

Estimation of the normal boiling point of organic compounds via a new group contribution method



Habib Ghasemitarbar, Kamyar Movagharnejad*

Babol University of Technology, Faculty of Chemical Engineering, Babol, Iran

ARTICLE INFO

Article history:

Received 11 February 2015

Received in revised form

18 November 2015

Accepted 24 November 2015

Available online 2 December 2015

Keywords:

Normal boiling point

Group contribution

Organic compound

Structural isomerism

Stereoisomerism

ABSTRACT

An improved second order group contribution method is developed to determine the normal boiling point of organic compounds. This method is based on the Joback's first order functional groups with some changes and some newly added functional groups developed from experimental data. Each functional group has a group contribution value, and the contribution values have been optimized using 2036 experimental data of organic components ranging from $N_C = 1-36$ and $M_W = 26-555$ ($\text{g}\cdot\text{mol}^{-1}$), including heavy and complex polycyclic aromatic compounds. In this method, it is tried to distinguish most of the structural isomers and stereoisomers by second order functional groups to predict their different normal boiling points. First and second order functional groups of the hydrocarbons and hydrocarbon derivatives containing carbon, hydrogen, oxygen, nitrogen, sulfur, fluorine, chlorine, bromine and iodine atoms, are included in this study. The *fminsearch* mathematical approach from MATLAB software is used in this study to select an optimal collection of functional groups (122 functional groups) and subsequently to develop the model. The results of the new method are compared to several well-known methods. The average absolute deviation of normal boiling point predictions for 2036 organic compounds is found to be 4.35194 K; while the percentage of the average absolute relative deviation is just 1.01075%.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

To carry out design calculations, every engineer needs to know thermo-physical properties such as normal boiling point, freezing point, heats of formation, densities and so on. Many of these properties have been measured and tabulated for pure substances, but thermo-physical properties under extreme conditions are not known usually, because the experiments in this regime are often too difficult and expensive. When the physical properties are not accessible, the estimation methods have to be employed. Although, these methods are not very adequate, sometimes there are no other choices.

The traditional approach is to develop mathematical models to predict the properties. Examples of these models include the large number of activity coefficient models like the UNIFAC model [1] or the equations of state which help us to predict the vapor–liquid equilibrium behavior. As we know, these models do not always work, and typically require some experimental data as the input. To

make models more predictive, group contribution methods have been developed [2]. A group contribution method is used to predict the properties of pure components and mixtures by using group or atom properties. Instead of knowing the properties of hundreds of compounds, only the data for a few dozens of groups have to be known. The vast majority of organic components, for example, are built of carbon, hydrogen, oxygen, nitrogen, halogens, and maybe sulfur or phosphorus, together with a single, a double, and a triple bond, there are only few atom types and three bond types to build thousands of components. The next slightly more complex building blocks of components are functional groups which are themselves built of few atoms and bonds. All organic molecules are formed by attaching other atoms or groups of atoms to the carbon backbone of a hydrocarbon. The physical and chemical properties of these derivative molecules are determined by the specific arrangement of atoms called functional groups [3,4].

The physical properties of compounds are described by a large group of structure related characteristics, such as normal boiling point and critical parameters. Most of these properties have been targeted by different correlations and approaches [5–8]. The boiling point is a fundamental thermo-physical property describing

* Corresponding author.

E-mail address: movagharnejad@yahoo.com (K. Movagharnejad).

the temperature of transition between the saturated vapor and liquid phases. It is clear that if specify the system pressure, the temperature at which pure species a boils is fixed. The normal boiling point of a substance, the temperature at which a pure species boils at a pressure of 1 atm, is well defined. The presence of an essentially nonvolatile solute in a liquid phase of mostly solvent, which the system is at a pressure of 1 atm. It was observed that it always requires a greater temperature to boil in mixture than in pure species. This phenomenon is termed boiling-point elevation. Pure component normal boiling point (NBP) of liquid compounds is of great importance for many applications in the design and simulation of various chemical and biochemical processes as well as for environmental and safety problems [9,10]. In some cases, especially for heavy hydrocarbons in the thermal cracking may occur at high temperatures, so the boiling points at lower pressures are specified. Boiling points of heavy hydrocarbons are usually measured at 1, 10, or 50 mm Hg. The normal boiling point, when available, is one of the most important characterization parameters for hydrocarbons and is frequently used in property estimation methods [9].

It is not always possible to find experimental values for the NBP since measurement is expensive and time consuming or sometimes even difficult or impossible, therefore accurate predictive methods can replace the measurements. The estimated properties are not as precise as the well-made measurements but for many purposes the quality of these estimated properties is sufficient. Predictive methods can also be used to check the results of the experimental work. Almost other thermo-physical properties are predictable from the normal boiling point [11–15], so the precise forecast of boiling point could be quite important. According to the well-known rule of Guldberg [16], the NBP of a liquid can be roughly estimated as the two-thirds of the critical temperature on the absolute scale. The rule is not obeyed by liquids of low boiling point, and even for liquids boiling in the range 293–523 K [17]. So other more accurate predictive methods have been proposed by researchers. Group contribution methods include a very important part of these predictive methods which are briefly described in the next section.

2. Group contribution methods for the prediction of normal boiling point

The following general considerations have been proposed to evaluate the prediction policy of the thermo-physical properties [18]:

- Generalized methods based on group contributions or corresponding states theories are usually used when no experimental data are available.
- If the substance of interest is listed in the input data file of any methods, that method usually gives better results for that particular substance.

The generalized methods based on the group contributions are usually divided into four different categories listed as follows:

2.1. Additive group contribution method

The simplest form of a group contribution method is to determine the component property by summing up the group contributions. In this simple form, it is assumed that the property is linearly dependent on the number of groups. it is also assumed that other groups in the molecule do not alter a groups contribution to the physical property. This simple approach is employed in the Joback and Reid method [19] for normal boiling point and it works

well in a limited range of components and property ranges but may lead to quite large errors outside of the applicable ranges. Joback developed the Lydersen's group contribution scheme [20], added several new functional groups, and calculated some new contribution values to improve the original method. His relation for NBP is given by the following equation:

$$T_b = 198.2 + \sum_k N_k(tbk) \quad (1)$$

where T_b is normal boiling point, K; N_k is the number of groups of type k in molecule, with contribution of tbk . Joback assumed a constant contribution of the added groups in homologous series like the n -alkanes, which does not describe the real behavior of the normal boiling points correctly [21]. Instead of the constant contribution a decrease of the contribution with increasing number of groups may give more accurate results. The Joback method usually gives acceptable estimation for mid-sized components, but leads to high deviations for large and small molecules.

2.2. Additive group contributions and correlations

In this group of methods the pure additive group contributions are correlated to give the desired property. This following thermodynamic relation may be used.

$$dH = TdS + VdP \quad (2)$$

It is known that during any phase change, temperature and pressure of the system remain constant, so:

$$T_b = \frac{\Delta H_{vap}(T_b)}{\Delta S_{vap}(T_b)} \quad (3)$$

several simple relations are available for the estimation of enthalpies of vaporization at the NBP of the liquid. Trouton's rule (Eq. (4)), for example, is only suitable for non-polar liquids, but the Giacalone equation (Eq. (5)) is fairly reliable for both polar and non-polar liquids:

$$\Delta H_{vap}(T_b) = 88 T_b \quad (4)$$

$$\Delta H_{vap}(T_b) = \left(\frac{RT_c T_b}{T_c - T_b} \right) \ln P_c \quad (5)$$

As the total interaction between the molecules in the liquid phase is much larger in comparison to the vapor phase, the enthalpy of vaporization can be approximated by the total intermolecular interaction in the liquid phase.

