Group-contribution based estimation of pure component properties

Jorge Marrero [a], Rafiqul Gani [b],*

[a] *Center of Pharmaceutical Chemistry, P.O. Box 16042, Havana, Cuba*
[b] *CAPEC, Department of Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Abstract

A new method for the estimation of properties of pure organic compounds is presented. Estimation is performed at three levels. The primary level uses contributions from simple groups that allow describing a wide variety of organic compounds, while the higher levels involve polyfunctional and structural groups that provide more information about molecular fragments whose description through first-order groups is not possible. The presented method allows estimations of the following properties: normal boiling point, critical temperature, critical pressure, critical volume, standard enthalpy of formation, standard enthalpy of vaporization, standard Gibbs energy, normal melting point and standard enthalpy of fusion. The group-contribution tables have been developed from regression using a data set of more than 2000 compounds ranging from $C = 3$–60, including large and complex polycyclic compounds. Compared to the currently used group-contribution methods, the new method makes significant improvements both in accuracy and applicability. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Group-contribution; Property prediction; Organic compounds

1. Introduction

The basis for the design and simulation of many chemical processing units is a set of physical and thermodynamic properties of compounds in a process that undergoes some form of transformation. It is not always possible, however, to find experimental values of properties for the compounds of interest in the literature. Since, it is not practical to measure them as the need arises, estimation methods are generally employed in this and other similar situations.

For the estimation of properties of pure compounds, group-contribution methods such as those given by Joback and Reid [1], Lydersen [2], Ambrose [3], Klincewicz and Reid [4], Lyman et al. [5], and Horvath [6] have been widely used. In these methods, the property of a compound is a function of structurally-dependent parameters, which are determined by summing the frequency of each group occurring in the molecule times its contribution. These methods provide the advantage of quick estimates

---

without requiring substantial computational resources. Many of these methods are, however, of questionable accuracy, and are unable to distinguish among isomers and have limited applicability due to the oversimplification of the molecular structure representation as a result of the use of a simple group-contribution approach and a relatively small data set used for estimation of group-contributions.

To overcome these limitations, several attempts have been reported in the literature. Constantinou et al. [7,8] proposed a quite complex estimation technique, which is based on conjugate forms (alternative formal arrangements of valence electrons). This technique provides accurate estimations of several properties of pure compounds and allows capturing the differences among isomers. However, the generation of conjugate forms is a nontrivial issue and requires a symbolic computing environment (Prickett et al. [9]). A less complex method has been proposed by Constantinou and Gani [10], which performs the estimation at two levels: the basic level uses contributions from first-order simple groups, while the second level uses a small set of second-order groups having the first-order groups as building blocks. The role of the second-order groups is to consider, to some extent, the proximity effects and to distinguish among isomers. Marrero-Morejón and Pardillo-Fontdevila [11] proposed another technique that considers the contributions of interactions between bonding groups instead of the contributions of simple groups, which allows the distinction of a large number of isomers.

Despite the advantages of the above-mentioned methods, their ranges of applicability are still quite restricted. Properties of large, complex and polyfunctional substances of interest in biochemical and environmental studies cannot be accurately estimated by using the current (available) methods. Due to the relatively small data sets used in the development of these methods, which usually includes just a few hundred relatively simple compounds, the predictive capability usually breaks down when dealing with large, polycyclic or polyfunctional molecules. Also, most of the existing group-contribution techniques do not include suitable groups for representing complex molecules such as the ones of biochemical or environmental importance. Motivated by these drawbacks, our efforts have been focused on developing a new group-contribution method that allows more accurate and reliable estimations of a wide range of chemical substances including large and complex compounds.

In our method, the estimation is performed at three levels. The basic level has a large set of simple groups that allows describing a wide variety of organic compounds. However, these groups capture only partially the proximity effects and are unable to distinguish among isomers. For this reason, the first level of estimation is intended to deal with simple and monofunctional compounds. The second level involves groups that permit a better description of proximity effects and differentiation among isomers. The second level of estimation is consequently intended to deal with polyfunctional, polar or nonpolar compounds of medium size, $C = 3$–6, and aromatic or cycloaliphatic compounds with only one ring and several substituents. The third level has groups that provide more structural information about molecular fragments of compounds whose description is insufficient through the first and second level groups. The third level of estimation allows estimation of complex heterocyclic and large ($C = 7$–60) polyfunctional acyclic compounds. The ultimate objective of the proposed multilevel scheme is to enhance the accuracy, reliability and the range of application for a number of important pure component properties.

2. Development of the new method

In the new method, the molecular structure of a compound is considered to be a collection of three types of groups: first-order groups, second-order groups and third-order groups. The first-orders groups

are intended to describe a wide variety of organic compounds, while the role of the second and third-order groups is to provide more structural information about molecular fragments of compounds whose description is insufficient through the first-order groups. Thus, the estimation is performed at three successive levels, where, the first level provides an initial approximation that is improved at the second level and further refined (if necessary) at the third level. The ultimate objective of this multilevel scheme of estimation is to enhance the accuracy, reliability and the range of application for a number of important pure component properties.

2.1. First-order groups

The basic or primary level of estimation has a large set of simple groups that allow describing a wide variety of organic compounds. For the identification of the first-order groups, we have assumed the following criteria.

1. As it has been mentioned above, the set of groups should allow the representation of a wide variety of chemical classes, including both aromatic and aliphatic hydrocarbons, alcohols, phenols, ketones, aldehydes, acids, esters, ethers, amines, anilines and many other different classes of organic compounds. Polyfunctional, large, complex and heterocyclic compound should be also described by this set of groups although a good performance in the property estimation is not expected.
2. Each group should be as small as possible because very large groups are not desirable. For instance, between groups such as CH_3COO and CH_3COOCH_2, the former should be incorporated into the set of groups and the latter should be left out. However, special care should be taken as it is not always a good practice to look for a minimal set of groups. The definition of excessively small-sized groups may adversely affect the performance of the method as observed earlier ([4,10] Larsen et al. 1987). For example, the description of esters by using only a single COO group does not give results as good as the one obtained by using several ester groups such as CH_3COO, CH_2COO, CHCOO and CCOO. For this reason, this series of ester groups has been included in the set of first-order groups as well as the single COO group, which is used for other esters that cannot be represented by the above four groups, e.g. acrylates. The same criterion has been applied for the identification of groups used for the representation of ketones, amines, sulfides and other chemical families.
3. A detailed first-order representation of aromatic compounds should be provided at a first level of estimation. In this case, the set of first-order groups contain groups of the form aC–R, such as aC–CO or aC–COO, to describe a large number of different aromatic substituents. In addition, two specific groups, aN and aCH, have been included to allow for the representation of pyridines and other nitrogen-containing aromatics as well as the aromatic carbon without substitution. Furthermore, three different corrections of the form aC has been defined to differentiate among (a) carbon atoms shared by different aromatic rings in a same fused system, e.g. naphthalene, (b) carbon atoms shared by both aromatic and nonaromatic rings in a same fused system, e.g. indane, and (c) any other substituted aromatic carbon that does not fall in the above categories, e.g. biphenyl.
4. The set of first-order groups should allow the distinction between groups occurring in cyclic and acyclic structures. It was attempted to describe both types of structures by using a common set of groups. However, it was found that much better property estimation is achieved for both cyclic and acyclic structures by using separate ring and nonring groups. Thus, we have included different ring-specific groups to describe nonaromatic cyclic structures such as cycloalkanes, cycloalkenes, cyclic ethers, thiophenes and others.

5. First-order groups should describe the entire molecule. In other words, there should be no fragment of a given molecule that cannot be represented by first-order groups. It should be also required that no atom of the molecule can be included in more than one group, which implies that no group is allowed to overlap any other first-order group. These requirements are equivalent to consider the set of groups describing a given compound as a full and disjoint splitting of the molecular structure. For example, trimethylurea should be represented by (1) NHCON, (3) CH_3. To use groups CH_3NH or CH_3N would be wrong because the nitrogen atoms would be covered more than one time.

6. The contribution of any first-order group should be independent of the molecule in which the group occurs, which satisfies one of the fundamental principles of the group-contribution approach.

Based on the above criteria of identification, a comprehensive set of first-order groups has been defined. The list of these groups, along with sample assignments and group occurrences, is presented in Appendix A (Table 6). It should be mentioned that some rules have to be followed in order to assign correctly the groups that occur in a given compound and avoid ambiguous descriptions. Having this in mind, we define the "weight" of a group as the sum of atomic weights of the atoms involved in that group. It is assumed that heavier groups hold more information about the molecular structure than lighter groups. Consequently, the golden rule is that, if the same fragment of a given compound is related to more than one group, the heavier group must be chosen to represent it instead of the lighter groups. For example, 3,3-dimethyl-2-butanone, $(CH_3)_3C–CO–CH_3$, is described correctly in the following way: (1) CH_3CO, (3) CH_3, (1) C. To describe it as (1) CCO, (4) CH_3 would be wrong. As shown in Fig. 1, there are only two exceptions for this rule, which in turn constitute the second and third rules. One is in the case of aromatic substituents for which groups of form aC–R must be used. An example is acetophenone that is rightly represented by (1) aC–CO, (5) aCH, (1) CH_3 and wrongly by (1) CH_3CO, (1) aC, (5) aCH. Similarly, it must be used (10) aCH, (1) aC–CH_2, (1) aC–O for describing benzyl phenyl ether rather than (10) aCH, (2) aC, (1) CH_2O. The other exception occurs for ureas and amides, for which specific functional groups are provided. For instance, the correct representation of 1,3-dimethyl urea must be (1) NHCONH, (2) CH_3 rather than (1) $CONHCH_3$, (1) CH_3, although the specific amide group is slightly heavier than the urea group.
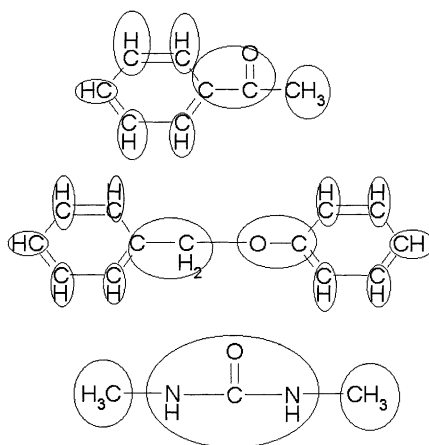

Fig. 1. Examples of correct descriptions by using first-order groups.

The proposed first-order groups, however, capture partially the proximity effects and differences among isomers. For this reason, the first level of estimation is intended to deal with simple and monofunctional compounds and further refinements are required to provide more accurate estimations for more complex compounds.

2.2. Second-order groups

The second level involves groups that permit a better description of polyfunctional compounds and differentiation among isomers. The following criteria have been considered for the identification of the second-order groups.

1. At the second level of approximation, there can be compounds that do not need any second-order group. The properties of these compounds (simple and monofunctional) are expected to be satisfactorily estimated after a first level of approximation.
2. Contrary to the case of first-order groups, the entire molecule does not need to be described with second-order groups. Second-order groups are intended to describe only molecular fragments that could not be adequately described by first-order groups, and thereby yielded a poor estimation at the first level. In this case, the second level makes use of groups that allow an improved representation of polyfunctional compounds. For example, groups such as $OH–CH_n–COO$, $CH_m(NH_n)–COOH$ or $NH_2–CH_m–CH_n–COOH$ have been defined to provide a better description of lactates and aminoacids, which usually results in poor estimates at a first level of approximation. Other molecular fragments that are not completely described at the first level are the alicyclic substituents for which groups such as $CH_{cyc}–OH$, $CH_{cyc}–NH_2$, $CH_{cyc}–COO$, $CH_{cyc}–COOH$ or $CH_{cyc}–SH$ have been included in the set of second-order groups.
3. As it has been suggested above, the set of second-order groups should allow the differentiation among isomers. Accordingly, specific groups are provided with this objective in mind. These groups allow the distinction of isomers not only in alkenes, alkenes and other open-chain structures, but also in aromatic compounds for which special groups such as $AROMRINGs^1s^2$, $AROMRINGs^1s^3$, $AROMRINGs^1s^4$, $AROMRINGs^1s^2s^3$ or $AROMRINGs^1s^2s^4$ have been included.
4. Second-order groups should be allowed to overlap each other. That is, a specific atom of the molecule may be included in more than one group, contrary to the case of first-order groups. The contribution of any group should be equal in whichever molecule the group occurs, which makes it possible to satisfy one of the fundamental principles of the group-contribution approach in the case of second-order groups.

