708.54
227	Ethylamine	657	[3]	669.65	678.44
228	<i>N,N</i> -Dimethylamine	673.15	[3]	685.37	701.25
229	Piperazine	728.15	[3]	678.38	683.26
230	Azabenzene	823	[4]	867.59	855.42
231	Tetrafluorethene	473.15	[3]	500.61	520.06
232	Imidole	823	[4]	836.50	838.93
233	Triacetaldehyde	510.93	[3]	569.49	573.94
234	3-Picoline	810	[3]	775.89	786.84
235	Butane nitrile	761	[4]	748.13	750.73
236	Cyanoacetic ester	733	[4]	783.41	784.63
237	2-Picoline	810.93	[3]	775.89	786.84
238	1,1-Dimethylcyclohexane	577	[3]	547.57	544.64
239	2,3,3-Trimethyl-1-Pentene	504	[4]	627.20	620.88
240	Isoprene	493.15	[3]	602.44	600.74
241	$\alpha$ -Methylstyrene	847.59	[3]	712.59	719.51
242	Butyl butyrate	623	[11]	614.64	592.88
243	Propylamine	591	[3]	648.30	648.47
244	$\alpha$ -Pinene	528.15	[3]	562.92	560.17
245	<i>trans</i> -2-Butene	597.04	[3]	606.58	607.41
246	1-Dodecanol	548.15	[3]	463.65	489.73
247	1,3-Butanediol	667.04	[3]	649.59	649.73
248	Isobutyl acrylate	613.15	[3]	642.54	638.83
249	Dimethyl terephthalate	843.15	[3]	825.19	823.65
250	1,7-Octadiene	493	[4]	531.05	517.32
251	1,3-Diisopropylbenzene	722	[4]	717.37	717.63
252	Benzoyl chloride	873	[4]	832.29	829.41
253	1,4-Dioxane	453.15	[3]	578.29	577.08
254	2-Ethylhexanal	463.15	[3]	483.91	496.49
255	Methylhexanone	728	[4]	635.11	629.39
256	2-(2-Methoxyethoxy)ethanol	488	[4]	503.68	500.18
257	2-Methoxyethyl ether	463	[4]	436.57	493.03
258	1,2-Dimethoxyethane,	473	[4]	533.99	521.67
259	Ethyl vinyl ether	451	[4]	575.22	560.78
260	2-(2-Ethoxyethoxy) ethanol	477.15	[3]	482.33	492.09
261	Pentanal	495.15	[3]	523.90	519.29
262	2,2'-Dihydroxyethyl ether	502.04	[3]	570.80	550.39

Table A1 (Continued)

