Estimation of boiling and melting points of light, heavy and complex hydrocarbons by means of a modified group vector space method

Estimation of Boiling and Melting Points of Light, Heavy and Complex Hydrocarbons by Means of a Modified Group Vector Space Method

Hengde Li, Hidenori Higashi, Kazuhiro Tamura*

Department of Chemistry and Chemical Engineering, Division of Material Engineering and Material Sciences, Graduate School of Natural Science & Technology, Kanazawa University, Kakuma-machi, Kanazawa 920-1192

* Corresponding author. FAX: +81-76-234-4829, e-mail: tamura@t.kanazawa-u.ac.jp

Abstract

A modified group vector space (GVS) method was developed for estimating the normal boiling points and melting points of alkanes, alkenes, alkynes, and cyclic and aromatic hydrocarbons including their isomers. The present method, based on group contributions as well as topological contributions, can represent the normal boiling points of isomeric compounds accurately. The group parameters for the modified GVS method were obtained from the correlation of the boiling and melting points of 1115 hydrocarbons.

Keywords: Normal boiling point; melting point; hydrocarbon; estimating method; topological index

1. Introduction

Normal boiling points and melting points of pure components are of primary importance for the design and development of many chemical processing units. In a distillation process petroleum mixtures are separated into their fractions by the boiling points. As well in a crystallization unit, the separation from liquid mixtures to pure solid is carried out using the melting points. If it is not always possible to find reliable experimental values in practice, the estimation method of normal boiling and melting points of the fractions is required. For the estimating methods found in the literature, there are function methods, group contribution methods and quality structural property research. Riazi and Al-Sahhaf [1] proposed a function method to estimate physical properties of single carbon number hydrocarbon groups like C_6-C_{50} n-alkanes, n-alkylcyclopentanes and n-alkylbenzenes using an exponential function:

$$\theta = \theta_\infty - \exp(a - bM^c) \tag{1}$$

where a, b, and c are the constants of the hydrocarbons, and θ and M represent a property such as T_b and molecular weight. θ is the limiting value for the property as M . This method can represent well various properties for these compounds, but cannot apply to estimate the properties of structural isomers.

The group contribution method (GC) for the estimation method of physical properties of pure compounds proposed by Joback and Reid [2] (JR method) is popular among the group-contribution methods reported previously. The normal boiling points T_b and melting points T_m of pure organic compounds estimated from the JR group contribution method can be expressed by

$$T_b = 198.2 + \sum N_i C_{bi} \tag{2}$$

$$T_m = 122.5 + \sum N_i C_{mi} \tag{3}$$

where N_i is the number of groups of type i, and C_{bi} and C_{mi} are respectively the group contributions of the boiling and melting points resulting from the group i of the molecule. The results obtained from Eqs. (2) and (3) are useful guides of the estimation but they are of questionable in accuracy and undistinguishable among structural isomers. Cordes and Rarey [3] presented a new method (GC^I) for the estimation of normal boiling points of non-electrolyte organic compounds which was based exclusively on the molecular structure of the compound. In their research, 86 main structural groups (25 main structural groups for hydrocarbons) and few second-order groups are used and the prediction of the boiling points was performed by using the following equation

$$T_b = \frac{\sum N_i C_i}{n^a + b} + c \tag{4}$$

where a, b, and c are the adjustable parameters, N_i the number of groups of type i, C_i the group contributions of group i, and n is the number of atoms in a molecule except for hydrogen. The estimated T_b of hydrocarbons by Eq. (4) shows better agreement with the experimental results than that of the JR method. However, this method is not taken into account the molecular structural effect among the isomeric compounds and cannot represent well the boiling points of structural isomers of the hydrocarbon.