The resulting list of second-order groups is given in Appendix A (Table 7). Sample assignments along with occurrences for each group are also provided. In order to explain how to use second-order groups, some guidance is provided below. Unlike first-order groups, second-order groups are allowed to overlap when they have atoms in common. It is necessary, however, to prevent a situation in which one group overlaps completely another group since it would lead to a redundant description of a same molecular fragment. Fig. 2 illustrates different situations of overlapping. The case of ʟ-alanine is an example of complete overlapping since the group $CH_m(NH_n)–COOH$ contains the group $CH_m–COOH$. Obviously, this situation should not be allowed and consequently the principal rule for assigning second-order groups is that, if one group embodies completely another group, the former must be chosen rather than the latter. For instance, in 2-amino-1-butanol, the only second-order group is $CH_m(OH)CH_n(–)$. The
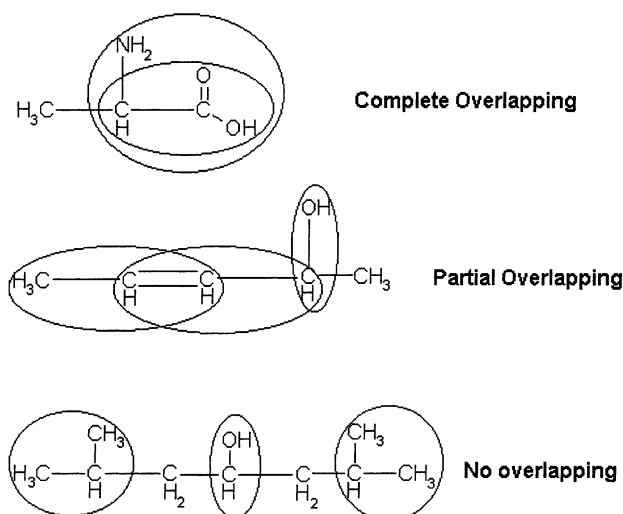
Fig. 2. Different situations related to overlapping of second-order groups.

group CHOH is not allowed in this case. Similarly, for 3,5-dimethylaniline, the only second-order group is AROMRINGs^{1}s^{3}s^5 and the groups AROMRINGs^{1}s^3 or AROMRINGs^{1}s^4 are not allowed.

The proposed second-order groups are, however, unable to provide a good representation of compounds containing more than one ring as well as, in some case, open-chain polyfunctional compounds with more than four carbon atoms in the main chain. For this reason, a further level is required to allow a better description of these types of compounds.

2.3. Third-order groups

The criteria used for the identification of third-order groups are analogous to those used for second-order groups except for, as it has been already mentioned, the types of compounds intended to be represented at the third level of approximation. Fig. 3 shows examples of third-order groups that have been included to describe molecular fragments in polycyclic compounds. These groups allow a quite detailed representation of systems of fused aromatic rings, systems of fused aromatic and nonaromatic rings, and systems of nonfused rings joined by chains in which can occur different functional groups. The list of third-order groups is given in Appendix A (Table 8) together with sample assignments and occurrences for each group. The principal rule for using these correctly groups is similar to that previously specified for second-order groups.

2.4. Proposed model

The property-estimation model has the form of the following equation:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \tag{1}$$

AROM.FUSED[2] + AROM.FUSED[4a] + AROM.FUSED[2]

AROM.FUSED[2] + 2 (aC-CHncyc)

aC-(CHn)m-aC (m>1)

Fig. 3. Examples of third-order groups used for representing polycyclic compounds.

In Eq. (1), C_i is the contribution of the first-order group of type-i that occurs N_i times, D_j the contribution of th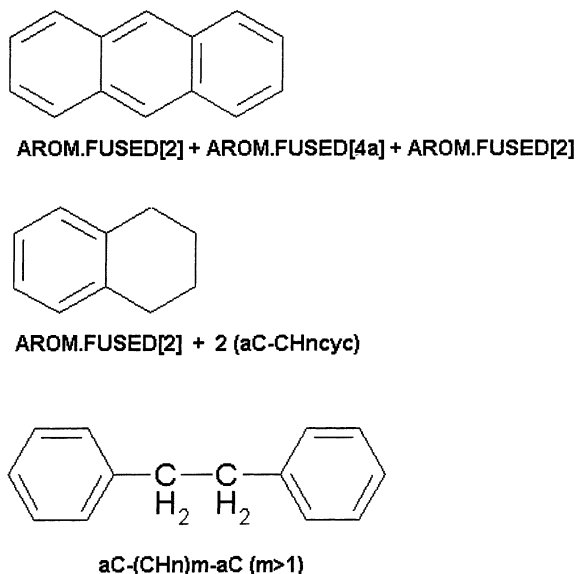e second-order group of type-j that occurs M_j times and E_k the contribution of the third-order group of type-k that has O_k occurrences in a compound. In the first level of estimation, the constants w and z are assigned zero values because only first-order groups are employed. In the second level, the constants w and z are assigned unity and zero values, respectively because only first- and second-order groups are involved while in the third level, both w and z are set to unity values. The left-hand side of Eq. (1) is a simple function $f(X)$ of the target property X. The selection of this function has been based on the following criteria.

1. The function has to achieve additivity in the contributions C_i, D_j and E_k.
2. It has to exhibit the best possible fit of the experimental data.
3. It should provide good extrapolating capability and therefore, a wide range of applicability.

According to these criteria, the selected functions are the same as used by Constantinou and Gani [10]. The target properties as well as their corresponding estimation functions are listed in Table 1. The symbols $T_{m1i}, T_{b1i}, T_{c1i}, P_{c1i}, V_{c1i}, G_{f1i}, H_{f1i}, H_{v1i}, H_{fus1i}$ represent the contributions (C_i) of the first-order groups for the corresponding properties. Similarly, $T_{m2j}, T_{b2j}, T_{c2j}, P_{c2j}, V_{c2j}, G_{f2j}, H_{f2j}, H_{v2j}, H_{fus2j}$ and $T_{m3k}, T_{b3k}, T_{c3k}, P_{c3k}, V_{c3k}, G_{f3k}, H_{f3k}, H_{fus3k}$ represent the contributions (D_j) and (E_k.) of the second- and third-order groups, respectively. The $T_{m0}, T_{b0}, T_{c0}, P_{c1}, P_{c2}, V_{c0}, G_{f0}, H_{f0}, H_{v0}, H_{fus0}$ are additional adjustable parameters of the estimation models or universal constants.

The determination of the adjustable parameters of the models, that is, the contributions C_i, D_j and E_k, has been divided into a three-step regression procedure.

1. Regression is carried out to determine the contributions (C_i) of the first-order groups and the universal constants of the models while w and z are set to zero.

Table 1
Selected function for each property

Property (X)	Left-hand side of Eq. (1) (function $f(X)$)	Right-hand side of Eq. (1) (group-contribution terms)
Normal melting point (T_m)	$\exp(T_m/T_{m0})$	$\sum_i N_i T_{m1i} + \sum_j M_j T_{m2j} + \sum_k O_k T_{m3k}$
Normal boiling point (T_b)	$\exp(T_b/T_{b0})$	$\sum_i N_i T_{b1i} + \sum_j M_j T_{b2j} + \sum_k O_k T_{b3k}$
Critical temperature (T_c)	$\exp(T_c/T_{c0})$	$\sum_i N_i T_{c1i} + \sum_j M_j T_{c2j} + \sum_k O_k T_{c3k}$
Critical pressure (P_c)	$(P_c - P_{c1})^{-0.5} - P_{c2}$	$\sum_i N_i P_{c1i} + \sum_j M_j P_{c2j} + \sum_k O_k P_{c3k}$
Critical volume (V_c)	$V_c - V_{c0}$	$\sum_i N_i V_{c1i} + \sum_j M_j V_{c2j} + \sum_k O_k V_{c3k}$
Standard Gibbs energy at 298 K (G_f)	$G_f - G_{f0}$	$\sum_i N_i G_{f1i} + \sum_j M_j G_{f2j} + \sum_k O_k G_{f3k}$
Standard enthalpy of formation at 298 K (H_f)	$H_f - H_{f0}$	$\sum_i N_i H_{f1i} + \sum_j M_j H_{f2j} + \sum_k O_k H_{f3k}$
Standard enthalpy of vaporization at 298 K (H_v)	$H_v - H_{v0}$	$\sum_i N_i H_{v1i} + \sum_j M_j H_{v2j}$
Standard enthalpy of fusion (H_{fus})	$H_{fus} - H_{fus0}$	$\sum_i N_i H_{fus1i} + \sum_j M_j H_{fus2j} + \sum_k O_k H_{fus3k}$

2. Then, w is set to unity, z is set to zero and another regression is performed using the C_i and the universal constants calculated in the previous step to determine the contributions (D_j) of the second-order groups.
3. Finally, both w and z are assigned to unity and, using the universal constants of the models (C_i and D_j obtained as results of the previous steps), the contributions (E_k.) of the third-order groups are determined.

   This stepped regression scheme ensures the independence among contributions of first-, second- and third-order. The contributions of the higher levels act as corrections to the approximations of the lower levels. The total list of groups and their contributions C_i, D_j and E_k for the nine listed properties are given in Appendix A (see Tables 6–8). The universal constants determined in the first step of the regression scheme are listed in Table 2. The optimization algorithm used for the data fitting was the Levenberg–Marquardt technique and the objective function was to minimize the sum of squares of the differences between experimental and estimated values of the target properties. The experimental data used in regression has been obtained from a comprehensive data bank of property values developed at CAPEC-DTU [12] through a systematic search of several data sources. Property values have been included in this collection after a rigorous analysis of their reliability.

Table 2
Values of the additional adjustable parameters

Adjustable parameter (universal constants)	Value
T_{m0}	147.450 K
T_{b0}	222.543 K
T_c	231.239 K
P_{c1}	5.9827 bar
P_{c2}	0.108998 bar$^{-0.5}$
V_{c0}	7.95 cm^3/mol
G_{f0}	-34.967 kJ/mol
H_{f0}	5.549 kJ/mol
H_{v0}	11.733 kJ/mol
H_{fus0}	-2.806 kJ/mol

3. Results and discussion

The key idea behind the new method is to consider that a first level of approximation is sufficient to estimate satisfactorily the property values of relatively simple compounds whereas, for properties of more complex compounds, additional levels are needed to obtain more accurate estimations. This multilevel approach is illustrated in Fig. 4, in which each region symbolizes a specific type of compounds. The first region contains simple and monofunctional compounds, which are represented by using only first-order groups. The second region includes aromatic and alicyclic compounds with only one ring as well as a considerable number of polyfunctional acyclic compounds, which are described through just first-order and second-order groups. Large, complex and heterocyclic compounds, which need third-order groups to be satisfactorily represented, are included in the third region. The percent values given in Fig. 4 for each region relates to the number of compounds used in this work for the corresponding region.

Table 3 presents the standard deviation for each property, the average absolute error and the average relative error for the first, second and third approximations. The number of experimental values used in the first regression step is also given. The statistics offered for the second and third approximations encompass all the data points, even those corresponding to compounds in which no second-order or third-order groups occur (and consequently not used in the second and third regression steps). Therefore, the average deviations given for the third approximation characterize the global results of the three subsequent approximations. Furthermore, due to the low number of available experimental values of enthalpies of vaporization at 298 K for complex and heterocyclic compounds, the contributions of third-order groups to this property have not been considered in this paper. A comparison of the average deviations obtained as a result of the second and third regression steps is shown in Table 4, which includes the actual number of data points used in each step, that is, the number of compounds in which second and third-order groups occur. For each set of compounds, the average deviations corresponding to both the current and previous step are presented in order to illustrate the improvement in accuracy achieved in each step.

The reliability of the estimation equations obtained from the regression steps has been tested for each property by performing a least-square analysis in which a randomly conformed subset of the *N*
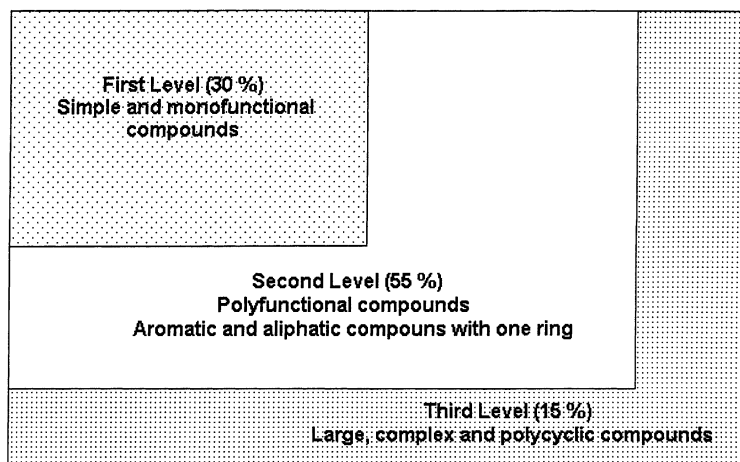

Fig. 4. A multilevel approach for property estimation from group-contributions.

Table 3
Global comparison of consecutive first, second and third approximations[a]

Property (X)	Data points	S.D.			AAE			ARE (%)		
		First	Second	Third	First	Second	Third	First	Second	Third
T_m (K)	1547	33.87	29.52	27.67	24.90	21.41	20.22	9.3	7.9	7.6
T_b (K)	1794	11.11	8.96	8.09	7.90	6.38	5.89	1.8	1.4	1.4
T_c (K)	783	17.25	8.50	6.99	8.75	5.67	4.93	1.4	0.9	0.8
P_c (bar)	775	1.73	1.53	1.39	1.02	0.87	0.79	2.9	2.6	2.3
V_c (cc/mol)	762	13.36	11.57	10.74	9.12	7.85	7.33	2.2	1.9	1.8
G_f (kJ/mol)	679	8.37	6.85	5.90	5.35	4.12	3.70	–	–	–
H_f (kJ/mol)	686	8.29	6.79	5.75	5.27	4.05	3.60	–	–	–
H_v (kJ/mol)	437	2.05	1.61	–	1.10	0.86	–	2.7	2.3	–
H_{fus} (kJ/mol)	711	4.16	3.88	3.65	2.58	2.32	2.17	18.3	16.4	15.7

[a] S.D. $= \sqrt{\sum(X_{est} - X_{exp})^2/N}$; AAE $= 1/N \sum |X_{est} - X_{exp}|$; ARE (%) $= 1/N \sum |(X_{est} - X_{exp})/X_{exp}| \times 100$ where N is the number of data points, X_{est} the estimated value of the property X, and X_{exp} the experimental value of the property X.

experimental data points has been excluded from the full data set. The mean-square residual J, defined as

$$J = \sqrt{\frac{\sum(X_i - Y_i)^2}{N}} \tag{2}$$

was calculated. In Eq. (2), N is the number of data points excluded from the full data set, X_i the property value of compound i estimated by the full regression, and Y_i the property value of the same compound estimated by the partial regression. For all the properties, the residuals are smaller than the estimation errors reported in Table 3, confirming the reliability of the method.