	Compound name	Exp. value	Reference	MLR	Degree 3
263	1,1-Diethoxyethane	503.15	[3]	515.36	509.67
264	2-Methyl-2-propenal	507.15	[3]	539.10	540.20
265	2-Ethoxyethanol	508.15	[3]	579.75	562.35
266	1-Nonene	510	[3]	523.18	510.75
267	Methylal	510.35	[3]	555.34	541.08
268	2-Butoxyethanol	511.15	[3]	537.05	520.58
269	Butylcyclopentane	523.15	[3]	542.16	530.64
270	<i>trans</i> -Decahydronaphthalene	528	[4]	526.31	522.96
271	beta-Pinene	528	[3]	522.65	525.46
272	1-Hendecanol	550	[3]	485.00	491.58
273	1-Nonanol	550	[3]	527.70	511.71
274	Methyl acetylacetate	623	[4]	710.37	718.56
275	Propanoic acid anhydride	558	[3]	609.33	589.86
276	2-Methoxyethanol	558.15	[3]	601.10	587.73
277	2-Pentene	561	[3]	585.22	580.42
278	<i>cis</i> -2-Methylcyclohexanol	569.15	[3]	591.93	587.77
279	<i>cis</i> -4-Methylcyclohexanol	570.15	[3]	591.93	587.77
280	<i>cis</i> -1,4-Dimethylcyclohexane	577	[3]	579.53	575.08
281	<i>cis</i> -1,2-Dimethylcyclohexane	577.15	[3]	579.53	575.08
282	2-Methylnitrobenzene	693	[4]	730.33	731.32
283	2,4-Dihydroxy-2-methylpentane	698	[4]	682.23	683.91
284	<i>cis</i> -1,3-Dimethylcyclohexane	579	[3]	579.53	575.08
285	1-Butanamine	585	[3]	626.94	619.27
286	2-Furancarboxaldehyde	588.71	[3]	596.58	589.39
287	Hexahydro-1H-azepine	603.15	[3]	606.36	599.98
288	Acetylene tetrabromide	608.15	[3]	632.51	637.16
289	2,4-Pentanedione	613.15	[3]	638.67	639.32
290	Propyl acrylate	615	[3]	643.87	632.98
291	1-Methyl-2-pyrrolidinone	619.15	[3]	639.25	646.13
292	1-Methoxy-2-propyl acetate	627.15	[3]	626.71	614.97
293	1,4-Butanediol	630	[3]	625.51	613.48
294	2-(2-Ethoxyethoxy)ethyl acetate	583	[4]	505.21	497.60
295	Nitroethane	633.15	[3]	640.47	641.83
296	2-Methyl-1-butene	638	[3]	615.92	616.92
297	2-Aminoethylethanolamine	641	[3]	617.23	603.06
298	Allylamine	647.04	[3]	634.81	631.47
299	1,3-Propylene glycol	651	[3]	646.87	642.43
300	2-Ethoxyethyl acetate	652.59	[3]	602.63	580.99
301	Ethylenediamine	658.15	[3]	653.17	653.80
302	1-Methyl-3-nitrobenzene	713	[4]	730.33	731.32
303	Acetic acid anhydride	603	[4]	652.04	646.66
304	2-Furanmethanol	664.15	[3]	685.80	691.87
305	2-Heptanone	666.15	[3]	611.03	594.31
306	Crotonic acid	669.26	[3]	636.73	634.34
307	Ethyl 2-hydroxypropanoate	673.15	[3]	693.82	702.11
308	1,3,5-Trioxacyclohexane	683	[4]	575.63	577.05
309	Hexanedioic acid	678	[4]	661.03	639.64
310	Ethylene chlorohydrin	698.15	[3]	734.68	736.73
311	Cyclobutane	700	[3]	632.91	632.07
312	Tartaric acid	700.93	[3]	733.98	743.06
313	3-Isopropyltoluene	709	[3]	736.00	739.83
314	3-Methyl-2-butanol	710	[3]	661.27	672.44
315	1,2-Epoxybutane	643	[4]	631.49	630.69
316	<i>N,N</i> -Dimethylformamide	683	[4]	564.10	561.31
317	<i>N</i> -Phenylacetamide	725.15	[3]	759.96	760.32
318	Methyl formate	729.26	[3]	744.09	762.48
319	1,1-Dichloroethane	731.15	[3]	761.14	770.60
320	<i>trans</i> -1,2-dichloroethylene	733	[3]	727.64	725.05
321	1,1,2-Trichloroethane	733.15	[3]	818.65	812.62
322	Acetonitrile	797	[4]	790.83	804.18
323	Vinyl chloride	745	[3]	711.83	715.42
324	Ethylphenylamine	752.15	[3]	739.33	744.14
325	Acrylonitrile	754.26	[3]	756.00	762.67
326	3-Aminotoluene	755.15	[3]	759.51	766.35
327	4-Aminotoluene	755.15	[3]	759.51	766.35
328	Nitrobenzol	753	[4]	737.14	744.69
329	1-Bromopropane	763.15	[3]	657.66	657.49
330	3-Hydroxypropionitrile	767.59	[3]	760.53	764.49
331	1,4-Benzenedicarboxylic acid	769	[3]	814.95	810.35
332	Ethyl bromide	784.26	[3]	679.01	687.56
333	<i>o</i> -Nitroaniline	794.15	[3]	735.20	736.57
334	<i>N</i> -Phenylacetamide	803.15	[3]	771.00	779.95
335	4-Picoline	810	[3]	775.89	786.84
336	Hexanedinitrile	823.15	[3]	810.13	794.59
337	Dichlorofluoromethane	825.15	[3]	741.66	748.85

Table A1 (Continued)