For example in a n -alkane homologous series, a given member differs from the preceding or succeeding member by the constant difference of $-CH_2-$ group. Hence, the volume of the molecule increases linearly with the number of $-CH_2-$ groups, but molecules are occasionally tend to tangle to a more or less spherical form, and the outer surface may increases non-linearly. It is assumed that the molecular interaction in the organic liquids should also be proportional to the molecular surface. Fig. 1 clearly shows that the ΔH_{vap} (Btu/lbmol) at the NBP for the n -alkanes is a function of the molecular weight and can be fitted with a single correlation equation (Eq. (6)) containing molecular weight.

$$\Delta H_{vap}(T_b) = 3048(M_w^{0.4106}) - 6263 \quad (6)$$

Where M_w is the molecular weight and $\Delta H_{vap}(T_b)$ is the enthalpy of vaporization at the NBP. As shown in Fig. 1, molecular weight can be used in properties estimation methods for additional property,

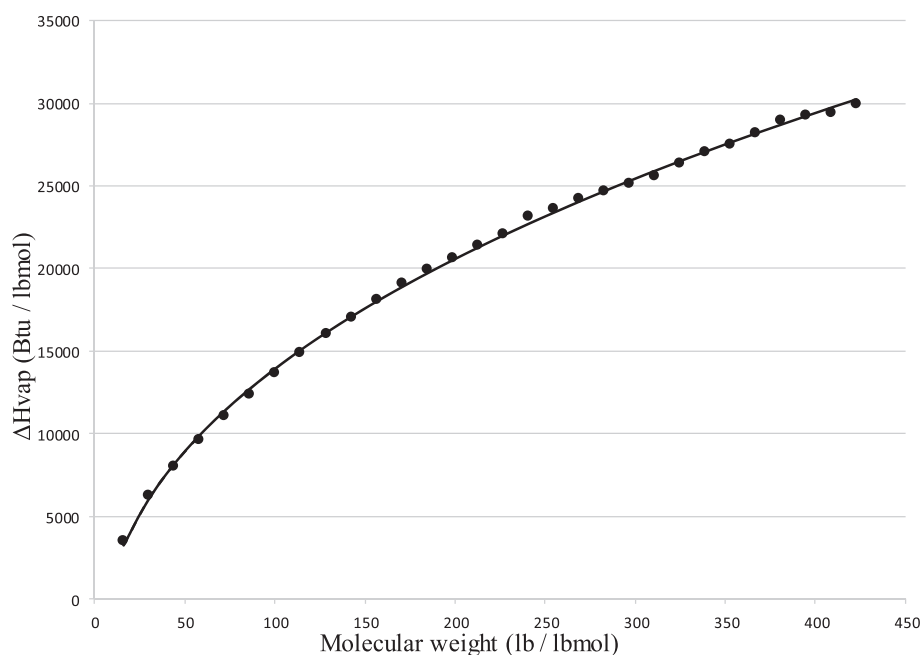


Fig. 1. Enthalpy of vaporization of the n-alkanes at the NBP as function of molecular weight. Solid lines (–): predicted curves with the Eq. (6). (●), experimental points of the n-alkanes [28–30].

which helping to improve accuracy of estimation methods.

Marrero-Morejon and Pardillo-Fontdevila [22] described a method for NBP prediction that was called a group interaction contribution technique or effectively a bond contribution method. Their basic equation of normal boiling point was given by the following equation:

$$T_b = M_W^{-0.404} \sum_k N_k(tbbk) + 156.00 \quad (7)$$

where M_W is the molecular weight and N_k is the number of atoms of type k in molecule, with contributions $tbbk$. In this method, the molecular weight is applied as an additional property. This approach often gives better results than pure additive equations because the relation with a known property introduces some knowledge about the molecule. Commonly used additional properties are the molecular weight, the number of atoms, chain length, and ring sizes and counts.

2.3. Group interactions

For the prediction of mixture properties it is not usually sufficient to use a purely additive method. Instead the property may be determined from the group interaction parameters shown by the following equation.

$$F = f(G_{ij}) \quad (8)$$

Where F stands for the property and G_{ij} stands for the group interaction values. A typical example of the group contribution method using group interaction values would be the UNIFAC method for activity coefficients estimation. A big disadvantage of the group interaction model is the need for many more model parameters. For example a simple additive model only needs ten parameters for ten groups, while a group interaction model needs 45 parameters.

2.4. Group contributions of higher orders

Some newer methods introduce the second-order groups which can be defined as the super-groups containing several first-order (standard) groups. The application of the second-order groups allows the introduction of new parameters for the position of groups. Another possibility is to modify the first-order group contributions if specific other groups are also present [23]. Constantinou and Gani (CG) [24] developed a second order group contribution method based on the UNIFAC groups. Their equation for normal boiling point is given by the following equation:

$$T_b = 204.359 \ln \left[\sum_k N_k(tb1k) + W \sum_j M_j(tb2j) \right] \quad (9)$$

where N_k is the number of first-order groups of type k in the molecule; $tb1k$ is the contribution for the first-order group of k ; M_j is the number of second-order groups of type j in the molecule; $tb2j$ is the contribution for the second-order group of j . The value of W is set to zero for first-order calculations and set to unity for second-order calculations.

The advantages of group contribution methods are so clear. However, most of group contribution methods have serious problems that they cannot distinguish isomers because the isomers have the same number and kind of groups so that the calculated results are inevitably the same. This work proposes a new group contribution method for the estimation of the NBP of organic compounds based on the Joback functional groups with some changes. Several functional groups were added using the experimental data for 2036 organic components, including large and complex polycyclic compounds. The new method can also be applied for isomer species and other specific hydrocarbons by using second order groups (see Table 2). Estimation of the NBP would be more reasonable for isomer species and other specific hydrocarbons in the new method than previous methods. In the following, we will show the results of these species for different methods.

Table 1
First-order group contributions for normal boiling temperature.

Group i	tbi [Kelvin]	Group i	tbi [Kelvin]
1 CH ₃ (1)	13.0945	27 COO (2) (Ester)	97.4615
2 CH ₂ (2)	22.9831	28 (C=O)—O—(C=O) (2) (Anhydride)	174.5641
3 CH (3)	23.7995	29 (C=O)—O—(C=O) (2) (ss) (Cyclic Anhydride)	229.7723
4 C (4)	19.1978	30 =O (1)	−30.0125
5 CH ₂ (ss) (2)	23.3777	31 NH ₂ (1)	51.8523
6 CH (ss) (3)	20.1251	32 NH (2)	41.4335
7 C (ss) (4)	17.7436	33 NH (ss) (2)	44.9923
8 =C (ss) (3)	22.1725	34 N (ss) (3)	38.8612
9 =CH ₂ (1)	7.1245	35 N (3)	31.8994
10 =CH (2)	29.2596	36 =N (2)	26.7784
11 =C (3)	36.1612	37 =N— (ds) (2)	22.0363
12 =C= (2)	54.6957	38 C≡N (1)	86.4988
13 ≡CH (1)	20.1071	39 NO ₂ (1)	61.9870
14 ≡C (2)	41.4620	40 SH (1)	88.6627
15 =CH (ds) (2)	25.8530	41 S (2)	91.9444
16 =C (ds) (3)	34.2365	42 S (ss) (2)	95.3134
17 OH (1) (Alcohol)	75.7146	43 =S (1)	88.7389
18 OH (1) (Cyclic Alcohol)	74.3934	44 S=O (2)	133.7269
19 AOH (1) (Phenol)	72.4050	45 S=O (ss) (2)	293.1187
20 O (2) (Ether)	30.1699	46 O=S=O (2)	220.2567
21 O (ss) (2) (Ether)	38.9549	47 O=S=O (ss) (2)	304.6647
22 C=O (2) (Ketone)	89.2602	48 F (1)	−0.1286
23 C=O (ss) (2) (Ketone)	105.1172	49 Cl (1)	37.9323
24 CH=O (1) (Aldehyde)	88.9145	50 Br (1)	70.3924
25 COOH (1) (Acid)	174.1229	51 I (1)	106.2119
26 HCOO (1) (Formate)	95.6576		

All of the primary properties may be described as a function of the molecular structural information described in terms of first-order, second-order and third-order groups. It can be noted that simpler molecules may be completely described by first-order groups but more complex molecules may contain second and even third-order groups. Various estimation methods for the prediction of normal boiling points from molecular structure are available. Good descriptions of these methods may be found elsewhere [21,25–27]. In this work, we will compare our results with several popular estimation methods such as Joback and Reid [19], Marrero and Pardillo [22], Constantinou and Gani [24] and Marrero and Gani [36]. All of these methods only require the knowledge about the molecular structure and therefore can be placed in a common group.

