Thermochimica Acta, 194 (1992) 289–296
Elsevier Science Publishers B.V., Amsterdam

Critical temperature correlation using a combined method of group contribution and topological index

Yan Xinjian, Xu Zhihong and Dong Qian

Laboratory of Computer Chemistry, Institute of Chemical Metallurgy,
Chinese Academy of Sciences, P.O. Box 353, Beijing 100080 (People's Republic of China)

(Received 26 April 1991)

Abstract

In this paper, a new approach which combines the group contribution and topological index methods is proposed for correlating the thermodynamic properties of organic compounds. The critical temperatures of alkanes, alkenes, alcohols, acids, ketones, amines and sulphur ethers have been correlated with this method, and an average error of 0.45% has been obtained for 200 compounds.

INTRODUCTION

The thermodynamic properties of compounds are indispensable for industrial design, simulation and research. In general, the existing experimental data can satisfy only a small part of the requirements of chemical engineers. A difficult situation that engineers often face is the lack of data for chemical industrial processes. Therefore, correlation methods are very important.

The most popular method in the chemical industry for estimating thermodynamic properties of organic compounds is the group contribution (GC) method by virtue of its simplicity and adaptability [1,2]. The basic idea of the method derives from the fact that organic compounds are composed of a limited number of groups, and the assumption that the compound properties can be approximately estimated by the addition of these group properties. The disadvantage of the method is that all group contribution methods are necessarily approximations because any group within a molecule is not completely independent of the other groups in that molecule [2]. Thus, it is impossible to generate highly accurate results by this method.

In another approach, correlation methods of compound properties with topological indices (TI) have developed rapidly since the 1970s [3,4]. Many indices have now been employed for predicting refractive indices, normal boiling points, heats of formation, heats of combustion, molar volumes, critical properties, etc. Compared with the group contribution method, fewer parameters are used and more integral characteristics of the compounds are taken into account in the TI method. The main problem of the TI method is that it is inadequate for compounds containing heteroatoms. Therefore, it is very difficult to get a TI correlation which is applicable to several kinds of compounds simultaneously.

The GC and TI methods have their own advantages and disadvantages. The aim of this work is to try to combine the two methods for studying thermodynamic properties. The critical temperatures (T_c) of some kinds of organic compounds are correlated by this method, in which a set of simple classified groups are used to study the group contributions to T_c. Concurrently, a modified TI method is proposed and used, which can give a good correlation for heteroatoms.

BASIC CONSIDERATIONS

Fluid properties can be affected by many factors: the properties of individuals, i.e. the properties of atoms or groups; the local characteristics of a molecule, determined by the interactions of neighbouring groups within the molecule; the characteristics of a molecular frame (integral shape); and the interactions of different molecules (for high pressure gases and liquids). In principle, the thermodynamic properties of organic compounds are dependent on these factors.

In practice, the individual component properties of a molecule are often represented by group properties, and the molecular properties can be approximately obtained by the additivity of these group properties. This is just what the principle of GC methods covers. The GC methods with simple group classifications cannot produce reliable estimation results because of the interactions among atoms within the molecule. Therefore, complex group classifications in some GC methods are employed for improving the estimation ability. However, this not only brings up the difficulty of fitting a large number of group parameters owing to few experimental data, but also makes it more complicated for the chemist and engineer to use the GC methods. It can be expected that group classification will be simplified if other approaches are adopted for depicting the local and integral characteristics of the molecule.

One possible way of describing the local characteristics of a molecule would be to generate a topological index for each group of the molecule. The value of a group's topological index should be determined by the affection of the group environment that closely relates to the interactions between this group and its neighboring groups. Based on this idea the authors successfully correlated standard heats of formation for very branched alkanes [8].

As pointed out in the introduction, TI provides one type of description of the molecular integral characteristics. Though TI can also be used to correlate the properties of organic compounds by itself, the authors hold that it is better to treat TI as a correcting factor and to use it together with other methods.

It is demonstrated by the UNIFAC method [2] that the interactions between two molecules can be approximately described by the interactions among groups in the two molecules. From a practical viewpoint, this scheme is more successful in estimating properties of mixtures on phase equilibrium.

The above four factors do not have equal weight. In most cases the atom or group properties of the compounds play a dominant role in determining the compound properties used in the chemical industry. Other factors can be used for correcting purposes. For liquid mixtures, all four factors may need to be considered. For some properties of pure compounds, improved correlation may be obtained if the two factors of GC and TI are taken into account in data fitting. In the following example, GC and TI are adopted for correlating the critical temperatures of the chain organic compounds.