	Compound name	Exp. value	Reference	MLR	Degree 3
338	1,2-Dichloropropane	830	[3]	758.11	761.07
339	Benzoic acid	805	[4]	791.60	798.61
340	1,3,5-Trichlorobenzene,	850	[3]	897.25	845.55
341	<i>p</i> -Nitroaniline	773	[4]	735.20	736.57
342	Phthalic anhydride	857.04	[3]	866.08	862.51
343	Hexachlorobutadiene	883.15	[3]	834.14	816.57
344	Methyl chloride	905	[3]	890.00	859.46
345	1,4-Dichlorobenzene	920	[3]	854.25	880.02
346	1,2-Dichlorobenzene	913	[4]	854.25	880.02
347	2-Methylnaphthalene	802	[3]	793.86	796.18
348	<i>cis</i> -1,2-Dichloroethylene	733	[3]	727.64	725.05
349	Ethylene	723.15	[3]	696.01	705.72
350	Ethyl chloride	792	[3]	743.63	751.52
351	Ethane	745	[3]	686.13	703.16
352	Acetone	738	[4]	696.44	711.67
353	Propane	723	[3]	664.77	673.03
354	Chloroprene	593.15	[3]	662.97	658.21
355	<i>cis</i> -2-Butene	598.15	[3]	606.58	607.41
356	Diethylamine	585.15	[3]	642.66	641.28
357	Cyclopentadiene	913.15	[3]	789.12	797.29
358	2-Methyl-2-butene	563	[3]	592.56	598.40
359	<i>p</i> -Hydroquinone	788.7	[3]	886.08	826.80
360	2-Methyl-1,3-butadiene	493	[3]	557.72	558.01
361	2-Hexanone	697.04	[3]	632.38	622.27
362	<i>m</i> -Cresol	832.04	[3]	820.36	872.41
363	<i>o</i> -Cresol	872.04	[3]	820.36	872.41
364	2,4-Dimethylpentane	610	[3]	627.51	628.30
365	Isopropyl butyrate	708	[3]	660.07	657.35
366	3-Methylhexane	553.15	[3]	603.44	593.30
367	2,6-Xylenol	872.04	[3]	813.55	867.67
368	Vinylcyclohexene	543	[3]	547.51	543.44
369	3,4,4-Trimethyl-2-pentene	598	[3]	603.84	602.22
370	Isobutyl isobutyrate	705.15	[3]	662.79	664.73
371	Benzyl acetate	734	[3]	775.37	780.24
372	Glyceryl triacetate	706	[3]	730.63	731.12
373	Dicyclopentadiene	783.15	[3]	748.43	744.65
374	Diethyl phthalate	730.15	[3]	782.49	774.10
375	Phenyl benzoate	833	[3]	872.03	855.43
376	2-(2-Butoxyethoxy)ethanol	477.59	[3]	439.62	493.46
377	Diglycolic acid	503	[3]	649.02	626.52
378	1-Hendecene	510	[3]	480.47	491.28
379	<i>cis</i> -Decahydronaphthalene	523.15	[3]	526.31	522.96
380	Tetrahydro-2-furancarbinol	555.37	[3]	594.60	588.67
381	<i>trans</i> -2-Methylcyclohexanol	569.15	[3]	591.93	587.77
382	Morpholine	583.15	[3]	628.34	628.07
383	3,3-Dimethylpentane	610	[3]	654.71	647.27
384	4-Methyl-3-penten-2-one	617.59	[3]	602.87	606.24
385	Vinyl ether	633.15	[3]	561.74	547.47
386	Ethanamine	657	[3]	669.65	678.44
387	4-Hydroxynitrobenzene	729	[4]	730.33	731.32
388	Isocrotonic acid	669	[3]	636.73	634.34
389	Isopropylamine	675.15	[3]	672.37	685.86
390	2-Methoxy-2-methylpropane	708	[4]	642.69	634.03
391	2-Propenoic acid	688	[4]	686.57	692.73
392	2-Aminotoluene	755.15	[3]	759.51	766.35
393	Ethylene diacetate	755.15	[3]	713.98	716.99
394	Allyl chloride	663	[4]	708.80	705.13
395	1,3-Benzenedicarboxylic acid	769	[3]	814.95	810.35
396	Methyl bromide	810.37	[3]	825.37	818.85
397	1,2-Dimethyl phthalate	829	[3]	825.19	823.65
398	1,2,4-Trichlorobenzol	844.26	[3]	897.25	845.55
399	<i>m</i> -Nitroaniline	794	[4]	735.20	736.57
400	Dichloromethane	878	[4]	804.16	804.54

Table A2

Experimental values and predicted values of the compounds in the validation set

No.	Compound name	Exp. value	Reference	MLR	Degree 3
1	1-Hexene	538	[4]	587.24	571.80
2	1-Hexadecene	513.15	[3]	373.71	554.21
3	2,3-Dimethyl-2-butene	673.7	[3]	578.54	589.59
4	Cyclohexane	533.15	[3]	583.62	577.15
5	Methylcyclohexane	558.15	[3]	581.57	576.11
6	Isopropylbenzene	697.04	[3]	742.81	753.03
7	1,3,5-Trimethylbenzene	823.15	[3]	747.83	748.34

Table A2 (Continued)