A fair comparison with other existing group-contribution methods is impossible since no other method exhibits the wide-ranging applicability of the proposed method. Moreover, the new method is able to deal with classes of compounds that cannot be handled by other widely used methods. The reason is that,

Table 4
Comparison of average deviations for second- and third-order approximations[a]

Property (X)	Data points (total)	Second			Third		
		Data points	Deviations		Data points	Deviations points	
			First	Second		First and second	Third
T_m	1547	960	9.7%	7.5%	181	10.0%	7.1%
T_b	1794	1107	1.9%	1.3%	141	2.6%	1.4%
T_c	783	412	1.6%	0.8%	52	1.8%	0.5%
P_c	775	411	2.6%	1.9%	64	5.0%	2.3%
V_c	762	408	2.5%	1.9%	62	2.7%	1.4%
G_f	679	358	5.8 kJ/mol	3.5 kJ/mol	57	9.5 kJ/mol	4.6 kJ/mol
H_f	686	353	5.6 kJ/mol	3.3 kJ/mol	58	9.3 kJ/mol	4.0 kJ/mol
H_v	437	218	2.5%	1.6%	–	–	–
H_{fus}	711	351	19.2%	15.5%	99	22.9%	17.7%

[a] Deviations are expressed as average relative errors for all properties except for G_f and H_f, which are expressed as average absolute errors.

Table 5
Comparison of accuracy between a classical group-contribution scheme and the proposed method[a]

Property (X)	Data points	S.D.		AAE		ARE (%)	
		JR	New	JR	New	JR	New
T_m (K)	1103	38.87	25.34	34.90	18.76	14.6	7.5
T_b (K)	1211	15.86	8.01	11.02	5.89	3.1	1.4
T_c (K)	587	18.73	6.87	10.96	4.87	2.1	0.9
P_c (bar)	573	3.71	1.36	2.45	0.74	5.6	2.2
V_c (cc/mol)	544	18.36	10.69	14.53	7.25	2.7	1.8
G_f (kJ/mol)	481	12.41	5.90	9.03	3.62	–	–
H_f (kJ/mol)	493	12.23	5.68	8.98	3.60	–	–
H_v (kJ/mol)	343	2.93	1.60	2.71	0.83	4.2	2.4
H_{fus} (kJ/mol)	499	6.84	3.62	3.06	2.11	46.3	15.6

[a] JR: Joback and Reid [1]; new: proposed method.

compared to other methods, a significantly larger data set has been used in the development of the new method as well as a larger and comprehensive set of groups. However, in order to make a quite acceptable comparison of the proposed method with another classical group-contribution method, we have calculated the contributions of the groups used by Joback and Reid [1] and recalculated our group-contributions using the estimation models reported in Table 1 and a common set of compounds that can be described by both group schemes. A comparison between the results obtained from both methods, the modified Joback's and the proposed one after the third approximation, is presented in Table 5. The new method clearly exhibits a much better accuracy. Application of the method is illustrated through several examples in Appendix B.

4. Conclusion

The application of three different sets of functional groups, one for a first-order approximation and two successive ones for refining the estimations for complex, large and heterocyclic compounds, has led to a new group-contribution method for the estimation of important physical and thermodynamic properties. Compared to other currently used estimation methods, the proposed method exhibits an improved accuracy and a considerably wider range of applicability to deal with chemically, biochemically and environmentally-related compounds. Even for lower molecular weight organic compounds, the larger set of first-order groups provides not only a wider range of application, but also an improved accuracy. In principle, the second-order and third-order groups have the same objective of providing additional structural information. The main difference between them is their size and the size of the compounds where they are used. Our results indicate that it is better to have a set of smaller second-order groups for the smaller, complex molecules and a set of larger third-order groups for the large, polyfunctional compounds. A computer program has also been developed for automatic selection of groups, given the molecular structure of the compound, and the interactive estimation of the listed properties.

Appendix A

See Tables 6–8.

Table 6
First-order groups and their contributions along with sample assignments

	Group	Example	T_{m1i}	T_{b1i}	T_{c1i}	P_{c1i}	V_{c1i}	G_{f1i}	H_{f1i}	H_{v1i}	H_{fus1i}
1	CH₃	*n*-Tetracontane (2)	0.6953	0.8491	1.7506	0.018615	68.35	2.878	−42.479	0.217	1.660
2	CH₂	*n*-Tetracontane (38)	0.2515	0.7141	1.3327	0.013547	56.28	8.064	−20.829	4.910	2.639
3	CH	2-Methylpentane (1)	−0.3730	0.2925	0.5960	0.007259	37.50	8.254	−7.122	7.962	0.134
4	C	2,2-Dimethylbutane (1)	0.0256	−0.0671	0.0306	0.001219	16.01	16.413	8.928	10.730	−1.232
5	CH₂=CH	1-Hexene (1)	1.1728	1.5596	3.2295	0.025745	111.43	95.738	57.509	4.031	1.268
6	CH=CH	2-Hexene (1)	0.9460	1.5597	3.0741	0.023003	98.43	92.656	69.664	9.456	4.441
7	CH₂=C	2-Methyl-1-butene (1)	0.7662	1.3621	2.7717	0.021137	91.40	85.107	61.625	8.602	2.451
8	CH=C	2-Methyl-2-butene (1)	0.1732	1.2971	2.5666	0.019609	83.89	88.691	81.835	14.095	3.032
9	C=C	2,3-Dimethyl-2-butene (1)	0.3928	1.2739	2.6391	0.014114	90.66	93.119	95.710	19.910	2.616
10	CH₂=C=CH	1,2-Butadiene (1)	1.7036	2.6840	5.4330	0.035483	143.57	229.906	198.840	11.310	7.076
11	CH₂=C=C	3-Methyl-1,2-butadiene (1)	1.5453	2.4014	4.8219	0.029678	146.36	226.710	208.490	*****	7.435
12	CH=C=CH	2,3-Pentadiene (1)	1.2850	2.5400	*****	*****	*****	*****	*****	*****	6.000
13	CH≡C	1-Pentyne (1)	2.2276	1.7618	3.7897	0.014010	84.60	230.029	224.902	6.144	−1.548
14	C≡C	3-Decyne (1)	2.0516	1.6767	4.5870	0.010888	74.66	216.013	228.282	12.540	6.128
15	aCH	Benzene (6)	0.5860	0.8365	2.0337	0.007260	42.39	26.732	12.861	3.683	1.948
16	aC fused with aromatic ring	Naphthalene (2)	1.8955	1.7324	5.4979	0.003564	35.71	20.379	20.187	6.631	0.845
17	aC fused with non-aromatic subring	Indane (2)	1.2065	1.1995	3.1058	0.006512	34.65	33.912	30.768	6.152	1.095
18	aC except as above	Benzophenone (1)	0.9176	1.5468	4.5344	0.012859	26.47	23.331	24.701	6.824	−0.531
19	aN in aromatic ring	Pyridine (1)	2.0438	1.3977	4.0954	−0.003339	36.47	89.902	70.862	9.420	2.555
20	aC–CH₃	Toluene (1)	1.0068	1.5653	3.4611	0.020907	97.33	24.919	−19.258	8.279	2.969
21	aC–CH₂	Ethylbenzene (1)	0.1065	1.4925	2.9003	0.018082	87.19	31.663	4.380	11.981	0.948
22	aC–CH	Cumene (1)	−0.5197	0.8665	1.9512	0.011795	73.51	30.393	18.440	13.519	−1.037
23	aC–C	*tert*-Butylbenzene (1)	−0.1041	0.5229	0.8576	0.011298	67.20	40.127	35.297	16.912	−2.856
24	aC–CH=CH₂	Styrene (1)	1.2832	2.4308	5.7861	0.030637	134.69	114.531	77.863	*****	4.013
25	aC–CH=CH	1-Propenylbenzene (1)	1.7744	2.9262	6.5062	0.026282	128.84	111.216	88.084	*****	8.274
26	aC–C=CH₂	α-Methylstyrene (1)	1.2612	2.1472	4.9967	0.026371	110.74	115.728	90.927	*****	3.324
27	aC–C≡CH	Phenylacetylene (1)	1.7495	2.3057	6.4572	0.019507	112.08	263.205	257.448	*****	2.514
28	aC–C≡C	1-Phenyl-1-propyne (1)	*****	2.7341	*****	*****	*****	*****	*****	*****	*****
29	OH	1,4-Butanediol (2)	2.7888	2.5670	5.2188	−0.005401	30.61	−144.051	−178.360	24.214	4.786
30	aC–OH	Phenol (1)	5.1473	3.3205	9.3472	−0.008788	50.77	−131.327	−164.191	34.099	8.427
31	COOH	1,5-Pentanedioic acid (2)	7.4042	5.1108	14.6038	0.009885	90.66	−337.090	−389.931	17.002	10.692
32	aC–COOH	Benzoic acid (1)	12.4296	6.0677	15.4515	0.017100	119.10	−312.422	−361.249	*****	14.649
33	CH₃CO	2-Butanone (1)	2.9588	3.1178	7.0058	0.025227	127.99	−120.667	−180.604	15.195	8.062
34	CH₂CO	3-Pentanone (1)	2.5232	2.6761	5.7157	0.019619	112.79	−120.425	−163.090	19.392	8.826
35	CHCO	2,4-Dimethyl-3-pentanone (1)	1.1565	2.1748	4.4743	0.012487	97.16	−116.799	−139.909	20.350	7.205
36	CCO	2,2,4,4-Tetramethyl-3-pentanone (1)	1.0638	1.7287	*****	*****	*****	*****	*****	*****	*****
37	aC–CO	Acetophenone (1)	2.9157	3.4650	9.4806	0.011007	90.69	−91.812	−106.965	25.036	4.852
38	CHO	1-Hexanal (1)	3.0186	2.5388	5.8013	0.010204	71.08	−100.882	−130.816	12.370	11.325
39	aC–CHO	Benzaldehyde (1)	2.4744	3.5172	9.4795	0.019633	122.91	−80.222	−107.159	*****	7.273
40	CH₃COO	Butyl acetate (1)	2.1657	3.1228	6.3179	0.033812	148.91	−306.733	−387.458	19.342	7.910
41	CH₂COO	Methyl Butyrate (1)	1.6329	2.9850	5.9619	0.026983	132.89	−298.332	−364.204	21.100	9.479
42	CHCOO	Ethyl isobutyrate (1)	1.0668	2.2869	4.7558	0.021990	125.52	−301.414	−352.057	24.937	9.317
43	CCOO	Ethyl 2,2-dimethylpropionate (1)	0.3983	1.6918	*****	*****	*****	*****	*****	23.739	*****
44	HCOO	Propyl formate (1)	2.0223	2.5972	5.6064	0.015249	93.29	−276.878	−327.678	15.422	8.115