CRITICAL TEMPERATURE CORRELATION

Critical temperature is widely used in the chemical industry as the basic property of organic compounds. The correlations available are primarily of the group contribution type. In most of them, an experimental value (usually the normal boiling point) for each compound is used to reflect the characteristics of the compound structure to obtain high accuracy. However, experimental data are not usually available for the compounds of interest to chemists. Fedors [7] proposed a method in which no experimental value is used. Unfortunately, the method is not accurate enough. Here, we use the following correlation for the critical temperature:

$$T_c = (\mathrm{GC}) + (\mathrm{TIC}) \tag{1}$$

where T_c is the critical temperature (K), (GC) is the contribution of the groups and (TIC) is the contribution of the molecular topological feature. The forms of (GC) and (TIC) are given in eqns. (2) and (3). The coefficients in both equations are obtained by fitting the literature data [1,5] of alkanes, alkenes, alcohols, acids, ketones, amines and sulphur ethers:

$$(\mathrm{GC}) = 603.225\left(1.0 + \sum B_i\right)\left[\log_{10}\left(\sum G_i\right)\right] \tag{2}$$

$$(\mathrm{TIC}) = 4.03725\,|\mathrm{TP}|^{c}/\sqrt{n} \tag{3}$$

where $c = 3.09330$, G_i and B_i are parameters of group i, n is the number of groups in a compound and TP is the extended ADSCI (Average

TABLE 1

Group parameters

No.	Group	G	E	$B \times 10^2$
1	$-CH_3$	1.55732	0.53724	0.06005
2	$>CH_2$	0.82833	0.44894	0.06960
3	$>CH-$	−0.31402	0.49894	−0.85195
4	$>C<$	−2.31852	0.57519	1.45858
5	$-OH$	4.56540	−2.97935	−0.61210
6	$-NH_2$	3.51974	−1.22101	−0.86889
7	$-SH$	4.12021	−0.94129	2.00135
8	$=CH-$	0.72046	0.59769	0.66138
9	$=C<$	−3.24965	1.67920	1.80860
10	$=CH_2$	1.41938	0.72917	−0.00376
11	$>C=O$	4.43964	−0.35463	−4.10921
12	$-S-$	4.04635	−0.41392	−2.26873
13	$-COOH$	9.40452	−1.52420	−5.67222

Distance Sum Connectivity Index). Equation (4) is the ADSCI proposed by Balaban [6]

$$J = [Q/(L+1)] \sum \left(1/\sqrt{D_i D_j}\right) \tag{4}$$

where the summation term in the above equation only applies to the edges (the bonds between groups), i and j are the end groups of each edge, D is the distance sum of a specific group from other groups in a given molecular distance matrix, L the cyclomatic number, and Q the edge number of a molecular graph.

The original ADSCI is good for alkanes, but inadequate for other kinds of compounds. Therefore the ADSCI is extended into a new form

$$\mathrm{TP} = [Q/(L+1)] \sum \left[(E_i + E_j)/\sqrt{D_i D_j}\right] \tag{5}$$

where E_i and E_j are the effects of group types i and j, respectively, on the topological index, which can be obtained by fitting literature data. If all E_i and E_j equal 0.5, eqn. (5) will be reduced to eqn. (4).

Table 1 shows the parameters of 13 groups, which were obtained by fitting literature data of 202 compounds. The E values of the groups from No. 1 to No. 4 in the table are close to 0.5, whereas the values of others deviate obviously from 0.5. This means that the original ADSCI is applicable for alkanes, but may be not suitable for other kinds of compounds. Table 2 lists the correlated results of eqn. (1) and a GC method. The results of the GC method were obtained by using eqn. (2) with another set of specifically fitted parameters, which are not presented to save space.

In order to examine the proposed technique in this article, firstly 9 compounds were randomly taken from the 202 compounds, then the

TABLE 2

Correlated results for critical temperatures

Compound class	Number of compound	Eqn. (1)		GC method	
		DT (K) [a]	Err. (%) [b]	DT (K) [a]	Err. (%) [b]
Alkanes	62	2.328	0.421	7.858	1.415
Alkenes	15	1.527	0.321	3.844	0.832
Acids	13	1.593	0.247	2.629	0.419
Alcohols	25	3.520	0.607	13.588	2.399
Amines	23	2.978	0.524	5.763	1.103
Sulphur alcohols	23	3.857	0.680	7.728	1.386
Ketones	10	1.581	0.277	2.963	0.518
Sulphur ethers	31	2.249	0.373	3.951	0.665
Average	202	2.57	0.45	6.84	1.23

[a] DT = calculation value − literature value.

[b] Err. (%) = 100 (DT/lit. value).

remaining 193 compounds were used to refit a new set of parameters for eqn. (1); the parameters are not listed here. In Table 3, part A shows the predicted results for the above 9 compounds via eqn. (1) with these new parameters; part B gives the correlated results of the same 9 compounds by eqn. (1) with the original parameters in Table 1.