No.	Compound name	Exp. value	Reference	MLR	Degree 3
8	1-Pentanol	573.15	[3]	613.12	599.87
9	1-Octanol	555	[3]	549.06	528.73
10	Phenol	878	[11]	827.16	876.07
11	Diisopropylamine	588.71	[3]	648.11	655.90
12	1,2-Dichloroethane	711	[11]	779.78	769.54
13	Isobutyl formate	593.15	[3]	704.10	711.48
14	2,3-Dimethylbutane	669	[11]	648.87	657.81
15	1,3-Cyclohexadiene	633	[11]	764.47	771.11
16	2-Methylbiphenyl	775	[11]	829.95	823.38
17	Dipropyl ether	488	[4]	546.00	529.93
18	Dihexyl ether	458.15	[3]	417.88	504.76
19	Octanoic acid	570	[3]	588.17	560.00
20	Hexylacetylene	498	[4]	474.46	492.79
21	Isopentyl nitrite	481	[4]	600.49	592.09
22	Cyclopropane	770.93	[3]	657.55	661.53
23	1-Propyne-3-ol	388.15	[3]	572.27	566.27
24	Dibutyl sebacate	638.15	[3]	457.75	494.43
25	Ethoxy ethane	433.15	[3]	588.71	575.05
26	1-Dodecene	528.15	[3]	555.54	516.29
27	1-Chloropentane	533.15	[3]	679.57	662.31
28	Ethyl acetylacetate	568.15	[3]	689.01	688.55
29	<i>trans</i> -4-Methylcyclohexanol	570.15	[3]	591.93	587.77
30	Ethyleneimine	593.15	[3]	704.94	717.22
31	Triethylene glycol	644	[3]	473.37	490.09
32	2-Isopropyltoluene	650	[3]	736.00	739.83
33	2-Methyl-1-propanamine	651.15	[3]	651.02	655.80
34	1-Nitropropane	694.15	[3]	619.12	612.90
35	3,5,5-Trimethyl-2-cyclohexane-1-one	733.15	[3]	617.64	626.44
36	Acetaldehyde	758.15	[3]	587.95	585.74
37	Phenylacetylene	763	[3]	656.54	647.94
38	1-Chloropropane	793.15	[3]	722.28	722.37
39	Vinylidene chloride	843	[3]	758.34	763.58
40	<i>cis</i> -1-Propenylbenzene	848	[3]	681.89	679.35
41	2,3-Dimethylphenol	872	[3]	813.55	867.67
42	3,4-Dimethylphenol	872.04	[3]	813.55	867.67
43	Formic acid	874.26	[3]	738.97	754.30
44	Aniline	813	[2]	766.32	778.86
45	Propylene	728.15	[3]	651.29	655.71
46	Maleic anhydride	749.82	[3]	838.71	853.23
47	1,3-Butadiene	702.04	[3]	616.46	609.80
48	1-Butene	657.04	[3]	629.94	626.27
49	1,5-Pentanediol	608.15	[3]	604.16	586.06
50	<i>cis</i> -2-Hexene	526	[3]	563.87	555.74
51	Ethylcyclobutane	483	[3]	609.51	602.46
52	<i>p</i> -Cresol	832.04	[3]	820.36	872.41
53	2-Methylhexane	566	[4]	603.44	593.30
54	<i>n</i> -Butyl acrylate	565.93	[3]	622.51	604.45
55	Styrene	763.15	[3]	726.61	729.15
56	Isopentyl propionate	698	[3]	638.72	627.85
57	<i>p</i> -Cymene	709.26	[3]	736.00	739.83
58	2-Ethylhexyl acrylate	530.93	[3]	561.18	536.87
59	1,1-Diphenylethane	713.15	[3]	818.13	816.76
60	Dibutyl phthalate	675.15	[3]	697.07	656.62
61	Ethyl methyl ether	463.15	[3]	610.06	601.61
62	Peroxyacetic acid	473.15	[3]	654.12	652.72
63	Cyclohexenylethylene	543	[3]	637.17	624.75
64	Butanoic acid anhydride	552.59	[3]	566.63	541.44
65	Cyclohexanamine	566.15	[3]	586.45	580.73
66	Tetraethylenepentamine	594	[4]	522.77	507.17
67	Nonanoic acid	589	[3]	566.82	537.71
68	Tetrahydrofuran	594.26	[3]	605.60	603.70
69	Diethylenetriamine	631	[3]	609.70	594.61
70	<i>N,N</i> -Dimethylbenzenamine	644.26	[3]	737.59	746.91
71	2-Butanamine	651	[3]	651.02	655.80
72	1,2,3,4-Tetramethylbenzene	700	[3]	741.02	735.04
73	2-Nitropropane	698	[4]	643.19	649.12
74	Diisopropyl ether	678	[4]	594.15	587.86
75	2-Hydroxybenzoic acid methyl ester	728	[3]	855.64	879.58
76	4-Methyl-2-pentanone	721	[4]	656.46	658.93
77	Propionitrile	785	[3]	769.48	778.49
78	2-Hydroxybenzoic acid	818.15	[3]	850.51	878.66
79	<i>trans</i> -1-Methylstyrene	848	[3]	681.89	679.35
80	2-Chloropropane	863	[4]	700.61	713.28
81	2,5-Dimethylphenol	872	[3]	813.55	867.67
82	4-Hydroxy-4-methyl-2-pentanone	876.48	[3]	698.77	700.47
83	1,3-Dichlorobenzene	920	[3]	854.25	880.02