45	aC–COO	Methyl benzoate (1)	1.3348	3.1952	6.7311	0.018948	105.53	−291.662	−307.727	25.206	8.149
46	aC–OOCH	Phenyl formate (1)	*****	0.4621	*****	*****	*****	*****	*****	*****	*****
47	aC–OOC	Phenyl acetate (1)	4.8044	3.0854	*****	*****	*****	*****	*****	*****	5.875
48	COO except as above	Ethyl acrylate (1)	1.5038	2.1903	4.7346	0.013087	81.17	−299.803	−331.397	*****	10.573
49	CH$_3$O	Methyl butyl ether (1)	1.3643	1.7703	3.4393	0.020084	88.20	−90.329	−156.062	5.783	5.089
50	CH$_2$O	Di-n-butyl ether (1)	0.8733	1.3368	2.4217	0.017954	74.03	−105.579	−152.239	9.997	4.891
51	CH–O	sec-Butyl ether (1)	0.2461	0.8924	0.7889	0.014487	60.06	−101.207	−147.709	14.620	4.766
52	C–O	tert-Butylether (1)	−0.4446	0.4983	0.2511	0.005613	52.96	−92.804	−121.608	13.850	2.458
53	aC–O	Methyl phenyl ether (1)	1.3045	1.8522	3.6588	0.005115	47.27	−83.354	−101.783	16.151	−0.118
54	CH$_2$NH$_2$	Ethylamine (1)	3.2742	2.7987	8.1745	0.011413	117.62	68.812	−10.703	15.432	13.482
55	CHNH$_2$	sec-Butylamine (1)	30.8394	2.0948	4.2847	0.013049	76.36	61.452	0.730	16.048	6.283
56	CNH$_2$	tert-Butylamine (1)	11.7400	1.6525	2.8546	0.010790	80.01	55.202	2.019	17.257	*****
57	CH$_3$NH	Dimethylamine (1)	2.4034	2.2514	4.5529	0.015863	77.04	88.512	24.740	11.831	4.490
58	CH$_2$NH	Dipropylamine (1)	1.7746	1.8750	3.2422	0.020482	95.15	88.874	23.610	13.067	7.711
59	CHNH	Diisopropylamine (1)	1.7577	1.2317	2.0057	0.005329	99.16	73.101	21.491	14.048	2.561
60	CH$_3$N	Methyldiethylamine (1)	0.9607	1.3841	3.0106	0.021186	94.94	125.906	55.024	9.493	6.008
61	CH$_2$N	Triethylamine (1)	0.0442	1.1222	2.1673	0.027454	74.05	121.247	65.331	12.636	1.756
62	aC–NH$_2$	Aniline (1)	3.9889	3.8298	10.2155	0.005335	81.40	66.470	17.501	23.335	6.542
63	aC–NH	N-methyl aniline (1)	1.4837	2.9230	8.4081	−0.005596	86.37	98.195	53.274	23.026	0.624
64	aC–N	N,N-dimethyl aniline (1)	1.7618	2.1918	5.8536	−0.000838	108.39	143.280	115.606	22.249	−2.576
65	NH$_2$ except as above	Cyclobutylamine	3.3478	2.0315	4.7420	0.000571	63.39	42.687	−8.556	13.425	6.158
66	CH=N	Acetaldazine (2)	8.8492	1.5332	*****	*****	*****	*****	*****	*****	*****
67	C=N	Ketazine (2)	1.4621	1.4291	*****	*****	*****	*****	*****	*****	*****
68	CH$_2$CN	Propionitrile (1)	2.5760	4.5871	12.9827	0.036523	133.62	134.997	99.245	21.923	7.303
69	CHCN	Isobutyronitrile (1)	2.1393	3.9774	8.4309	0.029034	134.73	142.475	151.390	24.963	9.464
70	CCN	2,2-Dimethylpropionitrile (1)	3.3807	2.8870	5.8829	0.024654	120.74	142.295	124.770	24.967	4.166
71	aC–CN	Benzonitrile (1)	5.1346	4.1424	10.4124	0.020978	119.08	162.175	148.968	*****	6.788
72	CN except as above	Acrylonitrile (1)	3.2747	3.0972	8.1381	0.024346	94.91	130.986	124.917	16.639	6.867
73	CH$_2$NCO	Ethyl isocyanate (1)	4.2256	3.4891	*****	*****	*****	*****	*****	*****	*****
74	CHNCO	Isopropyl isocyanate (1)	*****	3.1220	*****	*****	*****	*****	*****	*****	*****
75	CNCO	tert-Butyl isocyanate (1)	9.1492	*****	*****	*****	*****	*****	*****	*****	*****
76	aC–NCO	Phenyl isocyanate (1)	2.2327	3.1853	6.5884	0.025065	141.24	*****	*****	*****	*****
77	CH$_2$NO$_2$	1-Nitropropane (1)	3.2131	4.5311	10.9507	0.021056	157.57	25.783	−65.620	29.640	10.989
78	CHNO$_2$	2-Nitropropane (1)	0.7812	3.8069	9.5487	0.014899	143.36	16.407	−60.750	29.173	*****
79	CNO$_2$	2-Methyl-2-nitropropane (1)	5.6280	3.3059	*****	*****	*****	*****	*****	*****	−4.187
80	aC–NO$_2$	Nitrobenzene (1)	4.3531	4.5750	12.1243	0.018311	133.06	57.352	−22.931	24.863	7.572
81	NO$_2$ except as above	Nitrocyclohexane (1)	3.0376	3.2069	*****	*****	*****	*****	*****	*****	6.302
82	ONO	Butyl nitrite (1)	*****	1.8896	*****	*****	*****	*****	*****	*****	*****
83	ONO$_2$	n-Butyl nitrate (1)	2.5974	3.2656	*****	*****	*****	*****	*****	*****	9.353
84	HCON(CH$_2$)$_2$	Diethylformamide (1)	*****	5.8779	*****	*****	*****	*****	*****	*****	*****
85	HCONHCH$_2$	Ethylformamide (1)	*****	7.4566	*****	*****	*****	*****	*****	46.490	*****
86	CONH$_2$	Butyramide (1)	13.2124	6.5652	25.1184	0.001467	138.71	−127.512	−201.369	44.240	16.840
87	CONHCH$_3$	Methylacetamide (1)	5.4720	5.0724	20.5590	0.023455	190.71	−102.912	−203.069	*****	17.429
88	CONHCH$_2$	Ethylacetamide (1)	5.8825	6.6810	*****	*****	*****	*****	−183.613	52.723	*****
89	CON(CH$_3$)$_2$	Dimethylacetamide (1)	4.1720	6.0070	15.4603	0.043090	244.71	−56.412	−188.069	38.290	11.553
90	CONCH$_3$CH$_2$	Methylethylacetamide (1)	*****	*****	*****	*****	*****	*****	−48.210	*****	*****
91	CON(CH$_2$)$_2$	Diethylacetamide (1)	*****	5.0664	*****	*****	*****	*****	*****	*****	*****
92	CONHCO	Diacetamide (1)	9.1763	7.6172	*****	*****	*****	*****	*****	*****	*****
93	CONCO	Methyldiacetamide	3.2657	5.6487	*****	*****	*****	*****	*****	*****	*****

Table 6 (*Continued*)

	Group	Example	T_{m1i}	T_{b1i}	T_{c1i}	P_{c1i}	V_{c1i}	G_{f1i}	H_{f1i}	H_{v1i}	H_{fus1i}
94	aC–CONH$_2$	Benzamide	12.8071	8.3775	*****	*****	*****	*****	*****	*****	16.811
95	aC–NH(CO)H	*N*-phenylformamide (1)	5.6631	7.3497	19.8979	0.023447	162.08	−44.595	−125.052	*****	8.658
96	aC–N(CO)H	*N*-methyl-*N*-phenylmethanamide (1)	3.3602	5.1373	*****	*****	*****	*****	*****	*****	*****
97	aC–CONH	*N*-methylbenzamide (1)	6.5160	7.5850	*****	*****	*****	*****	*****	*****	10.959
98	aC–NHCO	*N*-(2-methylphenyl)acetamide (1)	9.8204	7.4955	*****	*****	*****	*****	*****	*****	4.370
99	aC–NCO	Phenylmethylacetamide (1)	7.2552	*****	*****	*****	*****	*****	*****	*****	*****
100	NHCONH	*N,N'*-dimethylurea (1)	9.3110	8.9406	*****	*****	*****	*****	*****	*****	9.862
101	NH$_2$CONH	Methylurea (1)	14.2020	16.3539	*****	*****	*****	*****	*****	*****	12.845
102	NH$_2$CON	*N,N*-dimethylurea (1)	13.0856	2.0796	*****	*****	*****	*****	*****	*****	10.958
103	NHCON	Trimethylurea (1)	8.4447	7.1529	*****	*****	*****	*****	*****	*****	12.098
104	NCON	Tetramethylurea (1)	3.5041	4.1459	*****	*****	*****	*****	*****	*****	9.557
105	aC–NHCONH$_2$	Phenylurea (1)	13.4695	5.7604	*****	*****	*****	*****	*****	*****	16.703
106	aC–NHCONH	*N,N'*-diphenylurea	23.2570	1.1633	*****	*****	*****	*****	*****	*****	18.460
107	NHCO except as above	*N*-chloroacetamide (1)	3.0882	*****	*****	*****	*****	*****	*****	*****	*****
108	CH$_2$Cl	1-Chlorobutane (1)	1.9253	2.6364	6.2561	0.021419	112.12	−19.484	−65.056	11.754	6.353
109	CHCl	2-Chloropropane (1)	1.0224	2.0246	4.3756	0.015640	100.78	−31.933	−65.127	12.048	*****
110	CCl	2-Chloro-2-methylpropane (1)	1.8424	1.7049	3.7063	0.009187	87.01	−37.848	−62.881	16.597	−0.082
111	CHCl$_2$	1,1-Dichloroethane (1)	2.5196	3.3420	7.8956	0.028236	159.79	−24.214	−80.812	17.251	6.781
112	CCl$_2$	2,2-Dichloropropane (1)	3.6491	2.9609	*****	*****	*****	*****	*****	20.473	1.823
113	CCl$_3$	1,1,1-Trichloroethane	4.4493	3.9093	8.8073	0.036746	204.71	−44.122	−105.369	20.550	3.492
114	CH$_2$F	1-Fluorobutane (1)	1.5597	1.5022	3.3179	0.023315	87.71	−180.212	−227.469	8.238	7.139
115	CHF	2-Fluorobutane (1)	1.1289	1.3738	2.6702	0.020040	78.08	−228.239	−261.901	*****	3.917
116	CF	2-Fluoro-2-methylpropane (1)	2.5398	1.0084	2.1633	−0.010120	*****	*****	*****	6.739	*****
117	CHF$_2$	1,1-Difluoroethane (1)	2.1689	2.2238	3.5702	0.031524	102.71	−411.239	−463.901	*****	7.011
118	CF$_2$	Perfluorohexane (4)	0.1312	0.5142	0.8543	0.018572	95.09	*****	*****	1.621	*****
119	CF$_3$	Hexafluoroethane (2)	1.4828	1.1916	1.7737	0.048565	108.85	−615.333	−673.875	7.352	2.526
120	CCl$_2$F	Tetrachloro-1,2-difluoroethane (2)	3.2035	2.5053	5.1653	0.037948	171.04	−249.020	−306.765	8.630	3.114
121	HCClF	1-Chloro-1,2,2,2-tetrafluoroethane (1)	*****	2.0542	*****	*****	*****	*****	*****	*****	*****
122	CClF$_2$	1,2-Dichlorotetrafluoroethane (2)	1.7510	1.7227	3.0593	0.041641	146.01	−396.814	−458.074	8.086	2.156
123	aC–Cl	Chlorobenzene (1)	1.7134	2.0669	5.7046	0.016033	92.67	1.985	−17.002	11.224	4.435
124	aC–F	Hexafluorobenzene (6)	0.9782	0.7945	1.5491	0.014037	54.36	−141.306	−160.965	3.965	2.003
125	aC–I	Iodobenzene (1)	2.1905	3.7739	12.4470	0.014403	131.08	91.505	95.048	*****	2.814
126	aC–Br	Bromobenzene (1)	2.4741	2.8414	8.4199	0.010199	104.12	42.977	38.917	14.393	5.734
127	I⁻ except as above	Iodoethane (1)	1.9444	3.1778	8.5775	−0.004637	104.28	43.910	47.632	14.171	6.103
128	Br⁻ except as above	Bromoethane (1)	1.7641	2.4231	4.5036	−0.001460	77.99	5.528	−1.703	9.888	4.826
129	F⁻ except as above	Benzyl fluoride (1)	1.2308	0.8504	0.8976	0.012034	24.62	−182.973	−201.968	*****	3.096
130	Cl⁻ except as above	Ethyl chloroacetate (1)	1.5454	1.5147	4.0947	0.007923	57.77	−29.876	−46.963	*****	5.181
131	CHNOH	Propionaldehyde oxime (1)	3.9813	4.5721	*****	*****	*****	*****	*****	*****	*****
132	CNOH	Diethyl ketoxime (1)	3.5484	4.0142	*****	*****	*****	*****	*****	*****	*****
133	aC–CHNOH	Phenyl oxime (1)	10.5579	*****	*****	*****	*****	*****	*****	*****	*****
134	OCH$_2$CH$_2$OH	2-Ethoxyethanol (1)	2.3651	4.8721	10.4579	0.025986	159.33	−233.335	−343.903	31.493	8.454
135	OCHCH$_2$OH	2-Ethoxy-1-propanol (1)	*****	4.2329	*****	*****	*****	*****	*****	*****	*****
136	OCH$_2$CHOH	1-Methoxy-2-propanol (1)	1.5791	3.6653	*****	0.018783	147.66	−239.423	−333.385	*****	12.594
137	–O–OH	*tert*-Butylhydroperoxide (1)	4.8181	3.1669	5.8307	−0.002815	58.01	−75.568	−125.111	*****	*****
138	CH$_2$SH	Ethanethiol (1)	2.2992	3.1974	7.7300	0.017299	105.68	27.469	−8.021	16.815	10.068
139	CHSH	2-Propanethiol (1)	0.9704	2.5910	5.8527	0.008968	109.36	27.030	3.510	17.098	4.266
140	CSH	2-Methyl-2-propanethiol (1)	4.2329	2.0902	4.6431	0.005118	94.01	27.338	12.589	18.397	−0.623