3. Development of the new group contribution model

The organic compounds data bank which was established for this work contains information related to the name, chemical formula, compounds class, molecular weight, experimental NBP and number of first- and second-order groups of organic compounds. The organic compounds include carbon, hydrogen, oxygen, nitrogen, sulfur, fluorine, chlorine, bromine and iodine atoms.

The database contains 2036 entries, the compounds include linear and branched alkanes (175), cycloalkanes (107), alkenes (166), cycloalkenes (42), alkynes (40), benzene derivatives (160), aromatics (58), alcohols (64), phenols (32), ethers (104), ketones (57), aldehydes (24), esters (141), formates (14), acids (25), other C.H.O compounds (23), amines (102), other nitrogen compounds (109), sulfur compounds (80), halogenated compounds (301), and other compounds (212). The numbers in the parentheses specify the number of components of each class of compounds. The experimental NBPs were used for modeling, are were obtained from the Dortmund Data Bank (DDB) and various standard texts [28–30].

Group contributions are obtained from the known experimental data of well-defined pure components. The given pure component properties are then assigned to the groups by statistical correlations

like e.g. (multi-)linear regression. Analysis of the different estimation methods shows which important steps during the development of a new method are:

1. Evaluation of the quality of available experimental data, elimination of wrong data, finding of the outliers.
2. Construction of the functional groups.
3. Searching additional simple and easily accessible properties that can be used to correlate the sum of group contributions with the examined property.
4. Finding a good but simple mathematical equation to relate the group contribution sum with the desired property.
5. Fitting the final group contribution weighting parameters.

Just a single-parameter contribution was considered for each group. The NBP function was constructed by all group contributions. Joback functional groups [30] and a few of other groups were applied into this work. For a molecule with N_i and N_j first and second-order groups, respectively, the NBP is given by the Eqs. (10)–(14) with values of the individual parameters given in Tables 1–4. The normal boiling temperature is given in Kelvins:

$$T_1(K) = \sum_i N_i(tbi) + (Q_1) \sum_j N_j(tbj) \quad (10)$$

$$T_2(K) = Q_2(aM_W^3 + bM_W^2 + cM_W) + d \quad (11)$$

$$Q_1 = N_C + \sum_{i=1}^{N_A} X_i U_i \quad (12)$$

$$Q_2 = N_C + \sum_{i=1}^{N_A} X_i V_i \quad (13)$$

Here, normal boiling point function is expressed as follows:

Table 2

Second-order group contributions for normal boiling temperature.

Group j	Description	tbj [Kelvin]	Example	Number of groups
1 (CH ₃) ₂ CH–	Two methyl groups connected to the CH group in a chain.	0.00458	O-Diisopropylbenzene	2
2 CH ₃ (CH)<	One methyl group connected to the CH in a chain.	–0.04893	2,4,6-Trimethylnonane	2
3 CH ₃ –CH ₂ (CH–CH ₃)–	One methyl group connected to the CH into CH ₃ –CH ₂ –CH in a chain.	0.12454	2,2-Oxybisbutane	2
4 (CH ₃) ₃ C–	Three methyl groups connected to the C in a chain.	0.46135	2,2,4,6,6-Pentamethylheptane	2
5 (CH ₃) ₂ C<	Two methyl groups connected to the C in a chain.	0.01029	4,4-Dimethylheptane	1
6 CH ₃ –CH ₂ [C(CH ₃) ₂]–	Two methyl groups connected to the C into CH ₃ –CH ₂ –C in a chain.	1.17176	2,2-Dimethyl-1-butanol	1
7 CH ₃ –CH ₂ (C–CH ₃)<	One methyl group connected to the C into CH ₃ –CH ₂ –C in a chain.	–0.28861	3-Methyl-3-phenylpentane	1
8 CH ₃ (C–)<	One methyl group connected to the C in a chain.	0.17777	1,1,2,2-Tetrachloropropane	1
9 CH ₂ =CH–	One methylene group connected to the CH in a chain by a double bond.	1.06565	1,3-Butadiene	2
10 CH ₂ =(C–CH ₃)–	One methyl group and one methylene group connected to the C in a chain by a double bond.	1.08778	2,5-Dimethyl-1,5-hexadiene	2
11 CH ₂ =C<	One methylene group connected to the C in a chain by a double bond.	0.50576	2,3-Dichloro-1,3-butadiene	2
12 CH ₃ –CH=	One methyl group connected to the CH=in a chain.	0.16043	Cis-2-trans-4-hexadiene	2
13 CH ₃ –CH ₂ –CH=	One CH ₃ –CH ₂ group connected to the CH=in a chain.	–0.00941	Trans-3-hexene	2
14 CH ₃ –CH ₂ (C–)=	One CH ₃ –CH ₂ group connected to the C=in a chain.	0.52765	2-Ethyl-1-pentene	1
15 CH ₃ –CH ₂ (C–CH ₃)=	One methyl group connected to the C=into CH ₃ –CH ₂ –C=in a chain.	0.09676	2,3-Dimethyl-2-pentene	1
16 CH ₃ (C–)=	One methyl group connected to the C=in a chain.	0.51997	2,5-Hexanedione	2
17 (CH ₃) ₂ C=	Two methyl groups connected to the C=in a chain.	0.56329	2,5-Dimethyl-2,4-hexadiene	2
18 CH≡C–	One methylidyne group connected to the C in a chain by a triple bond.	0.56935	5-Methyl-1-hexyne	1
19 CH ₃ –C≡C–	One methyl group connected to the C≡C in a chain.	0.53211	2-Hexyne	1
20 CH ₃ –CH ₂ –C≡C–	One CH ₃ –CH ₂ group connected to the C≡C in a chain.	0.11879	3-Heptyne	1
21 CH ₃ –O–	One methyl group connected to the oxygen atom in a chain.	1.04785	Dimethylmalonate	2
22 CH ₃ –CH ₂ –O–	One CH ₃ –CH ₂ group connected to the oxygen atom in a chain.	0.41932	1,1-Diethoxyethane	2
23 (CH ₃) ₂ N–	Two methyl groups connected to the nitrogen atom in a chain.	0.04168	Tetramethylhydrazine	2
24 CHa– a e (0–3)	One carbon atom connected to the main chain or ring, exclusive of the above foregoing functional groups (1–23).	0.06768	1,3-Dimethylnaphthalene	2
25 CH _a CH _b – a,b e (0–3)	Two carbon atoms connected to the main chain or ring, exclusive of the above foregoing functional groups (1–23).	–0.10259	3-Ethyl-3-phenylpentane	2
26 CH _a CH _b CH _c – a,b,c e (0–3)	Three carbon atoms connected to the main chain or ring, exclusive of the above foregoing functional groups (1–23).	–0.13493	Dipropyl succinate	2
27 CH _a CH _b CH _c CH _d – a,b,c,d e (0–3)	Four carbon atoms connected to the main chain or ring, exclusive of the above foregoing functional groups (1–23).	–0.16428	3-Hexylcyclopentene	1
28 CH _a CH _b CH _c ... CH _d – a,b,c,d ... e (0–3)	The other groups higher than four carbon atoms, connected to the main chain or ring, exclusive of the above foregoing functional groups (1–23).	–0.36236	3-Octylcyclopentene	1
29 Ncarbon	Number of carbons into the other groups higher than four carbon atoms.	0.03839	3-Octylcyclopentene	8
30 Alkene bond	The alkene chain connected to the ring.	2.06190	3-Phenyl-2-propenol	1
31 Ring	Number of ring.	–0.05102	4,5,9,10-Tetrahydropyrene	4
32 C–C	Number of the common single bonds between connected rings.	0.06173	4,5,9,10-Tetrahydropyrene	3
33 C=C	Number of the common double bonds between connected rings.	0.00001	4,5,9,10-Tetrahydropyrene	2
34 C–C (out)	Number of the common single bonds between many rings.	0.03212	Triphenylmethane	3
35 Double bond	Number of double bond.	–1.24364	Methyl isopropenyl ketone	2
36 Triple bond	Number of triple bond.	–2.56347	3-Nonyne	1
37 Cis-	Number of cis- position.	0.04210	Cis-1,3,5-trimethylcyclohexan	2
38 Trans-	Number of trans- position.	0.29954	trans,trans-2,4-Hexadiene	2
39 Alpha	Number of atoms into the alpha position, for cis hydrocarbons alpha is 2 and trans hydrocarbons alpha is 1.	–0.00006	trans,trans-2,4-Hexadiene	2
40 Beta	Number of atoms into the beta position, for cis hydrocarbons beta is 0 and trans hydrocarbons beta is 1.	–0.30682	trans,trans-2,4-Hexadiene	1
41 C1 ring	All of the ring members must be numbered, the numbering begins from the side chain with lower carbon atoms into a clockwise and this is number of side chain on the first carbon atoms into the ring.	–0.04231	1,1,2-Trimethylcyclopentane	2
42 C2 ring	Number of side chain on the second carbon atoms into the ring.	0.15008	1,1,2-Trimethylcyclopentane	1
43 C3 ring	Number of side chain on the third carbon atoms into the ring.	–0.02775	1,2,3-Trimethylbenzene	1
44 C4 ring	Number of side chain on the fourth carbon atoms into the ring.	0.00619	P-Ethyltoluene	1
45 C5 ring	Number of side chain on the fifth carbon atoms into the ring.	0.08082	1-Ethyl-2,3,6-trimethylbenzol	1