## References

- [1] API, API Publication 581: Risk-Based Inspection Base Resource Document, American Petroleum Institute, 2000.
- [2] L.M. Egolf, P.C. Jurs, Estimation of Autoignition Temperatures of Hydrocarbons, Alcohols, and Ester from Molecular Structure, *Ind. Eng. Chem. Res.* 31 (1992) 1798–1807.
- [3] AIChE, DIPPRO<sup>®</sup>, DIPPR Project 801 Pure Component Data, 1996 (public version).
- [4] The Hazardous Chemical Database, The World Wide Web: <http://ull.chemistry.uakron.edu/erd/index.html>.
- [5] R.J. Lewis, Hazardous Chemicals Desk Reference, fifth ed., Wiley-Interscience, 2002.
- [6] R.J. Lewis, SAX'S Dangerous Properties of Industrial Materials, 11th ed., John Wiley & Sons, 2004.
- [7] IPCS INCHEM., The World Wide Web: <http://www.inchem.org/pages/icsc.html>.
- [8] ASTM International, ASTM Standard TEST Method E659-78, The American Society for Testing and Materials, West Conshohocken, PA, 2005.
- [9] J. Tetteh, E. Metcalfe, S.L. Howells, Optimisation of radial basis and back-propagation neural networks for modeling auto-ignition temperature by quantitative-structure property relation, *Chemom. Intell. Lab. Syst.* 32 (1996) 177–191.
- [10] B.E. Mitchel, P.C. Jurs, Prediction of autoignition temperatures of organic compounds from molecular structure, *J. Chem. Inf. Comput. Sci.* 37 (1997) 538–547.
- [11] T. Susuki, Quantitative structure–property relationships for auto-ignition temperatures of organic compounds, *Fire Mater.* 18 (1994) 81–88.
- [12] T. Susuki, K. Ohtaguchi, K. Koide, Correlation and prediction of autoignition temperatures of hydrocarbons using molecular properties, *J. Chem. Eng. Jpn.* 25 (1992) 606–608.
- [13] Y.S. Kim, S.K. Lee, J.H. Kim, J.S. Kim, K.T. No, Predictions of autoignitions (AITS) for hydrocarbons and compounds containing heteroatoms by the quantitative structure–property relationship, *R. Soc. Chem.* 2 (2002) 2087–2092.
- [14] A.S.T.M. International, ASTM Computer Program for Chemical Thermodynamic and Energy Release Evaluation—CHETAH v 8.0, The American Society for Testing and Materials, West Conshohocken, PA, 2006.
- [15] T.A. Albahri, Flammability characteristics of pure hydrocarbons, *Chem. Eng. Sci.* 58 (2003) 3629–3641.
- [16] T.A. Albahri, R.S. George, Artificial neural network investigation of the structural group contribution method for predicting pure components auto ignition temperature, *Ind. Eng. Chem. Res.* 42 (2003) 5708–5714.
- [17] D.E. Swarts, M. Orchin, Spontaneous ignition temperature of hydrocarbons, *Ind. Eng. Chem.* 49 (3) (1957) 432–436.
- [18] W.A. Affens, J.E. Johnson, H.W. Carhart, Effect of chemical structure on spontaneous ignition of hydrocarbons, *J. Chem. Eng. Data* 6 (4) (1961) 613–619.
- [19] G.A.F. Seber, C.J. Wild, *Nonlinear Regression*, Wiley, 1988, pp. 191–194.
- [20] H. Demuth, M. Beale, M. Hagan, *Neural Network Toolbox—User's Guide (Version 5)*, The MathWorks, Inc., 2006.
- [21] A.B. Umetrics, *Multi- and Megavariate Data Analysis*, Umetrics AB, 2001, p. 107.