141	aC–SH	Benzenethiol (1)	2.8464	3.2675	9.5115	0.010086	95.08	48.905	41.648	17.413	4.513
142	–SH (except as above)	Cyclohexanethiol (1)	0.9600	2.3323	7.7987	0.006399	57.89	15.818	11.339	9.813	5.829
143	CH$_3$S	Dimethylsulfide (1)	1.7150	2.9892	6.9733	0.018013	122.03	35.845	−3.337	14.296	7.497
144	CH$_2$S	Diethylsulfide (1)	1.0063	2.6524	6.4871	0.015254	106.60	42.684	21.492	16.965	4.096
145	CHS	Diisopropylsulfide (1)	0.7892	2.0965	*****	*****	*****	*****	*****	19.038	*****
146	CS	di-*tert*-Butylsulfide (1)	1.1170	1.6412	*****	*****	*****	*****	*****	19.996	*****
147	aC–S–	Phenyl methyl sulfide (1)	0.9646	2.9731	*****	*****	*****	*****	*****	*****	*****
148	SO	Dimethyl sulfoxide (1)	5.3663	6.2796	19.8953	−0.005534	82.36	−52.231	−71.050	*****	13.403
149	SO$_2$	Dimethyl sulfone (1)	7.0778	7.0976	17.2586	−0.000784	89.95	−257.608	−305.498	*****	17.748
150	SO$_3$ (sulfite)	Dimethyl sulfite (1)	*****	3.9199	8.6910	0.004240	115.80	*****	−430.833	*****	*****
151	SO$_3$ (sulfonate)	Dimethyl sulfonate (1)	5.8426	6.7785	*****	*****	*****	*****	*****	*****	*****
152	SO$_4$ (sulfate)	Dimethyl sulfate (1)	3.6976	5.5627	18.9366	−0.027208	144.58	−519.853	−621.412	*****	*****
153	aC–SO	Phenyl methyl sulfoxide (1)	3.9911	6.1185	*****	*****	*****	*****	*****	*****	*****
154	aC–SO$_2$	Diphenyl sulfone (1)	5.2948	8.4333	*****	*****	135.47	−314.643	−370.493	*****	3.281
155	PH (phosphine)	Dimethylphosphine (1)	*****	2.0536	*****	*****	*****	*****	*****	*****	*****
156	P (phosphine)	Trimethylphosphine (1)	*****	1.0984	*****	*****	*****	*****	*****	*****	*****
157	PO$_3$ (phosphite)	Triethylphosphite (1)	1.0306	2.7900	*****	*****	*****	*****	*****	*****	*****
158	PHO$_3$ (phosphonate)	Dimethylphosphonate (1)	*****	5.6433	*****	*****	*****	*****	*****	*****	*****
159	PO$_3$ (phosphonate)	Trimethylphosphonate (1)	*****	4.5468	*****	*****	*****	*****	*****	*****	*****
160	PHO$_4$ (phosphate)	Diethylphosphate (1)	2.7461	5.1567	*****	*****	*****	*****	*****	*****	*****
161	PO$_4$ (phosphate)	Trimethylphosphate (1)	2.0330	3.7657	16.9914	−0.029036	85.59	*****	−1060.325	*****	*****
162	aC–PO$_4$	Triphenylphosphate (1)	−1.7840	2.3522	*****	*****	*****	*****	−1005.161	*****	4.256
163	aC–P	Triphenylphosphine (1)	0.2337	2.9272	38.6148	−0.126108	−142.79	*****	72.339	*****	−5.654
164	CO$_3$ (carbonate)	Diethylcarbonate (1)	3.6593	2.8847	6.6804	0.007235	93.56	−447.186	−516.282	21.613	8.363
165	C$_2$H$_3$O	Ethyl oxirane (1)	1.3135	2.8451	6.6418	0.021238	125.43	11.149	−52.241	*****	*****
166	C$_2$H$_2$O	2,2-Dimethyl oxirane (1)	*****	2.6124	6.0159	0.010678	194.36	1.890	−51.390	*****	*****
167	C$_2$O	Trimethyl oxirane (1)	*****	2.2036							
168	CH$_2$ (cyclic)	Cyclopentane (5)	0.5699	0.8234	1.8815	0.009884	49.24	13.287	−18.575	3.341	1.069
169	CH (cyclic)	Methylcyclopentane (1)	0.0335	0.5946	1.1020	0.007596	44.95	6.107	−12.464	6.416	2.511
170	C (cyclic)	1,1-Dimethylcyclohexane (1)	0.1695	0.0386	−0.2399	0.003268	33.32	−0.193	−2.098	7.017	−0.921
171	CH=CH (cyclic)	Cyclobutene (1)	1.1936	1.5985	3.6426	0.013815	83.91	86.493	59.841	7.767	1.185
172	CH=C (cyclic)	1-Methylcyclopentene (1)	0.4344	1.2529	3.5475	0.010576	70.98	67.056	64.295	7.171	2.559
173	C=C (cyclic)	1,2-Dimethylcyclopentene (1)	0.3048	1.1975	*****	*****	*****	*****	*****	*****	*****
174	CH$_2$=C (cyclic)	Methylene cyclohexane (1)	0.2220	1.5109	4.4913	0.019101	83.96	*****	*****	*****	5.351
175	NH (cyclic)	Cyclopentimine (1)	3.4814	2.1634	5.9726	−0.003678	51.80	72.540	23.138	13.700	8.655
176	N (cyclic)	*N*-methylpyrrolidine (1)	0.6040	1.6541	4.3905	−0.001179	31.41	83.779	65.622	*****	0.269
177	CH=N (cyclic)	Imidazole (1)	5.5779	6.5230	*****	*****	*****	*****	*****	*****	3.993
178	C=N (cyclic)	2-Methyl-1H-imidazole (1)	6.6382	6.6710	*****	*****	*****	*****	*****	*****	*****
179	O (cyclic)	Tetrahydropyran (1)	1.3828	1.0245	2.7409	−0.000387	17.69	−114.062	−137.353	6.877	3.806
180	CO (cyclic)	Cyclobutanone (1)	3.2119	2.8793	12.6396	−0.000207	57.38	−156.672	−180.166	17.124	6.137
181	S (cyclic)	2-Methyl-thiophene (1)	1.6023	2.3256	5.5523	0.001540	45.45	12.020	15.453	12.262	5.170
182	SO$_2$ (cyclic)	Cyclobutadiene sulfone (1)	6.1006	*****	24.3995	0.002487	96.66	−241.601	−283.839	*****	9.934

Table 7

Second-order groups and their contributions along with sample assignments

	Group	Example	T_{m2j}	T_{b2j}	T_{c2j}	P_{c2j}	V_{c2j}	G_{f2j}	H_{f2j}	H_{v2j}	H_{fus2j}
1	$(CH_3)_2CH$	2-Methylpentane (1)	0.1175	−0.0035	−0.0471	0.000473	1.71	−0.418	−0.419	−0.399	0.396
2	$(CH_3)_3C$	2,2,4,4-Tetramethylpentane (2)	−0.1214	0.0072	−0.1778	0.000340	3.14	−2.776	−1.967	−0.417	0.554
3	$CH(CH_3)CH(CH_3)$	2,3,4-Trimethylpentane (2)	0.2390	0.3160	0.5602	−0.003207	−3.75	6.996	6.065	0.532	−1.766
4	$CH(CH_3)C(CH_3)_2$	2,2,3,4,4-Pentamethylpentane (2)	−0.3276	0.3976	0.8994	−0.008733	−10.06	8.938	8.078	0.623	0.351
5	$C(CH_3)_2C(CH_3)_2$	2,2,3,3,4,4-Hexamethylpentane (2)	3.3297	0.4487	1.5535	−0.016852	−8.70	10.735	10.535	5.086	−1.089
6	$CH_n=CH_m–CH_p=CH_k$ (k, m, n, p in 0..2)	1,3-Butadiene (1)	0.7451	0.1097	0.4214	0.000792	−7.88	−6.562	−11.786	1.632	1.408
7	$CH_3–CH_m=CH_n$ (m, n in 0..2)	2-Methyl-2-butene (3)	0.0524	0.0369	−0.0172	−0.000101	0.50	−0.120	−0.048	0.064	0.070
8	$CH_2–CH_m=CH_n$ (m, n in 0..2)	1,4-Pentadiene (2)	−0.1077	−0.0537	0.0262	0.000815	0.14	1.006	1.449	−0.060	−0.632
9	$CH_p–CH_m=CH_n$ (m, n in 0..2; p in 0..1)	3-Methyl-1-butene (1)	−0.2485	−0.0093	−0.1526	−0.000163	−2.67	3.857	3.964	0.004	−0.368
10	CHCHO or CCHO	2-Methylbutyraldehyde (1)	0.5715	−0.1286	−1.0434	0.005789	10.36	−0.525	1.514	−0.550	−0.369
11	CH_3COCH_2	2-Pentanone (1)	−0.0968	−0.0215	−0.0338	−0.000111	−4.08	−1.543	0.033	−0.403	0.105
12	CH_3COCH or CH_3COC	3-Methyl-2-pentanone (1)	−0.6024	−0.0803	−0.3658	−0.001892	3.02	2.202	4.994	0.723	1.005
13	CHCOOH or CCOOH	2-Methyl butanoic acid (1)	−3.1734	−0.3203	−4.7275	0.006916	10.56	3.920	7.422	5.475	
14	CH_3COOCH or CH_3COOC	Isopropyl acetate (1)	0.2114	−0.2066	−0.5537	−0.000569	4.28	−11.779	−12.295	−1.871	1.208
15	CO–O–CO	Propanoic anhydride (1)	−1.2441	−0.0500	−0.3576	0.001812	2.98	−16.075	−14.140	*****	−2.666
16	CHOH	2-Butanol (1)	−0.3489	−0.2825	−0.6768	0.000246	−3.04	−5.614	−4.422	−0.206	−0.599
17	COH	2-Methyl-2-butanol (1)	0.3695	−0.5325	−1.5224	0.003224	13.98	−25.382	−25.929	−1.579	−0.459
18	CH_3COCH_nOH (n in 0..2)	3-Hydroxy-2-butanone (1)	0.9886	−0.2987	−0.3940	−0.002912	5.17	6.621	8.244	*****	*****
19	NCCHOH or NCCOH	2-Hydroxypropionitrile (1)	−1.1810	0.2981	0.3414	−0.000516	0.68	4.833	0.000	*****	−0.149
20	$OH–CH_n–COO$ (n in 0..2)	Ethyl lactate (1)	−0.1526	−0.2310	*****	*****	*****	*****	*****	*****	*****
21	$CH_m(OH)CH_n(OH)$ (m, n in 0..2)	Ethylene glycol (1)	−0.0414	0.8854	1.9395	−0.004712	7.54	−1.051	−0.592	−6.611	−0.306
22	$CH_m(OH)CH_n(–)$ (m, n, p in 0..2)	2-Amino-1-butanol (1)	−0.5941	0.5082	1.2342	0.002581	5.58	−1.506	−0.959	*****	−0.041
23	$CH_m(NH_2)CH_n(NH_2)$ (m, n in 0..2)	Ethylenediamine (1)	0.3258	−0.0064	−3.3555	0.000726	20.82	0.344	−1.443	2.384	−1.575
24	$CH_m(NH)CH_n(NH_2)$ (m, n in 1..2)	Diethylenetriamine (1)	−1.8403	0.2318	−1.1598	0.000157	−26.31	3.848	3.608	*****	*****
25	$H_2NCOCH_nCH_mCONH_2$ (m, n in 1..2)	Butanediamide (1)	11.5351	*****	*****	*****	*****	*****	*****	*****	*****
26	$CH_m(NH_n)–COOH$ (m, n in 0..2)	L-Alanine (1)	12.3481	*****	62.4740	−0.002696	17.78	3.145	6.598	*****	7.032
27	$HOOC–CH_n–COOH$ (n in 1..2)	Malonic acid (1)	0.9327	−0.1222	1.9595	−0.001479	12.46	−5.217	−6.058	*****	4.264
28	$HOOC–CH_n–CH_m–COOH$ (n, m in 1..2)	Succinic acid (1)	7.5057	*****	0.7686	0.000090	15.17	−4.281	−6.929	*****	29.245
29	$HO–CH_n–COOH$ (n in 1..2)	2-Hydroxyisobutyric acid (1)	−0.4531	−0.4625	*****	*****	*****	*****	*****	*****	*****
30	$NH_2–CH_n–CH_m–COOH$ (n, m in 1..2)	β-Alanine (1)	14.1593	*****	*****	*****	*****	*****	*****	*****	*****
31	$CH_3–O–CH_n–COOH$ (n in 1..2)	Methoxyacetic acid (1)	−2.3026	0.9198	0.4750	−0.001445	7.91	−2.678	−1.727	*****	*****
32	HS–CH–COOH	2-Mercaptopropionic acid (1)	−2.1535	*****	*****	*****	*****	*****	*****	*****	*****
33	$HS–CH_n–CH_m–COOH$ (n, m in 1..2)	β-Thiolactic acid (1)	−2.7514	*****	−0.2697	0.000655	20.43	−7.376	7.292	*****	−3.623
34	$NC–CH_n–CH_m–CN$ (n, m in 1..2)	1,2-Dicyanoethane (1)	4.0747	1.8957	1.9699	0.002330	24.82	18.974	5.661	*****	−8.038
35	$OH–CH_n–CH_m–CN$ (n, m in 1..2)	3-Hydroxypropanenitrile (1)	−0.9493	1.3434	0.2311	−0.001022	14.54	0.558	−3.906	*****	−4.371
36	$HS–CH_n–CH_m–SH$ (n, m in 1..2)	1,2-Ethanedithiol (1)	0.2232	0.1815	2.1272	0.001321	−10.31	6.728	0.794	−0.683	−0.931
37	$COO–CH_n–CH_m–OOC$ (n, m in 1..2)	Ethylene glycol diacetate (1)	−0.5946	0.3401	1.5418	−0.003385	−2.33	1.306	4.025	1.203	*****
38	$OOC–CH_m–CH_m–COO$ (n, m in 1..2)	Dimethylsuccinate (1)	2.5962	0.5794	*****	*****	*****	*****	*****	*****	2.303
39	$NC–CH_n–COO$ (n in 1..2)	Methylcyanoacetate (1)	−0.2509	1.2171	2.7051	−0.001999	−0.73	*****	*****	*****	1.100
40	$COCH_nCOO$ (n in 1..2)	Methylacetoacetate (1)	0.6304	0.2427	0.7502	−0.000231	1.69	10.556	−7.261	*****	*****
41	$CH_m–O–CH_n=CH_p$ (m, n, p in 0..3)	Ethyl vinyl ether (1)	−0.0811	0.1399	0.2900	−0.000432	−4.54	−10.098	−9.411	0.372	3.169
42	$CH_m=CH_n–F$ (m, n in 0..2)	1-Fluoro-1-propene (1)	−0.2568	0.0591	*****	*****	*****	*****	*****	*****	2.823
43	$CH_m=CH_n–Br$ (m, n in 0..2)	1-Bromo-1-propene (1)	−0.4329	−0.3192	*****	−0.010021	2.63	14.470	17.014	*****	2.212
44	$CH_m=CH_n–I$ (m, n in 0..2)	1-Iodo-1-propene (1)	*****	−0.3486	*****	*****	*****	*****	*****	*****	*****
45	$CH_m=CH_n–Cl$ (m, n in 0..2)	1-Chloro-2-methylpropene (1)	0.0446	−0.0268	−0.0188	0.000152	2.80	8.207	9.715	*****	−0.480
46	$CH_m=CH_n–CN$ (m, n in 0..2)	Acrylonitrile (1)	0.1027	0.0653	−1.1249	0.000893	3.82	−8.304	−16.903	*****	−0.405