(continued on next page)

Table 2 (continued)

Group j	Description	tbj [Kelvin]	Example	Number of groups
46 C6 ring	Number of side chain on the sixth carbon atoms into the ring.	0.10018	1-Ethyl-2,3,6-trimethylbenzol	1
47 C1=C2	A double bond between first carbon and second carbon into the ring.	−1.23644	Naphthalene	1
48 C2=C3	A double bond between second carbon and third carbon into the ring.	0.02962	Biphenyl	2
49 C3=C4	A double bond between third carbon and fourth carbon into the ring.	−0.76038	4,7,8,9-Tetrahydroindene	2
50 C4=C5	A double bond between fourth carbon and fifth carbon into the ring.	−0.56452	Biphenyl	2
51 C5=C6	A double bond between fifth carbon and sixth carbon into the ring.	0.47129	Naphthalene	2
52 C6=C1	A double bond between sixth carbon and first carbon into the ring.	−0.98187	Biphenyl	2
53 C5=C1	A double bond between fifth carbon and first carbon into the ring.	−1.53187	1-Butyl cyclopentene	1
54 Alpha-6	Number of side chain, oxygen, nitrogen and sulfur atoms in alpha position into the six membered ring of bicyclo compounds.	−0.00013	1,5-Dimethyl naphthalene	2
55 Beta-6	Number of side chain, oxygen, nitrogen and sulfur atoms in beta position into the six membered ring of bicyclo compounds.	−0.00190	2,3-Dimethyl naphthalene	1
56 Gamma-6	Number of side chain, oxygen, nitrogen and sulfur atoms in gamma position into the six membered ring of bicyclo compounds.	−0.01857	2,3-Dimethyl naphthalene	1
57 Alpha-5	Number of side chain, oxygen, nitrogen and sulfur atoms in alpha position into the five membered ring of bicyclo compounds.	−0.60505	3-Methylindole	2
58 Beta-5	Number of side chain, oxygen, nitrogen and sulfur atoms in beta position into the five membered ring of bicyclo compounds.	−0.51369	2-Methylindole	1
59 Left ring side chain	Number of side chain, oxygen, nitrogen and sulfur atoms into the left ring of polycyclic compounds.	0.00001	2,6-Dimethyl quinoline	2
60 Right ring side chain	Number of side chain, oxygen, nitrogen and sulfur atoms into the right ring of polycyclic compounds.	0.00001	2,6-Dimethyl quinoline	1
61 Middle ring side chain	Number of side chain, oxygen, nitrogen and sulfur atoms into the middle ring of polycyclic compounds.	0.04370	1,3-Diphenylbenzene	2
62 Left ring double bond	Number of double bond into the left ring of polycyclic compounds.	0.02789	7,8-Benzoquinoline	2
63 Right ring double bond	Number of double bond into the right ring of polycyclic compounds.	0.03390	7,8-Benzoquinoline	2
64 Mid ring double bond	Number of double bond into the middle ring of polycyclic compounds.	0.37051	7,8-Benzoquinoline	3
65 Oxygen	Number of oxygen atoms.	−0.70342	1,1-Ethanediol diacetate	4
66 Nitrogen	Number of nitrogen atoms.	−1.13838	Ethylenediamine	2
67 Sulfur	Number of sulfur atoms.	−0.94615	1,5-Pentanedithiol	2
68 Chlorine	Number of chlorine atoms.	0.04350	Hexachlorobenzene	6
69 Fluorine	Number of fluorine atoms.	0.04398	Perfluoroheptane	16
70 Bromine	Number of bromine atoms.	0.29198	Tribromomethane	3
71 Iodine	Number of iodine atoms.	1.34393	Diiodomethane	2

$$T_b(K) = T_1(K) + T_2(K) \tag{14}$$

where tbi and tbj reveals i (first-order) and j (second-order) group

Table 3
Q₁ and Q₂ parameters for equations of T₁(K) and T₂(K).

Hydrocarbon atoms	X _i		U _i	V _i
	Existence	Nonexistence		
Carbon	1	0	3.70994	0.84921
Halogen	1	0	12.28953	6.12625
Oxygen	1	0	−3.21100	−1.36996
Nitrogen	1	0	−15.81486	−2.74061
Sulfur	1	0	9.41461	−0.00053
Bromine	1	0	−8.83995	−4.39828
Chlorine	1	0	−3.39475	−2.55976
Fluorine	1	0	−7.11286	−7.88435
Iodine	1	0	−5.48580	−9.16336
Hydrogen	0	1	16.29488	−7.99679

Table 4
Adjustable parameters for cubic equation T₂(K).

Parameter	a	b	c	d
M _w < 40	−2.31803E-03	1.28367E-01	−2.02275	182.46129
40 ≤ M _w < 80	1.87799E-05	−3.44053E-18	−7.72077E-02	203.60493
80 ≤ M _w < 140	−2.47302E-06	2.78437E-04	3.48865E-02	199.26548
140 ≤ M _w < 190	−1.00099E-06	−3.17991E-05	5.58225E-02	189.14878
190 ≤ M _w < 330	6.67661E-07	−4.91020E-04	8.34812E-02	187.75233
M _w ≥ 330	−8.64682E-09	−1.98700E-05	7.08590E-17	189.28368

contributions, respectively, N_i and N_j represents the number of groups i and j, respectively. N_A and N_C is total number of atoms and carbon atoms in a molecule, respectively. M_w is the molecular weight and a, b, c, d are the adjustable parameters for cubic equation (T₂) as a function of the molecular weight, which shown in the following Table 4. Besides, Q₁ and Q₂ are the other adjustable parameters and their values are relevant to the variant atoms present in the molecules for modify those cubic equation and section of second order groups that shown in the following Table 3.

For mathematical models that are nonlinear in the parameters and/or when the maximum likelihood objective function is used, the best estimates of the adjustable parameters must be obtained via iterative methods. These methods, initialized by a tentative first estimate of the parameter vector, θ₀, are based on a strategy for modifying the current estimate and include a convenient but arbitrary termination criterion to stop the algorithm. Thus, the results must be considered only approximate; moreover, if several local minima of the objective function are present, the attainment of the absolute minimum is not guaranteed since these methods usually provide a suboptimal solution, i.e., they converge to a local minimum depending on θ₀. A good solution to this problem is, when possible, to make use of a preliminary estimation procedure to produce a good initial estimate of the parameter values. If this is not possible, it can be useful to compare the results obtained using different initial estimates of the parameter vector [31].

MATLAB software is used to analyze the data, develop the algorithms, and create the mathematical models and applications. The language, tools, and built-in math functions enable you to

explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java.