TABLE 3

Comparison between predicted and correlated T_c values

No.	Compound	T_c lit. value (K)	Part A [a]		Part B [b]	
			DT (K) [c]	Err. (%) [d]	DT (K) [c]	Err. (%) [d]
1	3,3-Dimethylpentane	536.3	−4.20	−0.78	−3.97	−0.74
2	2,4-Dimethylhexane	554.1	3.82	0.69	3.68	0.66
3	3,3,5-Trimethylheptane	609.5	3.51	0.58	3.27	0.54
4	3-Methyl-1-butene	464.8	−2.47	−0.53	−1.33	−0.29
5	Octanoic acid	695.0	−1.27	−0.18	0.48	0.07
6	3-Methyl-1-butanol	579.6	−8.63	−1.49	−8.15	−1.41
7	2-Ethyl-1-hexanamine	627.0	−1.15	−0.18	−0.79	−0.13
8	2-Butanone	535.6	0.68	0.13	0.53	0.10
9	1-[(1-Methyl-ethyl) thio]-propane	594.0	1.45	0.24	1.42	0.24
	Average error		3.02	0.54		

[a] Part A is predicted from eqn. (1).

[b] Part B is correlated from eqn. (1).

[c] DT = calculated value − literature value.

[d] Err. (%) = 100 (DT/lit. value).

TABLE 4

Correlated T_c results for alkanes

No.	Compound [a]	Lit. (K) [b]	DT (K) [c]	Err. (%) [d]	(GC) (K) [e]	(TIC) (K) [f]
1	Ethane	305.4	−3.84	−1.26	298.0	3.6
2	Propane	369.7	0.58	0.16	360.1	10.2
3	Butane	425.6	−0.91	−0.21	410.4	14.3
4	2-Me.propane	407.7	5.93	1.45	383.0	30.6
5	2-Me.butane	461.1	−0.21	−0.05	428.6	32.3
6	Pentane	470.0	−0.54	−0.12	452.8	16.7
7	2,2-Di.propane	433.8	6.49	1.50	363.3	76.9
8	2,2-Di.butane	488.2	−1.00	−0.21	414.8	72.4
9	2,3-Di.butane	500.3	−2.33	−0.47	444.8	53.2
10	Hexane	507.9	−0.43	−0.08	489.4	18.1
11	2-Me.pentane	498.5	−0.46	−0.09	467.5	30.5
12	3-Me.pentane	504.4	−1.42	−0.28	467.5	35.4
13	2,2,3-Tri.butane	530.4	0.32	0.06	433.5	97.2
14	Heptane	540.1	0.42	0.08	521.6	18.9
15	2-Me.hexane	531.0	−0.95	−0.18	501.5	28.5
16	3-Me.hexane	535.4	0.30	0.06	501.5	34.2
17	2,2-Di.pentane	520.4	−1.56	−0.30	458.1	60.8
18	2,3-Di.pentane	537.6	−3.18	−0.59	480.7	53.7
19	2,4-Di.pentane	520.8	3.67	0.70	480.7	43.7
20	3,3-Di.pentane	536.3	−3.97	−0.74	458.1	74.3
21	3-Et.pentane	540.7	1.24	0.23	501.5	40.4
22	2,2,3,3-Te.butane	567.8	0.87	0.15	419.1	149.6
23	2-Me.heptane	560.6	−2.01	−0.36	531.7	26.9
24	3-Me.heptane	564.1	−0.48	−0.08	531.7	32.0
25	4-Me.heptane	562.2	3.58	0.64	531.7	34.1
26	2,2-Di.hexane	549.9	−3.51	−0.64	495.4	51.0
27	2,3-Di.hexane	564.5	−2.98	−0.53	512.4	49.1
28	2,4-Di.hexane	554.1	3.68	0.66	512.4	45.4
29	2,5-Di.hexane	549.9	0.20	0.04	512.4	37.7
30	3,3-Di.hexane	562.8	−1.17	−0.21	495.4	66.2
31	3,4-Di.hexane	569.7	−1.95	−0.34	512.4	55.4
32	3-Et.hexane	566.1	5.42	0.96	531.7	39.8
33	Octane	569.1	0.69	0.12	550.5	19.3
34	2-Me.-3-Et.pentane	567.0	3.71	0.65	512.4	58.3
35	3-Me.-3-Et.pentane	576.5	−1.37	−0.24	495.4	79.7
36	2,2,3-Tri.pentane	563.5	0.44	0.08	473.2	90.7
37	2,2,4-Tri.pentane	543.8	2.27	0.42	473.2	72.9
38	2,3,3-Tri.pentane	573.5	−2.39	−0.42	473.2	97.9
39	2,3,4-Tri.pentane	566.4	−2.32	−0.41	492.5	71.5
40	2,2-Di.heptane	576.8	−4.68	−0.81	528.3	43.8
41	2,2,3-Tri.hexane	588.0	−1.73	−0.29	507.9	78.4
42	2,2,4-Tri.hexane	573.5	3.99	0.70	507.9	69.6
43	2,2,5-Tri.hexane	568.0	−1.99	−0.35	507.9	58.2
44	Nonane	594.8	1.27	0.21	576.6	19.5
45	2-Me.octane	587.0	−2.58	−0.44	558.8	25.6
46	2,2,3,3-Te.pentane	607.5	0.37	0.06	463.3	144.6
47	2,2,3,4-Te.pentane	592.6	2.04	0.34	486.8	107.9