47	CH_n=CH_m–COO–CH_p (*m, n, p* in 0..3)	Ethyl Acrylate (1)	0.2117	−0.0430	−0.0880	0.000044	0.21	−12.085	−12.509	*****	−0.014
48	CH_m=CH_n–CHO (*m, n* in 0..2)	Propenaldehyde (1)	−0.7191	0.1102	*****	*****	*****	*****	*****	*****	*****
49	CH_m=CH_n–COOH (*m, n* in 0..2)	Acrylic Acid (1)	2.4103	0.0667	−1.7762	−0.000763	4.36	10.194	9.090	*****	1.291
50	aC–CH_n–X (*n* in 1..2) X: Halogen	Benzyl bromide (1)	0.8092	0.4537	2.2630	0.002464	−4.88	−8.081	−8.570	*****	*****
51	aC–CH_n–NH_m (*n* in 1..2; *m* in 0..2)	Benzyl amine (1)	−1.0802	0.2590	1.4069	−0.000034	2.50	−2.044	−3.447	4.608	−0.639
52	aC–CH_n–O– (*n* in 1..2)	Benzyl ethyl ether (1)	0.8607	−0.0425	0.2698	−0.000417	−7.49	6.043	5.486	*****	0.969
53	aC–CH_n–OH (*n* in 1..2)	Benzyl alcohol (1)	0.8981	0.1005	−1.0107	0.002944	−0.25	*****	*****	*****	−2.754
54	aC–CH_n–CN (*n* in 1..2)	Benzyl cyanide (1)	0.1088	1.0587	2.4950	−0.000796	−11.01	25.157	16.950	*****	*****
55	aC–CH_n–CHO (*n* in 1..2)	Phenyl acetaldehyde (1)	1.9470	−0.0177	*****	*****	*****	*****	*****	*****	*****
56	aC–CH_n–SH (*n* in 1..2)	Phenyl methanethiol (1)	1.2057	0.1702	0.8705	0.000183	2.00	16.725	7.568	*****	0.890
57	aC–CH_n–COOH (*n* in 1..2)	Phenyl acetic acid (1)	0.3666	0.1584	*****	*****	*****	*****	*****	*****	−4.086
58	aC–CH_n–CO– (*n* in 1..2)	Phenyl acetone (1)	−0.2363	0.3094	*****	*****	*****	*****	*****	*****	*****
59	aC–CH_n–S– (*n* in 1..2)	Benzyl methyl sulfide (1)	0.4506	0.1030	*****	*****	*****	*****	*****	*****	*****
60	aC–CH_n–OOC–H (*n* in 1..2)	Benzyl formate (1)	*****	0.2238	1.7860	0.004195	−3.40	3.020	4.145	*****	*****
61	aC–CH_m–NO_2 (*n* in 1..2)	Phenyl nitromethane (1)	*****	0.5390	*****	*****	*****	*****	*****	*****	*****
62	aC–CH_n–$CONH_2$ (*n* in 1..2)	Phenyl ethanamide (1)	2.2421	−0.2197	*****	*****	*****	*****	*****	*****	*****
63	aC–CH_n–OOC (*n* in 1..2)	Benzyl acetate (1)	−0.6997	0.0886	1.1629	−0.000384	−7.02	1.556	4.066	*****	*****
64	aC–CH_n–COO (*n* in 1..2)	Methyl phenyl acetate (1)	−0.2636	0.0352	*****	*****	*****	*****	*****	*****	*****
65	aC–SO_2–OH	Benzenesulfonic acid (1)	−1.1057	*****	*****	*****	*****	*****	*****	*****	*****
66	aC–$CH(CH_3)_2$	Cumene (1)	0.0642	0.0196	0.1565	−0.001446	−2.04	1.238	−0.751	1.030	−0.270
67	aC–$C(CH_3)_3$	*tert*-Butylbenzene (1)	0.0790	0.0494	0.8016	−0.006495	−5.70	0.354	−0.192	*****	−0.878
68	aC–CF_3	Perfluorotoluene (1)	−10.8058	−1.5974	*****	*****	*****	*****	*****	*****	*****
69	(CH_n=C)(cyclic)–CHO (*n* in 0..2)	Furfural (1)	−1.0516	0.4267	2.4070	−0.002650	0.39	−6.438	−12.517	*****	−1.670
70	(CH_n=C)$_{cyc}$–COO–CH_m (*n, m* in 0..3)	Methyl furanyrate (1)	−6.9427	0.0879	*****	*****	*****	*****	*****	*****	*****
71	(CH_n=C)$_{cyc}$–CO– (*n* in 0..2)	2-Acetylfuran (1)	0.6572	0.6115	*****	*****	*****	*****	*****	*****	*****
72	(CH_n=C)$_{cyc}$–CH_3 (*n* in 0..2)	1,2-Dimethylcyclopentene (2)	0.0416	0.0173	−0.2509	−0.000624	0.03	28.972	24.560	*****	2.235
73	(CH_n=C)$_{cyc}$–CH_2 (*n* in 0..2)	2-Ethylfuran (1)	−0.3151	−0.0504	−1.1019	0.003921	−4.43	−22.533	−12.044	*****	0.961
74	(CH_n=C)$_{cyc}$–CN (*n* in 0..2)	3-Cyanofuran (1)	1.5819	−0.2474	*****	*****	*****	*****	*****	*****	*****
75	(CH_n=C)$_{cyc}$–Cl (*n* in 0..2)	2-Chlorofuran (1)	−0.8604	−0.5736	*****	*****	*****	*****	*****	*****	*****
76	CH_{cyc}–CH_3	Methylcyclopentane (1)	−0.1326	−0.1210	−0.1233	0.000779	2.79	4.178	4.452	0.096	0.033
77	CH_{cyc}–CH_2	Ethylcyclohexane (1)	−0.4669	−0.0148	0.3816	0.001694	−2.95	5.332	4.428	−0.428	−1.137
78	Ch_{cyc}–CH	Isopropylcyclopentane (1)	−0.3548	0.1395	0.1093	0.000124	6.19	6.084	−4.128	0.153	2.421
79	Ch_{cyc}–C	*tert*-Butylcyclohexane (1)	−0.1727	0.1829	*****	*****	*****	*****	*****	*****	*****
80	Ch_{cyc}–CH=CH_n (*n* in 1..2)	Vinylcyclopentane (1)	0.6817	−0.1192	*****	*****	*****	*****	*****	*****	*****
81	Ch_{cyc}–C=CH_n (*n* in 1..2)	Limonene (1)	−1.0631	−0.0455	−0.2832	0.002114	−16.97	6.768	10.390	*****	*****
82	Ch_{cyc}–Cl	Chloro cyclopentane (1)	0.5124	0.2667	*****	*****	*****	*****	*****	*****	*****
83	Ch_{cyc}–F	Fluoro cyclohexane (1)	2.8497	−0.1899	*****	*****	*****	*****	*****	*****	*****
84	Ch_{cyc}−OH	Cyclohexanol (1)	1.3691	−0.3179	0.8973	0.004640	−7.73	−3.024	−8.050	2.134	*****
85	Ch_{cyc}–NH_2	Cyclohexylamine (1)	1.5069	−0.3576	−0.9610	0.000039	−2.50	2.046	3.446	−4.607	0.328
86	Ch_{cyc}–NH–CH_n (*n* in 0..3)	*N*-methylcyclohexylamine (1)	0.0370	−0.7458	−2.0833	−0.014535	−51.50	−11.965	14.531	*****	0.402
87	Ch_{cyc}−N–CH_n (*n* in 0..3)	*N,N*-dimethylcyclohexanamine (1)	*****	0.1218	*****	*****	*****	*****	*****	*****	*****
88	Ch_{cyc}–SH	Cyclohexanethiol (1)	−0.3312	−0.0569	−0.6447	−0.000199	−2.00	−16.723	−7.569	*****	−0.878
89	Ch_{cyc}–CN	Cyanocyclopentane (1)	*****	0.4649	*****	*****	*****	*****	*****	*****	*****
90	Ch_{cyc}–COOH	Cyclopropanecarboxylic acid (1)	−2.0822	0.1506	*****	*****	*****	*****	*****	*****	*****
91	Ch_{cyc}–CO	Methyl cyclohexyl ketone (1)	0.7743	0.1300	*****	*****	*****	*****	*****	−0.616	*****
92	Ch_{cyc}−NO_2	Nitrocyclohexane (1)	−0.8578	0.6540	*****	*****	*****	*****	*****	*****	*****
93	Ch_{cyc}–S–	Methyl cyclopentyl sulfide (1)	−0.8638	0.0043	*****	*****	*****	*****	*****	*****	*****
94	Ch_{cyc}–CHO	Cyclohexanecarboxaldehyde (1)	0.5076	−0.2692	*****	*****	*****	*****	*****	*****	*****
95	Ch_{cyc}–O–	Methoxycyclohexane (1)	−0.3978	−0.2787	*****	*****	*****	*****	*****	*****	*****

Table 7 (*Continued*)