The coefficients and group contributions were found by the *fminsearch* algorithm in MATLAB programming. *Fminsearch* finds the minimum of a scalar function of several variables (objective function), starting at an initial estimate. This is generally referred to as unconstrained nonlinear optimization. This algorithm uses the Nelder–Mead simplex algorithm as described in Lagarias et al. [32]. *Fminsearch* algorithm uses a simplex of $n + 1$ points for n -dimensional vectors X . The algorithm first makes a simplex around the initial guess X_0 by adding 5% of each component $X_0(i)$ to X_0 , and using these n vectors as elements of the simplex in addition to x_0 (It uses 0.00025 as component i if $X_0(i) = 0$).

The procedure for estimation of normal boiling point by this method can be summarized as follows:

- Step 1: Obtain the total number of atoms, carbon atoms and atom types in the molecule.
- Step 2: Determine the structures of first- and second-order groups in the molecule.
- Step 3: Obtain the needed group contributions from the Tables 1 and 2
- Step 4: Calculate of Q_1 and Q_2 parameters by Eqs. (12) and (13), respectively, from Table 4.
- Step 5: Sum up these group contributions to calculate $T_1(K)$ from the Eq. (10).
- Step 6: Obtain the molecular weight.
- Step 7: Obtain the needed adjustable parameters from the Table 3.
- Step 8: Calculate of cubic equation, $T_2(K)$ from the Eq. (11).
- Step 9: Compute the normal boiling point from the Eq. (14).

Values of first-order group contributions for the 51 groups are given in Table 1. The notation of the table is such that when an atom is bonded to an element other than hydrogen, (1) means a single bond, (2) means 2 single bonds, = means a double bond and ≡ means a triple bond, (ss) or single site means that the group exists in a ring such as in naphthenics, (ds) or double site means that the group exists in a ring such as in aromatics and, AOH means that the OH is connected to the ring.

Isomers are defined as molecules with the same chemical formula but different chemical structures. That is, isomers contain the same number of atoms of each element, but have different arrangements of their atoms in space. There are two main forms of isomers; structural isomers and stereoisomers (spatial isomerism) [33]. In structural isomers, sometimes referred to as constitutional isomers, the atoms and functional groups are joined together in different ways. Structural isomers have different IUPAC¹ names and may or may not belong to the same functional groups [34].

In stereoisomers the bond structure is the same, but the geometrical positioning of the atoms and functional groups in space differs. This class includes enantiomers, which are non-superimposable mirror-images of each other, and diastereomers, which are not. Enantiomers always contain chiral centers and diastereomers often do, but there are some diastereomers that neither are chiral nor contain chiral centers. Cis and trans isomers are a form of stereoisomers describing the relative orientation of functional groups within a molecule. Cis and trans isomers often have different physical properties [35].

Differences between the isomers, in general, arise from the differences in the shape of the molecule or the overall dipole

Table 5

The results with regard to develop the new method.

Method	New method	Training set	Test set
Substances	2036	1833	203
%AARD	1.01075	0.98903	1.20691
AAD (K)	4.35194	4.25127	5.26093
%ARD ≤ 1%	1233	1120	113
1% < %ARD ≤ 2%	458	411	47
2% < %ARD ≤ 3%	247	223	24
3% < %ARD ≤ 4%	90	71	19
%ARD > 4%	8	8	0
AD ≤ 5 K	1328	1210	118
5 K < AD ≤ 10 K	454	402	52
10 K < AD ≤ 15 K	232	203	29
15 K < AD ≤ 20 K	22	18	4
AD > 20 K	0	0	0

moment. These differences can be very small, as in the case of the NBP of straight-chain alkenes, such as 2-pentene, which is 310.8 K in the cis isomer and 309.49 K in the trans isomer. The differences between cis and trans isomers can be larger if polar bonds are present, as in the 1,2-dichloroethenes. The cis isomer in this case has a NBP of 333.45 K, while the trans isomer has a NBP of 320.65 K [28]. By applying the second-order functional groups it would be possible to obtain different NBPs for some structural isomers and stereoisomers. The obtained contribution values for isomers are presented in Table 2.

4. Results and discussion

In most cases, part of the available data is used for model development, this data set is called training set, another part is used for model validation and is called the test set. Normally, training sets give fairly good predictions, because the model itself has been formed using these data sets, but this does not mean that the method is yet safe to use in practical situations. A recommended next step is to test the quality of predictions using an independent test set. The results of the new model was analyzed by splitting the

Table 6

Deviations of the new models for the different classes of hydrocarbons.

Components	Number of Components	New method	
		AAD	%AARD
Alkanes	175	3.65755	0.77380
Cyclo alkanes	107	3.12578	0.74110
Alkenes	166	2.38099	0.67937
Cyclo alkenes	42	4.72519	1.18632
Alkynes	40	2.09492	0.56264
Benzene derivatives	160	4.22411	0.85999
Aromatics	58	3.53772	0.61585
Alcohols	64	6.58256	1.52284
Phenols	32	7.12344	1.42240
Ethers	104	4.08150	1.03578
Ketones	57	3.74212	0.87387
Aldehydes	24	5.02739	1.16442
Esters	141	4.15201	0.89048
Formates	14	2.09240	0.53552
Acids	25	5.86088	1.18827
Others C.H.O compounds	23	6.64100	1.49551
Amines	102	5.35945	1.14139
Others nitrogen compounds	109	5.52026	1.32990
Sulfur compounds	80	4.02528	0.98858
Chlorine compounds	88	5.17911	1.25819
Fluorine compounds	49	3.97291	1.28834
Bromine compounds	37	5.73414	1.35148
Iodine compounds	18	3.03072	0.62551
Others halogen compounds	109	4.91272	1.22905
Others compounds	212	5.10746	1.17732

¹ International Union of Pure and Applied Chemistry.

Table 7
Summary the results of the NBP estimations for different methods.

Method	New method	Joback and Reid	Constantinou and Gani	Marrero and Pardillo	Marrero and Gani
Substances	1628	1628	1628	1628	1628
%AARD	0.95859	3.69394	2.71643	1.89573	2.37457
AAD (K)	4.17953	17.96056	11.74632	8.49819	9.68867
%ARD \leq 1%	1026	368	498	691	723
1% < %ARD \leq 2%	359	326	393	426	410
2% < %ARD \leq 3%	180	255	273	222	236
3% < %ARD \leq 4%	61	154	184	115	98
%ARD > 4%	2	525	280	174	161
AD \leq 5 K	1093	444	566	766	801
5 K < AD \leq 10 K	352	352	402	399	401
10 K < AD \leq 15 K	172	236	260	217	193
15 K < AD \leq 20 K	11	160	137	101	92
AD > 20 K	0	436	263	145	141

whole set of compounds into training and test set, 90% of the data, i.e. 1833 compounds was included in the training set and remaining of the data, i.e. 203 compounds in the test set. The results of new model development and deviations of new method for different types of organic compound classes are presented in the [Tables 5 and 6](#), respectively.

Different methods were compared to the present model to detect any possible weaknesses of the new method. Excellent results were obtained for all of the investigated chemical families. For all the data points included in our database, the objective function defined by the Eq. (A.4) is 1.01075%. [Table 7](#) summarizes the average absolute deviation (AAD) and percentage of average absolute relative deviation (%AARD) between the experimental and predicted values for different methods.

The second-order groups defined in this work are able to distinguish most isomers including cis- and trans-structures. A

comparison of the different models for isomers, is given in [Fig. 2](#), which shows the plots of the average absolute deviation (AAD) for isomers including a large number of structural isomers (see [Fig. 2a](#)), cis and trans isomers (see [Fig. 2b and c](#), respectively). Contributions for compounds containing carbon, hydrogen, oxygen, nitrogen, sulfur, fluorine, chlorine, bromine and iodine atoms were also reported, and the results were compared to other methods. [Table 8](#) shows the average absolute deviation (AAD) and percentage of average absolute relative deviation (%AARD) for different types of organic compound classes.