TABLE 4 continued

No.	Compound [a]	Lit. (K) [b]	DT (K) [c]	Err. (%) [d]	(GC) (K) [e]	(TIC) (K) [f]
48	2,2,4,4-Te.pentane	574.5	−7.78	−1.35	463.3	103.4
49	2,3,3,4-Te.pentane	607.5	0.54	0.09	486.8	121.3
50	Decane	616.1	3.85	0.62	600.4	19.6
51	3,3,5-Tri.heptane	609.5	3.27	0.54	538.6	74.2
52	2,2,3,3-Te.hexane	623.0	4.06	0.65	501.4	125.7
53	2,2,5,5-Te.hexane	581.4	−1.81	−0.31	501.4	78.2
54	*n*-Undecane	640.1	1.74	0.27	622.3	19.5
55	*n*-Dodecane	659.5	2.57	0.39	642.7	19.4
56	*n*-Tridecane	678.4	2.49	0.37	661.7	19.2
57	*n*-Tetradecane	696.8	1.68	0.24	679.5	19.0
58	*n*-Pentadecane	717.6	−2.59	−0.36	696.2	18.8
59	*n*-Hexadecane	734.3	−3.68	−0.50	712.1	18.6
60	*n*-Heptadecane	749.3	−3.91	−0.52	727.1	18.3
61	*n*-Octadecane	763.2	−3.77	−0.49	741.3	18.1
62	*n*-Nonadecane	776.0	−3.19	−0.41	755.0	17.9

[a] Me., methyl; Et., ethyl; Di., dimethyl; Tri., trimethyl; Te., tetramethyl.
[b] Literature value.
[c] DT = calculated value − literature value.
[d] Err. % = 100 (DT/lit. value).
[e] Contribution of the group part.
[f] Contribution of the topological part.

DISCUSSION AND CONCLUSION

The results in Table 2 show that eqn. (1) gives excellent correlated results; its average error is much less than the value obtained from group contribution only. According to Reid et al. [1], the methods of Ambrose and Joback (normal boiling points are used in both methods) are the best ones for T_c, their average errors being about 0.7 and 0.8%, respectively. Considering that no experimental value was used in eqn. (1), one may conclude that the TI method does reflect the molecular characteristics to some extent, and that the combined GC and TI method is meaningful. In Table 3, the errors in the predicted values are only slightly larger than those in the correlated results, which proves the reliability of our method.

The calculated results for alkanes are listed in Table 4 for discussion. It is noteworthy that the group part makes more contribution than the topological part. This is consistent with our previous consideration that the component properties of the compounds are the dominant factors, whereas TI is a structure-correcting factor that can describe the difference in structural isomers. For example, 2,2,3,3-tetrahexane has same groups as 2,2,5,5-tetrahexane, but their difference in T_c is about 42 K. This is a difficult case for the GC method, but good correlated results (Table 4) are obtained for these two molecules using eqn. (1).

In principle, the method proposed in this paper could be extended to other kinds of compounds. However, a ring parameter may be needed for ring compounds. The study of other kinds of compounds with this method is in progress.

REFERENCES

1 R.C. Reid, J.M. Prausnitz and B.E. Poling, The Properties of Gases and Liquids, McGraw–Hill, New York, 4th edn., 1987.
2 A. Fredenslund, J. Gmehling and P. Rasmussen, Vapor–Liquid Equilibria using UNIFAC, Elsevier, Amsterdam, 1977.
3 N. Trinajstic, Chemical Graph Theory, CRC Press, Boca Raton, 1983.
4 A.T. Balaban, I. Motoc, D. Bonchev and O. Mekenyan, Top. Curr. Chem., 114 (1983) 21.
5 K.H. Simmrock, R. Janowsky and A. Ohnsorge, Critical Data of Pure Substances, Chemistry Data Series Vol. II, Dechema, Frankfurt, 1986.
6 A.T. Balaban, Chem. Phys. Lett., 89 (1982) 399.
7 R.F. Fedors, Chem. Eng. Commun., 16 (1982) 149.
8 X. Yan, Z. Xu and Q. Dong, J. Fuel. Chem. Technol., 18 (1990) 77.