	Group	Example	T_{m2j}	T_{b2j}	T_{c2j}	P_{c2j}	V_{c2j}	G_{f2j}	H_{f2j}	H_{v2j}	H_{fus2j}
96	Ch_{cyc}–OOCH	Cyclohexyl ester formic acid (1)	*****	−0.2107	*****	*****	*****	*****	*****	*****	*****
97	Ch_{cyc}–COO	Ethyl cyclobutyrate (1)	*****	0.0926	*****	*****	*****	*****	*****	*****	*****
98	Ch_{cyc}–OOC	Cyclohexyl acetate (1)	−0.4666	−0.4495	−0.3450	−0.000692	−12.03	4.358	−15.751	*****	*****
99	C_{cyc}–CH_3	1,1-Dimethyl-cyclohexane (2)	0.1737	0.0722	0.1607	0.001235	1.95	0.107	0.238	0.808	−1.237
100	C_{cyc}–CH_2	1-Ethyl-1-methyl-cyclopentane (1)	−1.9233	0.0319	0.1090	−0.000610	−5.17	18.755	21.498	0.585	*****
101	C_{cyc}–OH	1-Methylcyclopentanol (1)	0.7334	−0.6775	−2.1303	−0.004683	−14.40	−18.970	−21.975	*****	0.235
102	>N_{cyc}–CH_3	N-methyl-2-pyrrolidone (1)	−0.0383	0.0604	−0.0003	0.000058	*****	*****	*****	*****	*****
103	>N_{cyc}–CH_2	N-ethylpyrrole (1)	1.0497	−0.3080	*****	*****	*****	*****	*****	*****	*****
104	AROMRINGs^{1}s^2	2-Methyl-phenol (1), 2-Et-toluene (1)	−0.6388	−0.1590	−0.3161	0.000522	2.86	1.577	1.486	1.164	−1.470
105	AROMRINGs^{1}s^3	3-Methyl-phenol (1), 3-Et-toluene (1)	−0.6218	0.0217	−0.0693	0.001790	6.54	−1.037	0.294	−1.910	−1.059
106	AROMRINGs^{1}s^4	4-Methyl-phenol (1), 4-Et-toluene (1)	0.9840	0.1007	0.0803	0.000467	3.70	−0.709	0.384	0.331	1.244
107	AROMRINGs^{1}s^{2}s^3	1,2,3-Trimethylbenzene (1)	−0.2762	−0.1647	1.0088	−0.005598	−9.58	7.731	5.743	1.433	0.473
108	AROMRINGs^{1}s^{2}s^4	1,2,4-Trihydroxybenzene (1)	−0.3689	−0.1387	0.0908	0.000255	−2.05	−2.767	−0.449	0.313	−0.302
109	AROMRINGs^{1}s^{3}s^5	3,5-Diethyltoluene (1)	−0.3841	−0.1314	−0.6412	0.004090	−7.67	−2.148	−7.538	−0.117	−2.530
110	AROMRINGs^{1}s^{2}s^{3}s^4	3-Ethyl-1,2,4-trimethylbenzene (1)	1.7722	0.2745	2.1116	−0.007612	−7.04	14.226	12.710	*****	−1.736
111	AROMRINGs^{1}s^{2}s^{3}s^5	1,2,3,5-Tetramethylbenzene (1)	0.4553	0.1645	0.9353	−0.001811	−0.04	4.926	5.220	*****	−2.246
112	AROMRINGs^{1}s^{2}s^{4}s^5	1,2,4,5-Tetramethylbenzene (1)	2.0561	0.0754	0.6241	−0.000500	−0.04	−0.474	−1.340	*****	8.034
113	PYRIDINEs^2	2-Methylpyridine (1)	−0.5769	−0.1196	−1.0256	0.007006	8.68	−9.713	−9.644	−1.683	−0.786
114	PYRIDINEs^3	3-Methylpyridine (1)	−0.2556	0.0494	0.5784	0.007006	8.68	−2.523	−2.446	0.277	3.671
115	PYRIDINEs^4	4-Methylpyridine (1)	1.6282	0.1344	0.6595	0.001283	14.28	−4.703	−6.466	0.397	5.975
116	PYRIDINEs^{2}s^3	2,3-Dimethylpyridine (1)	−0.1341	0.0032	*****	*****	*****	*****	*****	−0.939	*****
117	PYRIDINEs^{2}s^4	2,4-Dimethylpyridine (1)	−1.6848	−0.0817	*****	*****	*****	*****	*****	−1.269	*****
118	PYRIDINEs^{2}s^5	2,5-Dimethylpyridine (1)	−0.9802	−0.1564	*****	*****	*****	*****	*****	−1.719	*****
119	PYRIDINEs^{2}s^6	2,6-Dimethylpyridine (1)	0.3018	−0.5176	−2.2773	0.008029	−50.26	−16.570	−17.778	−3.419	−1.487
120	PYRIDINEs^{3}s^4	3,4-Dimethylpyridine (1)	0.1018	0.5477	*****	*****	*****	*****	*****	1.742	*****
121	PYRIDINEs^{3}s^5	3,5-Dimethylpyridine (1)	0.2811	0.3533	*****	*****	*****	*****	*****	0.572	*****
122	PYRIDINEs^{2}s^{3}s^6	2,3,6-Trimethylpyridine (1)	−0.3189	−0.3888	*****	*****	*****	*****	*****	−2.744	*****

Table 8
Third-order groups and their contributions along with sample assignments

	Group	Example	T_{m3k}	T_{b3k}	T_{c3k}	P_{c3k}	V_{c3k}	G_{f3k}	H_{f3k}	H_{fus3k}
1	HOOC–(CH$_n$)$_m$–COOH ($m > 2$, n in 0..2)	1,5-Pentanedioic acid (1)	−1.5257	1.6498	−1.6986	0.001544	−3.72	−4.708	−6.572	−7.583
2	NH$_n$–(CH$_n$)$_m$–COOH ($m > 2$, n in 0..2)	4-Aminobutyric acid (1)	11.2271	*****	*****	*****	*****	*****	*****	*****
3	NH$_2$–(CH$_n$)$_m$–OH ($m > 2$, n in 0..2)	4-Aminobutanol (1)	0.7732	1.0750	0.4950	0.000728	−23.74	3.079	4.171	−4.840
4	OH–(CH$_n$)$_m$–OH ($m > 2$, n in 0..2)	1,9-Nonanediol (1)	0.6674	0.7193	0.1725	−0.000327	−0.84	7.536	5.411	−0.272
5	OH–(CH$_p$)$_k$–O–(CH$_n$)$_m$–OH (m, $k > 0$; p, n in 0..2)	Dipropylene glycol (1)	−0.1073	1.1867	6.6872	0.001937	1.44	−8.397	−8.651	1.661
6	OH–(CH$_p$)$_k$–S–(CH$_n$)$_m$–OH (m, $k > 0$; p, n in 0..2)	2,2′-Diethyl-dihydroxy sulfide (1)	−1.3891	*****	2.6769	0.003792	−1.62	10.194	8.164	−3.479
7	OH–(CH$_p$)$_k$–NHx–(CH$_n$)$_m$–OH (m, $k > 0$; p, n, x in 0..2)	Diethanolamine (1)	−0.0781	0.2991	*****	0.003254	−0.69	1.662	1.753	0.301
8	CH$_p$–O–(CH$_n$)$_m$–OH ($m > 2$; n, p in 0..2)	Butoxypropanol (1)	*****	−0.4605	*****	*****	*****	*****	*****	*****
9	NH$_2$–(CH$_n$)$_m$–NH$_2$ ($m > 2$; n in 0..2)	1,5-Diaminopentane (1)	−0.0604	0.0060	−4.3195	0.006734	6.69	4.100	0.371	5.666
10	NH$_k$–(CH$_n$)$_m$–NH$_2$ ($m > 2$; k in 0..1; n in 0..2)	N,N-dimethylpropylenediamine (1)	−1.1888	−0.1819	*****	*****	*****	*****	*****	*****
11	SH–(CH$_n$)$_m$–SH ($m > 2$; n in 0..2)	1,5-Pentanedithiol (1)	0.6669	0.4516	*****	*****	*****	*****	*****	*****
12	NC–(CH$_n$)$_m$–CN ($m > 2$)	Glutaronitrile (1)	−0.3798	1.3440	0.0834	−0.011090	−36.89	−7.035	7.782	−0.607
13	COO–(CH$_n$)$_m$–OOC ($m > 2$; n in 0..2)	Glyceryl tridodecanoate (1)	−2.6542	*****	*****	*****	*****	*****	*****	*****
14	aC–(CH$_n$=CH$_m$)$_{cyc}$ (fused rings) (n, m in 0..1)	Indene (1), Acenaphtylene (2)	0.2479	−0.3741	−0.0185	0.000851	−8.87	−1.601	2.689	−2.703
15	aC–aC (different rings)	Biphenylene (2), Biphenyl (1)	1.1395	−0.4961	6.1894	−0.040100	−26.26	−4.459	−4.558	−0.385
16	aC–CH$_{n,cyc}$ (different rings) (n in 0..1)	Cyclohexylbenzene (1)	0.0570	−0.4574	−0.2474	−0.005826	−8.55	−5.267	−5.914	−0.442
17	aC–CH$_{n,cyc}$ (fused rings) (n in 0..1)	Tetralin (2), Indane (2)	−0.5640	−0.1736	0.5060	−0.003746	−11.56	−4.203	−4.863	−0.143
18	aC–(CH$_n$)$_m$–aC (different rings) ($m > 1$; n in 0..2)	Bibenzyl (1)	1.9902	0.3138	3.0321	0.003007	9.73	1.318	0.084	5.377
19	aC–(CH$_n$)$_m$–CH$_{cyc}$ (different rings) ($m > 0$; n in 0..2)	1-Cyclopentyl-3-phenylpropane (1)	*****	0.5928	*****	*****	*****	*****	*****	*****
20	CH$_{cyc}$–CH$_{cyc}$ (different rings)	Cyclohexylcyclohexane (1)	0.5460	0.4387	2.1761	0.002745	7.72	−67.517	−66.870	*****
21	CH$_{cyc}$–(CH$_n$)$_m$–CH$_{cyc}$ (different rings) ($m > 0$; n in 0..2)	1,2-Dicyclohexylethane (1)	0.4497	0.5632	*****	*****	*****	*****	*****	*****
22	CH multiring	Hexahydroindan (2), Decalin (2)	0.6647	0.1415	0.4963	−0.000985	−3.33	*****	*****	0.223
23	C multiring	Spiropentane (1)	0.0792	*****	*****	*****	*****	*****	*****	*****
24	aC–CH$_m$–aC (different rings) (m in 0..2)	Diphenylmethane (1)	0.6457	0.2391	0.1174	−0.002673	−4.67	−0.729	0.866	−0.958
25	aC–(CH$_m$=CH$_n$)–aC (different rings) (m, n in 0..2)	1,2-Diphenylethylene (1)	0.9608	0.7192	0.7039	−0.004661	14.31	−0.702	−2.291	3.275
26	(CH$_m$=C)$_{cyc}$–CH=CH–(C=CH$_n$)$_{cyc}$ (different rings)	1,2-Furanyl ethene (1)	16.2235	*****	*****	*****	*****	*****	*****	*****
27	(CH$_m$=C)$_{cyc}$–CH$_p$–(C=CH$_n$)$_{cyc}$ (different rings)	Difuranyl methane (1)	16.8558	*****	*****	*****	*****	*****	*****	*****
28	aC–CO–aC (different rings)	Benzophenone (1)	−1.0394	1.0171	−0.2678	−0.001837	−7.05	11.125	7.108	−4.091
29	aC–CH$_m$–CO–aC (different rings) (m in 0..2)	Benzyl phenone (1)	−0.4486	0.9674	*****	*****	*****	*****	*****	*****
30	aC–CO–(C=CH$_n$)$_{cyc}$ (different rings) (n in 0..1)	Phenyl-2-furanyl-methanone (1)	−0.1376	0.1126	*****	*****	*****	*****	*****	*****
31	aC–CO–CO–aC (different rings)	Diphenylethanedione (1)	0.4361	0.9317	*****	*****	*****	*****	*****	−3.687
32	aC–CO$_{cyc}$ (fused rings)	Phenolphthalein (1)	3.6847	0.5031	*****	*****	*****	*****	*****	2.047
33	aC–CO–(CH$_n$)$_m$–CO–aC (different rings) ($m > 0$; n in 0..2)	1,4-Diphenyl-1,4-butanedione (1)	4.9038	*****	*****	*****	*****	*****	*****	7.327
34	aC–CO–CH$_{n,cyc}$ (different rings) (n in 0..1)	Cyclohexyl phenyl methanone (1)	−7.0038	*****	*****	*****	*****	*****	*****	*****
35	aC–CO–NH$_n$–aC (different rings) (n in 0..1)	N-phenyl benzamide (1)	5.9653	*****	*****	*****	*****	*****	*****	2.510
36	aC–NH$_n$CONH$_m$–aC (different rings) (n, m in 0..1)	N,N'-diphenylurea (1)	1.5629	*****	*****	*****	*****	*****	*****	0.018
37	aC–CO–N$_{cyc}$ (different rings)	N-phenonyl-piperidine (1)	−9.1856	*****	*****	*****	*****	*****	*****	*****
38	aC–S$_{cyc}$ (fused rings)	Dibenzothiophene (2)	0.2612	0.2242	3.5541	0.004600	12.60	8.333	9.212	−0.784
39	aC–S–aC (different rings)	Diphenyl sulfide (1)	−1.8403	0.0185	*****	*****	*****	*****	*****	*****
40	aC–PO$_n$–aC (different rings) (n in 0..4)	Triphenylphosphate (3)	0.0393	*****	*****	*****	*****	*****	*****	*****
41	aC–SO$_n$–aC (different rings) (n in 1..4)	Diphenyl sulfone (1)	0.9514	−0.0850	*****	*****	*****	*****	*****	−2.485
42	aC–NH$_{n,cyc}$ (fused rings) (n in 0..1)	Carbazole (2)	3.4983	1.1457	3.5541	0.017201	0.44	−2.221	−16.080	0.196
43	aC–NH–aC (different rings)	Diphenylamine (1)	−0.3048	0.5768	0.9519	0.008484	1.42	−0.596	−1.994	1.934
44	aC–(C=N)$_{cyc}$ (different rings)	Phenyl-3-pyrazole (1)	−1.3060	−0.5335	*****	*****	*****	*****	*****	*****
45	aC–(N=CH$_n$)$_{cyc}$ (fused rings) (n in 0..1)	Benzoxazole (1)	−4.9289	−5.2736	*****	*****	*****	*****	*****	−0.599
46	aC–(CH$_n$=N)$_{cyc}$ (fused rings) (n in 0..1)	Benzoisoxazole (1)	−10.1007	*****	*****	*****	*****	*****	*****	*****

Table 8 (*Continued*)