5. Conclusion

A group contribution method for the estimation of the NBP of organic compounds was developed. The new second-order groups used in this work may distinguish between most isomers including

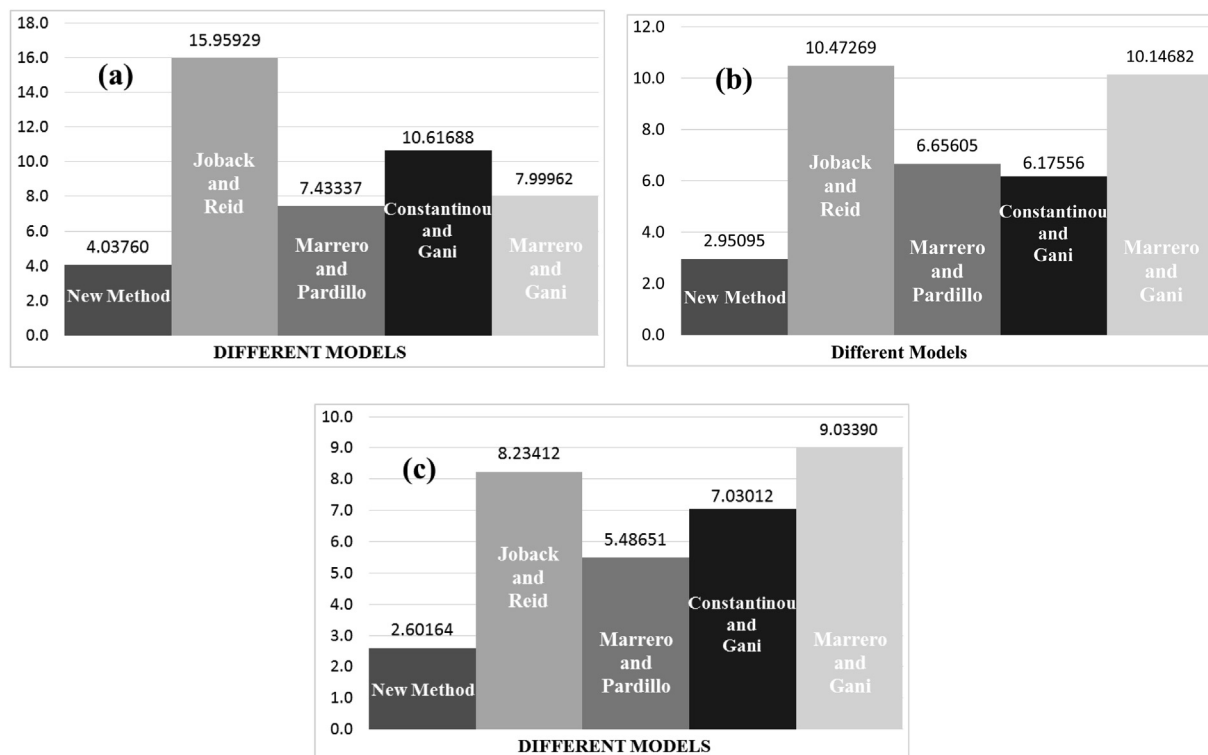


Fig. 2. A comparison of the different models for hydrocarbon isomers, average absolute deviation (AAD) for (a) 1084 structural isomers, (b) 67 cis isomers and (c) 78 trans isomers [28–30].

Table 8

Deviations of the different models for the different types of hydrocarbons.

Components	NC ^a	New method		Joback and Reid		Constantinou and Gani		Marrero and Pardillo		Marrero and Gani	
		AAD	%AARD	AAD	%AARD	AAD	%AARD	AAD	%AARD	AAD	%AARD
Alkanes	175	3.6576	0.7738	28.1716	4.9678	10.2879	1.9200	7.9663	1.4488	10.1646	2.0780
Cyclo alkanes	107	3.1258	0.7411	16.1762	3.0029	13.8683	3.0482	12.4492	3.1328	10.4992	2.1927
Alkenes	161	2.2796	0.6426	8.6351	2.2418	5.8308	1.5155	3.9252	1.0418	3.6926	0.9933
Cyclo alkenes	26	3.2686	0.8752	8.8728	2.5166	7.3426	1.9756	15.4689	4.0301	7.0496	1.8553
Alkynes	37	2.2299	0.5968	12.7209	3.1240	13.8054	3.5030	6.0044	1.5020	5.6551	1.7195
Benzene derivatives	148	4.1028	0.8257	35.9986	6.0925	10.7214	1.9265	10.0645	1.8045	7.4279	1.3787
Aromatics	47	3.7881	0.6586	15.6203	2.6579	22.5996	3.9129	10.1958	1.7274	9.0507	1.6489
Alcohols	59	6.2638	1.4507	24.8055	5.8151	10.7604	2.5091	14.5576	3.4049	5.7750	1.3722
Phenols	32	7.1234	1.4224	15.4724	3.0094	9.3878	1.8373	7.6740	1.5345	5.3262	1.0513
Ethers	91	3.9078	1.0075	9.4572	2.3408	11.3270	2.9801	4.7429	1.1801	5.3566	1.4658
Ketones	42	3.7176	0.8776	9.3834	2.0079	8.0835	1.8714	4.6888	1.0674	4.7082	1.0619
Aldehydes	19	4.4657	1.0487	10.7371	2.5042	9.8188	2.4650	3.7856	0.8150	7.0727	1.8038
Esters	117	4.0910	0.8532	20.1295	3.6968	8.7110	1.7974	8.4900	1.6191	7.5715	1.5983
Acids	24	5.9231	1.1958	12.4863	2.4965	8.8717	1.6087	5.0195	0.9687	10.0996	1.9377
Others C.H.O compounds	21	6.8381	1.5387	12.7237	2.9344	9.3549	2.1330	10.2795	2.2991	8.7045	1.9996
Amines	84	5.2813	1.1474	14.8350	3.3465	15.8591	3.4628	5.0348	1.0916	7.2100	1.6514
Others nitrogen compounds	45	4.9662	1.1353	16.3670	3.2036	12.9636	2.8147	8.7358	1.8383	10.8137	2.4387
Sulfur compounds	44	4.0695	0.9751	15.5398	3.7450	9.6501	2.3812	5.6717	1.3573	4.7998	1.2484
Chlorine compounds	65	5.3529	1.2348	16.8124	3.9334	9.1428	2.0979	9.5242	2.2157	7.1645	1.6396
Fluorine compounds	49	3.9729	1.2883	11.5881	3.6706	22.0825	7.4797	9.7558	3.2668	58.7693	17.9857
Bromine compounds	27	5.8307	1.3240	12.8455	2.9543	4.9534	1.0575	11.4274	2.4574	3.4431	0.7740
Iodine compounds	16	3.0474	0.6063	16.2138	3.6100	28.4793	8.1401	12.3962	2.6519	6.6298	1.3111
Others halogen compounds	61	4.7193	1.1806	16.6422	4.7858	16.1313	4.6918	10.9282	2.9499	16.3705	5.0052
Others compounds	131	5.1042	1.1663	16.3170	3.5198	15.5633	3.8082	10.7865	2.4086	14.2764	3.6074

^a The number of components to which the methods was applicable.

cis- and trans-structures of organic compounds to give more accurate predictions for NBP. The predictions are based exclusively on the molecular structure of the compound. The ability to distinguish and give different results for different isomers like ortho, meta and para isomers of an aromatic ring would be one of the main advantages of this work. Testing the accuracy of this work with 2036 organic compounds of the data bank shows that the new method is more precise than other second order existing methods. The results of the Tables 7 and 8 clearly show that the new method is in most cases far more accurate and never significantly worse than others.

Appendix A. Error functions

The mathematical relations for calculation of the absolute deviation (AD), percentage of absolute relative deviation (%ARD), average absolute deviation (AAD) and percentage of average absolute relative deviation (%AARD) are presented as follows:

$$AD(K) = |T_{b,pred} - T_{b,exp}| \quad (A.1)$$

$$\%ARD = \frac{|T_{b,pred} - T_{b,exp}|}{T_{b,exp}} \times 100 \quad (A.2)$$

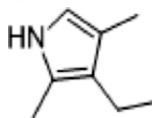
$$AAD = \frac{1}{N} \sum_N |T_{b,pred} - T_{b,exp}| \quad (A.3)$$

$$\%AARD = \frac{1}{N} \sum_N \frac{|T_{b,pred} - T_{b,exp}|}{T_{b,exp}} \times 100 \quad (A.4)$$

where pred and exp stand for the predicted values by methods and its experimental values, respectively. N is the number of data points.

Example 1

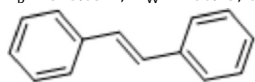
Estimation of the NBP of 2,4-Dimethyl-3-ethylpyrrole (experimental value: $T_b = 470.45$ K, $M_W = 123.198$, $C_8H_{13}N$).



Hydrocarbon atoms	X _i	U _i	V _i	
Carbon	1	3.70994	0.84921	
Nitrogen	1	-15.81486	-2.74061	
Q ₁ = N _C + ∑ _{i=1} ^{N_A} X _i U _i = -4.10		Q ₂ = N _C + ∑ _{i=1} ^{N_A} X _i V _i = 6.10		
First-order groups	Occurrences	Contribution		
CH ₃ (1)	3	13.0945		
CH ₂ (2)	1	22.9831		
=CH (ds) (2)	1	25.8530		
=C (ds) (3)	3	34.2365		
NH (ss) (2)	1	44.9923		
∑ _i N _i (tbi) = 235.82				
Second-order groups	Occurrences	Contribution		
CHa- a ε (0-3)	2	0.06768		
CHaCHb- a,b ε (0-3)	1	-0.10259		
Ring	1	-0.05102		
C1 ring	1	-0.04231		
C2 ring	1	0.15008		
C3 ring	1	-0.02775		
C4 ring	1	0.00619		
C2=C3	1	0.02962		
C5=C1	1	-1.53187		
Nitrogen	1	-1.13838		
∑ _j N _j (tbj) = -2.57				
Parameter	a	b	c	d
80 ≤ M _W < 140	-2.47302E-06	2.78437E-04	3.48865E-02	199.26548
Q ₂ (aM _W ³ + bM _W ² + cM _W) + d = 223.08				
T _b exp	T _b cal	AD		%ARD
470.45	469.47	0.97982		0.20827

Example 2

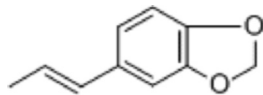
Estimation of the NBP of Trans-1,2-Diphenylethene (experimental value: $T_b = 579.65$ K, $M_W = 180.25$, $C_{14}H_{12}$).



Hydrocarbon atoms	X _i	U _i	V _i	
Carbon	1	3.70994	0.84921	
Hydrogen	0	16.29488	-7.99679	
Q ₁ = N _C + ∑ _{i=1} ^{N_A} X _i U _i = 17.70		Q ₂ = N _C + ∑ _{i=1} ^{N_A} X _i V _i = 14.84		
First-order groups	Occurrences	Contribution		
==CH (2)	2	29.2596		
==CH (ds) (2)	10	25.8530		
=C (ds) (3)	2	34.2365		
∑ _i N _i (tbi) = 385.52				
Second-order groups	Occurrences	Contribution		
CHaCHb- a,b ∈ (0-3)	1	-0.10259		
Alkene bond	1	2.06190		
Ring	2	-0.05102		
C-C (out)	3	0.03212		
Double bond	1	-1.24364		
Trans-	1	0.29954		
Alpha	1	-0.00006		
Beta	1	-0.30682		
C1 ring	2	-0.04231		
C2=C3	2	0.02962		
C4=C5	2	-0.56452		
C6=C1	2	-0.98187		
Left ring side chain	1	0.00001		
Right ring side chain	1	0.00001		
Left ring double bond	3	0.02789		
Right ring double bond	3	0.03390		
∑ _j N _j (tbj) = -2.23				
Parameter	a	b	c	d
140 ≤ M _W < 190	-1.00099E-06	-3.17991E-05	5.58225E-02	189.14878
Q ₂ (aM _W ³ + bM _W ² + cM _W) + d = 236.17				
T _b exp	T _b cal	AD		%ARD
579.65	582.19	2.54906		0.43975

Example 3

Estimation of the NBP of Isosafrole (experimental value: $T_b = 525.15$ K, $M_W = 162.188$, $C_{10}H_{10}O_2$).



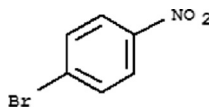
Hydrocarbon atoms	X_i	U_i	V_i
Carbon	1	3.70994	0.84921
Oxygen	1	-3.21100	-1.36996
$Q_1 = N_C + \sum_{i=1}^{N_A} X_i U_i = 10.49$		$Q_2 = N_C + \sum_{i=1}^{N_A} X_i V_i = 9.47$	
First-order groups	Occurrences	Contribution	
CH ₃ (1)	1	13.0945	
CH ₂ (ss) (2)	1	23.3777	
=CH (2)	2	29.2596	
=CH (ds) (2)	3	25.8530	
=C (ds) (3)	3	34.2365	
O (ss) (2) (Ether)	2	38.9549	
$\sum_i N_i(tbi) = 353.17$			

Example 3 (continued)

First-order groups	Occurrences	Contribution		
Second-order groups	Occurrences	Contribution		
CH ₃ –CH=	1	0.16043		
CHa– a ε (0–3)	1	0.06768		
Ring	2	–0.05102		
C=C	1	0.00001		
Double bond	1	–1.24364		
C3 ring	2	–0.02775		
C5 ring	1	0.08082		
C2=C3	2	0.02962		
C4=C5	2	–0.56452		
C6=C1	2	–0.98187		
Gamma-6	1	–0.01857		
Alpha-5	2	–0.60505		
Left ring side chain	1	0.00001		
Right ring side chain	2	0.00001		
Left ring double bond	3	0.02789		
Oxygen	2	–0.70342		
$\sum_j N_j(tbj) = -5.16$				
Parameter	a	b	c	d
$140 \leq M_W < 190$	–1.00099E-06	–3.17991E-05	5.58225E-02	189.14878
$Q_2(aM_W^3 + bM_W^2 + cM_W) + d = 226.56$				
T _b exp	T _b cal	AD		%ARD
525.15	525.54	0.39666		0.07553

Example 4

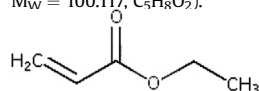
Estimation of the NBP of p-Bromonitrobenzene (experimental value: $T_b = 532.35$ K, $M_W = 202.007$, $C_6H_4BrNO_2$).



Hydrocarbon atoms	X _i	U _i	V _i	
Carbon	1	3.70994	0.84921	
Halogen	1	12.28953	6.12625	
Oxygen	1	-3.21100	-1.36996	
Nitrogen	1	-15.81486	-2.74061	
Bromine	1	-8.83995	-4.39828	
Q ₁ = N _C + ∑ _{i=1} ^{N_A} X _i U _i = -5.86		Q ₂ = N _C + ∑ _{i=1} ^{N_A} X _i V _i = 4.46		
First-order groups	Occurrences		Contribution	
=CH (ds) (2)	4		25.8530	
=C (ds) (3)	2		34.2365	
NO ₂ (1)	1		61.9870	
Br (1)	1		70.3924	
∑ _i N _i (tbi) = 304.26				
Second-order groups	Occurrences		Contribution	
Ring	1		-0.05102	
Double bond	1		-1.24364	
C1 ring	1		-0.04231	
C4 ring	1		0.00619	
C2=C3	1		0.02962	
C4=C5	1		-0.56452	
C6=C1	1		-0.98187	
Oxygen	2		-0.70342	
Nitrogen	1		-1.13838	
Bromine	1		0.29198	
∑ _j N _j (tbj) = -5.10				
Parameter	a	b	c	d
190 ≤ M _W < 330	6.67661E-07	-4.91020E-04	8.34812E-02	187.75233
Q ₂ (aM _W ³ + bM _W ² + cM _W) + d = 198.16				
T _{bexp}	T _{bcal}	AD		%ARD
532.35	532.34	1.89814E-06		3.56559E-07

Example 5

Estimation of the NBP of Ethyl acrylate (experimental value: $T_b = 373.15$ K, $M_W = 100.117$, $C_5H_8O_2$).