	Group	Example	T_{m3k}	T_{b3k}	T_{c3k}	P_{c3k}	V_{c3k}	G_{f3k}	H_{f3k}	H_{fus3k}
47	aC–O–CH$_n$–aC (different rings) (n in 0..2)	Benzyl phenyl ether (1)	1.0834	0.6571	*****	*****	*****	*****	*****	*****
48	aC–O–aC (different rings)	Diphenyl ether (1)	−0.4803	−0.8252	−0.9785	0.001162	−2.63	2.668	−5.074	1.193
49	aC–CH$_n$–O–CH$_m$–aC (different rings) (n, m in 0..2)	Benzyl ether (1)	−3.2676	0.2790	−1.4002	−0.004716	28.42	−4.229	−2.303	−3.971
50	aC–O$_{cyc}$ (fused rings)	Benzoxazole (1)	−0.3545	−0.6848	*****	*****	*****	*****	*****	−1.153
51	AROMFUSED[2]	Naphthalene (2)	0.2825	0.0441	−1.0095	−0.001332	−6.88	1.993	1.904	0.694
52	AROMFUSED[2]s^1	1-Methylnaphtalene (1)	−1.2836	−0.1666	0.1605	−0.002030	−3.17	−2.940	−2.274	−3.699
53	AROMFUSED[2]s^2	2,7-Dimethylnaphtalene (2)	0.3378	−0.2692	−0.6765	−0.002436	−3.85	−1.873	−1.316	2.037
54	AROMFUSED[2]s^{2}s^3	2,3-Dimethylnaphtalene (1)	1.8941	−0.2807	*****	*****	*****	*****	*****	2.150
55	AROMFUSED[2]s^{1}s^4	1,4-Dimethylnaphtalene (1)	−2.7585	−0.3294	*****	*****	*****	*****	*****	*****
56	AROMFUSED[2]s^{1}s^2	1,2-Dimethylnaphtalene (1)	−3.0362	−0.2931	*****	*****	*****	*****	*****	*****
57	AROMFUSED[2]s^{1}s^3	1,3-Dimethylnaphtalene (1)	−3.2228	−0.3360	*****	*****	*****	*****	*****	*****
58	AROMFUSED[3]	Phenalene (3), Pyrene (2)	1.6600	0.0402	−1.0430	0.004695	35.21	3.896	5.819	1.176
59	AROMFUSED[4a]	Anthracene (1)	7.0402	1.0466	3.3011	0.015244	−6.96	13.843	11.387	5.027
60	AROMFUSED[4a]s^1	9-Methylanthracene (1)	−3.3463	−7.8521	*****	*****	*****	*****	*****	*****
61	AROMFUSED[4a]s^{1}s^4	9,10-Dimethylanthracene (1)	6.8373	*****	*****	*****	*****	*****	*****	*****
62	AROMFUSED[4p]	Phenanthrene (1), Pyrene (2)	−1.5856	0.9126	2.8885	0.007280	−24.02	−16.040	−19.089	−3.417
63	AROMFUSED[4p]s^{3}s^4	9,10-Dimethylphenanthrene (1)	2.0821	*****	*****	*****	*****	*****	*****	*****
64	PYRIDINE.FUSED[2]	Quinoline (1)	−4.4725	−0.9432	1.1251	−0.005369	63.29	8.688	13.586	−4.967
65	PYRIDINE.FUSED[2–iso]	Isoquinoline (1)	−2.5898	−0.5844	3.9241	−0.011207	−2.71	−5.112	−0.314	−2.587
66	PYRIDINE.FUSED[4]	Acridine (1)	1.0358	0.1733	7.7134	−0.001275	−12.04	20.073	15.786	−1.365
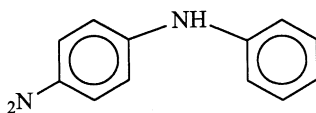
Fig. 5. Structure of *N*-phenyl-1,4-benzenediamine.

Appendix B

To illustrate the proposed method, we provide the estimation of the normal boiling point and normal melting point using six example compounds. The experimental data and estimations through Joback and Reid method [1] are also given.

Example 1. Estimation of the normal boiling point of *N*-phenyl-1,4-benzenediamine (experimental value: $T_b = 627.15$ K) (Fig. 5).

First-order groups	Occurrences	Contribution
AC–NH$_2$	1	3.8298×1
AC–NH	1	2.9230×1
aC	1	1.5468×1
aCH	9	0.8365×9
	$\sum_i N_i T_{b1i} = 15.8281$	
	$T_b = 222.543 \ln(15.8281) = 614.62$ K	
	(first-order approximately, error: 12.53 K)	

Second-order groups	Occurrences	Contribution
AROMRINGs^{1}s^4	1	0.1007×1
	$\sum_j M_j T_{b2j} = 0.1007$	
	$T_b = 222.543 \ln(15.8281 + 0.1007)$	
	$= 616.03$ K (second-order approximately, error: 11.12 K)	

Third-order groups	Occurrences	Contribution
AC–NH–aC	1	0.5768×1
	$\sum_k O_k T_{b3k} = 0.5768$	
	$T_b = 222.543 \ln(15.8281 + 0.1007 + 0.5768)$	
	$= 623.94$ K (third-order approximately, error: 3.21 K)	

Estimation through Joback
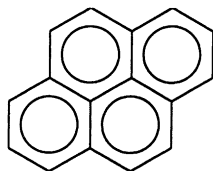and Reid [1]: 655.20 K;
error: − 28.05 K
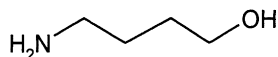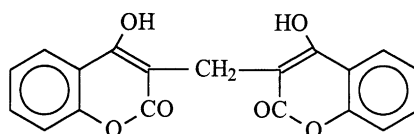
Fig. 6. Structure of pyrene.

Fig. 7. Structure of 4-aminobutanol.

Fig. 8. Structure of *N*-phenyl-1,4-benzenediamine.

Example 2. Estimation of the normal boiling point of pyrene (experimental value: $T_b = 677.15$) (Fig. 6).

First-order groups	Occurrences	Contribution
aC (fused with aromatic ring)	6	1.7324×6
aCH	10	0.8365×10
	$\sum_i N_i T_{b1i} = 18.7593$	
	$T_b = 222.543 \ln(18.7593)$	
	$= 652.43$ K (first-order approximately, error: 24.72 K)	
No second-order groups are involved		

Third-order groups	Occurrences	Contribution
AROMFUSED[3]	2	0.0402×2
AROMFUSED[4p]	2	0.9126×2
	$\sum_k O_k T_{b3k} = 1.9056$	
	$T_b = 222.543 \ln(18.7593 + 1.9056)$	
	$= 673.96$ K (third-order approximately, error: 3.19 K)	
Estimation through Joback and Reid [1]: 651.56 K; error: −24.41 K		

Example 3. Estimation of the normal boiling point of 4-aminobutanol (experimental value: $T_b = 478.15$ K) (Fig. 7).

First-order groups	Occurrences	Contribution
OH	1	2.5670×1
CH$_2$NH$_2$	1	2.7987×1
CH$_2$	3	0.7141×3
	$\sum_i O_i T_{b1i} = 7.508$	
	$T_b = 222.543 \ln(7.508)$	
	$= 448.64$ K (first-order approximately,	
	error: 29.51 K)	
No second-order groups are involved		

Third-order groups	Occurrences	Contribution
NH$_2$–(CH$_n$)$_m$–OH ($m > 2$)	1	1.0750×1
	$\sum_k O_k T_{b3k} = 1.0750$	
	$T_b = 222.543 \ln(7.508 + 1.0750)$	
	$= 478.42$ K (third-order approximately,	
	error: -0.27 K)	
Estimation through Joback and Reid [1]: 364.31 K; error: 113.84 K		

Example 4. Estimation of the normal melting point of 3,3′-methylenebis-4-hydroxycoumarin (Dicoumarol) (experimental value: $T_m = 563.15$ K) (Fig. 8).

First-order groups	Occurrences	Contribution
OH	2	2.7888×2
aC (fused with nonaromatic ring)	4	1.2065×4
ACH	8	0.5860×8
C=C (cyclic)	2	0.3048×2
CO (cyclic)	2	3.2119×2
O (cyclic)	2	1.3828×2
CH$_2$	1	0.2515×1
	$\sum_i N_i T_{m1i} = 25.1421$	
	$T_m = 147.450 \ln(25.1421)$	
	$= 475.46$ K (first-order approximately,	
	error: 87.69 K)	
No second-order groups are involved		

(Continued)
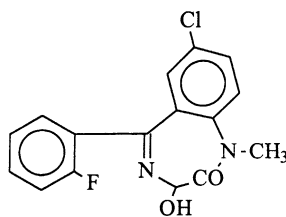
Fig. 9. Structure of flutemazepan.

Third-order groups	Occurrences	Contribution
AROMFUSED[2]	2	0.2825×2
aC–$(CH_n$=$CH_m)_{cyc}$ (in fused rings)	2	0.2479×2
aC–O (cyclic) (in fused rings)	2	-0.3545×2
$(CH_m$=$C)_{cyc}$–CH_p–$(C$=$CH_n)_{cyc}$ (in different rings)	1	16.8558×1
	$\sum_k O_k T_{m3k} = 17.2076$	
	$T_m = 147.450 \ln(25.1421 + 17.2076)$	
	$= 552.34$ K (third-order approximately, error: 10.81 K)	
Estimation through Joback and Reid [1]: 749.28 K; error: -186.13 K		

Example 5. Estimation of the normal melting point of 7-chloro-5-(2-fluorophenyl)-1,3-dihydro-3-hydroxy-1-methyl-2H-1,4-benzodiazepin-2-one (Flutemazepan) (experimental value: $T_m = 436.00$) (Fig. 9).

First-order groups	Occurrences	Contribution
OH	1	2.7888×1
aC–Cl	1	1.7134×1
aC–F	1	0.9782×1
CH_3	1	0.6953×1
AC	1	0.9176×1
aC (fused with nonaromatic ring)	2	1.2065×2
ACH	7	0.5860×7
CH (cyclic)	1	0.0335×1
CO (cyclic)	1	3.2119×1
N (cyclic)	1	0.6040×1
$(C$=$N)_{cyc}$	1	6.6382×1
	$\sum_i N_i T_{m1i} = 24.0959$	
	$T_m = 147.450 \ln(25.1421) = 469.19$ K	
	(first-order approximately, error: 33.19 K)	

(*Continued*)
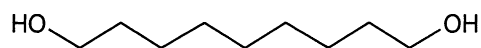
Fig. 10. Structure of 1,9-nonadiol.

Second-order groups	Occurrences	Contribution
AROMRINGs[1]s[2]	1	-0.6388×1
N_{cyc}–CH_3	1	-0.0383×1
CH_{cyc}–OH	1	1.3691×1
	$\sum_j M_j T_{m2j} = 0.6920$	
	$T_m = 147.450 \ln(24.0959 + 0.6920)$	
	$= 473.36$ K (second-order approximately, error: 37.36 K)	

Third-order groups	Occurrences	Contribution
AROMFUSED[2]s[3]	1	2.2589×1
aC–NH_n(cyclic) (in fused rings)	1	3.4983×1
aC–$(C{=}N)_{cyc}$ (in different rings)	1	-1.3060×1
aC–$(CH_n{=}N)_{cyc}$ (in fused rings)	1	-10.1007×1
	$\sum_k O_k T_{m3k} = -5.6495$	
	$T_m = 147.450 \ln(24.0959 + 0.6920 - 5.6495)$	
	$= 435.22$ K (third-order approximately, error: 0.78 K)	
Estimation through Joback and Reid [1]: 511.51 K; error: -75.51 K		

Example 6. Estimation of the normal melting point of 1,9-nonadiol (experimental value: $T_m = 318.95$ K) (Fig. 10).

First-order groups	Occurrences	Contribution
OH	2	2.7888×2
CH_2	9	0.2515×9
	$\sum_i N_i T_{m1i} = 7.8411$	
	$T_m = 222.543 \ln(7.8411)$	
	$= 303.66$ K (first-order approximately, error: 15.29 K)	
No second-order groups are involved		

(*Continued*)

Third-order groups	Occurrences	Contribution
HO–(CH$_n$)$_m$–OH ($m > 2$)	1 $\sum_k O_k T_{m3k} = 0.6674$ $T_m = 222.543 \ln(7.8411 + 0.6674)$ $= 315.70$ K (third-order approximately, error: 3.25 K)	0.6674×1
Estimation through Joback and Reid 　[1]: 312.83 K; error: 6.12 K		

References

[1] K.G. Joback, R.C. Reid, Chem. Eng. Comm. 57 (1987) 233–243.

[2] A.L. Lydersen, Estimation of critical properties of organic compounds, College Engineering University Wisconsin, Engineering Experimental Station Report 3, Madison, WI, April, 1955.

[3] D. Ambrose, Correlation and estimation of vapor–liquid critical properties. I. Critical temperatures of organic compounds, National Physical Laboratory, Teddington, UK, NPL Report Chem., 92, September 1978.

[4] K.M. Klincewicz, R.C. Reid, AIChE J. 30 (1984) 137–142.

[5] W.J. Lyman, W.F. Reehl, D.H. Rosenblatt, Handbook of Chemical Property Estimation Methods, American Chemical Society, Washington, DC, 1990.

[6] A.L. Horvath, Molecular Design, Elsevier, Amsterdam, 1992.

[7] L. Constantinou, S.E. Prickett, M.L. Mavrovouniotis, Ind. Eng. Chem. Res. 32 (1993) 1734–1746.

[8] L. Constantinou, S.E. Prickett, M.L. Mavrovouniotis, Ind. Eng. Chem. Res. 39 (1994) 395–402.

[9] S.E. Prickett, L. Constatinou, M.L. Mavrovouniotis, Mol. Simul. 11 (1993) 205–228.

[10] L. Constantinou, R. Gani, AIChE J. 40 (1994) 1697–1710.

[11] J. Marrero-Morejón, E. Pardillo-Fontdevilla, AIChE J. 45 (1999) 615–621.

[12] The CAPEC Database, Department of Chemical Engineering, DTU, Lyngby, Denmark, 2000.