Hydrocarbon atoms	X _i	U _i	V _i	
Carbon	1	3.70994	0.84921	
Oxygen	1	-3.21100	-1.36996	
Q ₁ = N _C + ∑ _{i=1} ^{N_A} X _i U _i = 5.49		Q ₂ = N _C + ∑ _{i=1} ^{N_A} X _i V _i = 4.47		
First-order groups	Occurrences		Contribution	
CH ₃ (1)	1		13.0945	
CH ₂ (2)	1		22.9831	
=CH ₂ (1)	1		7.1245	
=CH (2)	1		29.2596	
COO (2) (Ester)	1		97.4615	
∑ _i N _i (tbi) = 169.92				
Second-order groups	Occurrences		Contribution	
CH ₂ =CH-	1		1.06565	
CH ₃ -CH ₂ -O-	1		0.419326	
Double bond	2		-1.24364	
Oxygen	2		-0.70342	
∑ _j N _j (tbj) = -2.40				
Parameter	a	b	c	d
80 ≤ M _W < 140	-2.47302E-06	2.78437E-04	3.48865E-02	199.26548
Q ₂ (aM _W ³ + bM _W ² + cM _W) + d = 216.29				
T _b exp	T _b cal		AD	%ARD
373.15	372.97		0.17927	0.04804

List of symbols

a, b, c and d	Adjustable parameters
tbk, tbbk, tb1k, tbi	First order group contributions
tb2j, tbj	Second order group contributions
N_k, N_i	Number of first order groups of type k and i, respectively
M_j, N_j	Number of second order groups of type j
K	Kelvin
H	Enthalpy
S	Entropy
T	Absolute Temperature
V	Volume
P	Pressure
R	Universal gas constant
M_W	Molecular weight
G_{ij}	Group interaction value
N_C	Total number of carbon atoms
N_A	Number of atom types in molecule
N	Number of data points in each data set or subset
AD(K)	Absolute deviation
%ARD	Percentage of absolute relative deviation
AAD	Average absolute deviation
%AAD	Percentage of average absolute relative deviation

Subscripts

b	Boiling
vap	Vaporization
c	Critical
pred	Predicted
exp	Experimental

References

- [1] A. Fredenslund, L. Russell, J.M. Prausnitz, J.M. Prausnitz, Group-contribution estimation of activity coefficients in nonideal liquid mixtures, *AIChE J.* 21 (1975) 1086.
- [2] E.J. Maginn, *Molecular Theory and Modeling Chemical Engineering*, University of Notre Dame, IN 46556 USA, 1997, pp. 9–11.
- [3] E.J. Wood, A. Myers, *Essential Chemistry for Biochemistry*, second ed., The Biochemical Society, 1991, pp. 25–29.
- [4] T. Makee, J.R. Makee, *The Molecular Basis of Life*, third ed., McGraw-Hill, New York, 2004, pp. 8–10.
- [5] W.J. Lyman, W.F. Reehl, D.H. Rosenblatt, *Handbook of Chemical Property Estimation Methods*, McGraw-Hill, New York, 1982.
- [6] R.C. Reid, J.M. Prausnitz, B.E. Poling, *Properties of Gases and Liquids*, fourth ed., McGraw-Hill, New York, 1987.
- [7] C.H. Fisher, Calculation of critical temperatures from the number of carbons in organic molecules, *J. Am. Oil Chem. Soc.* 67 (1990) 101–102.
- [8] R.C. Mebane, C.D. Williams, T.R. Rybolt, Correlations of critical properties with computed molecular surface areas for 11 homologous series of 118 organic compounds, *Fluid Ph. Equilib.* 124 (1996) 111–122.
- [9] M.D. Koretsky, *Engineering and Chemical Thermodynamics*, second ed., John Wiley, 2013, p. 531.
- [10] M.R. Riazi, *Characterization and Properties of Petroleum Fractions*, Philadelphia, PA, first ed., 2005, p. 31, 88–89.
- [11] R. Sanghvi, S.H. Yalkowsky, Estimation of normal boiling point of organic compounds, *Ind. Eng. Chem. Res.* 45 (2006) 2856–2861.
- [12] H. Tadeusz, Prediction of thermodynamic properties of the systems formed by n-alkanes, aliphatic monoethers, and 1-chloroalkanes using a cell-hole group contribution model, *J. Phys. Chem. B* 108 (2004) 2383–2397.
- [13] L.P. Rebelo, J.N. Canongia, J.M. Esperanca, E. Filipe, On the critical temperature normal boiling point and vapor pressure of ionic liquids, *J. Phys. Chem. B* 109 (2005) 6040–6043.
- [14] J.D. Valderrama, P.A. Robles, Critical properties, normal boiling temperature and acentric factors of fifty ionic liquids, *Ind. Eng. Chem. Res.* 46 (2007) 1338–1344.
- [15] J.C. Dearden, Quantitative structure-property relationships for prediction of boiling point, vapor pressure and melting point, *Environ. Toxicol. Chem.* 22 (2003) 1696.
- [16] C.M. Guldberg, *Z. Phys. Chem.* 5 (1890) 374; J.R. Partington, *An Advanced Treatise on Physical Chemistry*, 1, Longmans Green and Co. Ltd, London, 1951, p. 646.
- [17] J. Corner, *Trans. Faraday Soc.* 36 (1940) 784.
- [18] T.K. Ghosh, D.H.L. Prasad, N.V.K. Dutt, K.Y. Rani, *Viscosity of Liquids*, Springer, Dordrecht, The Netherlands, 2007, pp. 407–408.
- [19] K.G. Joback, R.C. Reid, Estimation of pure component properties from group contributions, *Chem. Eng. Commun.* 57 (1987) 233–243.
- [20] A.L. Lydersen, *Estimation of Critical Properties of Organic Compounds*, University of Wisconsin College Engineering, Eng. Exp. Stn. Rep. 3, Madison, Wisconsin.
- [21] S.E. Stein, R.L. Brown, Estimation of normal boiling points from group contributions, *J. Chem. Inf. Comput. Sci.* 34 (1994) 581–587.
- [22] J. Marrero-Morejon, E. Pardillo-Fontdevila, Estimation of pure compound properties using group interaction contributions, *AIChE J.* 45 (3) (1999) 615–621.
- [23] Y. Nannoolal, J. Rarey, J. Ramjugernath, Estimation of pure component properties Part 2. Estimation of critical property data by group contribution, *Fluid Ph. Equilib.* 252 (1–2) (2007) 1–27.
- [24] L.R. Constantinou, R. Gani, New group contribution method for estimating properties of pure compounds, *AIChE J.* 40 (10) (1994) 1697–1710.
- [25] J.M. Prausnitz, R.N. Lichtenthaler, E.G.D. Azevedo, *Molecular Thermodynamics of Fluid Phase Equilibria*, third ed., Prentice-Hall, Upper Saddle River, 1999.
- [26] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, *The Properties of Gases and Liquids*, fifth ed., McGraw Hill, New York, 2000.
- [27] R.S. Boethling, D. Mackay, *Handbook of Property Estimation Methods for Chemicals*, CRC Press, Boca Raton, FL, 2000.
- [28] R.E. Pulley, T.E. Daubert, *Technical Data Book-petroleum Refining*, sixth ed., American Petroleum Institute, Washington, D.C, 1997.
- [29] C.L. Yaws, *Handbook of Chemical Compound Data for Process Safety*, Elsevier Science & Technology Books, 1997.
- [30] E.B. Poling, J.M. Prausnitz, J.P. O'Connell, *The Properties of Gases and Liquids*, fifth ed., McGraw-Hill, Inc., New York, 2001, Table C-1.
- [31] F. Caccavale, M. Iamarino, F. Pierri, V. Tufano, *Control and Monitoring of Chemical Batch Reactors*, Springer London Dordrecht Heidelberg New York, 2011, p. 69.
- [32] J.C. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, Convergence properties of the Nelder-Mead simplex method in low dimensions, *SIAM J. Optim.* 9 (Number.1) (1998) 112–147.
- [33] R.H. Petrucci, R.S. Harwood, F.G. Herring, *General Chemistry*, eighth ed., Prentice-Hall, 2002, p. 91.
- [34] J. Smith, J. Gorzynski, *General Organic and Biological Chemistry*, first ed., The McGraw-Hill Companies, 2010, p. 450.
- [35] L. Ernest, E.H. Wilen, S.H. Wilen, *Stereochemistry of Organic Compounds*, Wiley Interscience, 1994, pp. 52–53.
- [36] J. Marrero, R. Gani, Group-contribution based estimation of pure component properties, *Fluid Ph. Equilib.* 183–184 (2001) 183–208.