To overcome this limitation, complex group contribution methods have been developed. A second order group contribution method [4] was used to estimate the boiling points of alkanes from C_1 to C_{10} and also represent well the normal boiling points of alkane isomers, wherein 5 main groups and 17 complex second-order groups are employed. A two level group contribution method [5] for the estimation of properties of pure organic compounds was presented. The resultant equations were obtained in a logarithmic function for the estimation of the normal boiling points and melting points, and greatly improved the results obtained from the JR method. Furthermore, a three level group contribution method (GC^{II}) proposed by Marrero and Gani [6] assumes that the primary level uses contributions from simple groups, and the higher levels involve polyfunctional and structural groups that provide more information about molecular fragments whose description does not consider in the first-order groups. In their model, 182 first-order groups, 122 second-order groups and 66 third-order groups were used to estimate pure component properties, and the expressions of normal boiling points T_b and melting points T_m are

$$T_b = 222.543 \ln\left(\sum_i N_i C_{b1i} + \sum_j M_j C_{b2j} + \sum_k O_k C_{b3k} \right) \tag{5}$$

$$T_m = 147.450 \ln\left(\sum_i N_i C_{m1i} + \sum_j M_j C_{m2j} + \sum_k O_k C_{m3k} \right) \tag{6}$$

where C_{b1i}, C_{b2j} and C_{b3k} are the first-order, second-order and third-order group contributions of normal boiling point, C_{m1i}, C_{m2j} and C_{m3k} are the first-order, second-order and third-order group

contributions of melting point, and N_i, M_j and O_k are the number of the first-order, second-order and third-order groups in the molecule. On the other hand, Marrero and Pardillo proposed a group-interaction contribution (GIC) method [7], which considers the contributions of interactions between bonding groups in the molecule instead of the contribution of simple groups. The properpties of a compound are considered to be functions of structurally dependent parameters, which are thereby determined by summing the number frequency of each group-contribution (bonding group) occurring in the molecule times its contribution. The normal boiling points of pure organic compounds is expressed as

$$T_b = M^{-0.404} \sum N_i C_i + 156.00 \tag{7}$$

where M is the molecular weight. N_i and C_i represent the number frequency and contribution of bonding groups of type i. This method makes a important improvement in distinguishing among structural isomers since the number of basic bonding groups is more than the number of basic groups, e.g., the number of the basic bonding groups was 10 in their research for alkanes but that of basic groups was 4 in general group contribution methods.

In order to distinguish the T_b among structural isomers, Toropov [8] proposed a quality structural property research (QSPR) model for calculating the boiling points of cyclic hydrocarbons by using a maximal topological distance matrix from a molecular graph. Cao and Yuan [9] proposed a QSPR model for the estimation of the boiling points of paraffins and cycloalkanes by using the topological indices based on a vertex, distance and ring from a molecular graph. These QSPR models can calculate well the T_b of structural isomers for a limited number of hydrocarbons in their researches.

Xu and Yang [10, 11] considered the specific position of groups in a molecule and developed a group vector space (GVS) method for estimating boiling and melting points of organic compounds. In their research the group topological indices, obtained from a group vector space and based on a molecular graph, are combined with the group contributions to the boiling and melting points of hydrocarbons. This method improves the estimation of the boiling and melting points of hydrocarbons better than the GC methods described above in both accurate representation and capability of distinguishing isomers.

In the present paper, our purpose is aimed at further development in accurate representation of the normal boiling points and melting points over a wide range and variety of hydrocarbons and their isomers by means of a simple modification of the GVS method.

2. Modification of GVS method

The GVS method assumes that the topology of a molecule can be described by a topological matrix ($k\times$), where j and k represent the group (point) number and molecular dimension number in a molecular graph. The dimension number k is equal to the total sum of both the number of end points and the number of rings in the graph. Fig. 1 shows the chemical structures and groups of typical compounds and the topology of the group expressed in terms of the graph

and corresponding matrix. According to Xu and Yang, the topological matrix ($k \times$) is designated by the shortest distance evaluated from the following constrains:

(1) The distance from an end point to other any group equals to the actual shortest distance between them. Particularly when the group point is in itself, the distance corresponds to zero.

(2) From a ring to any group on the self-ring, the distance is equivalent to the number of the points on the ring. From a ring to any group outside the ring, the distance is defined by the number of the points on the ring plus the shorted distance outside the ring.

The matrix of any group j is represented by a k-dimensional vector matrix ($m_{1j}, m_{2j},...,m_{kj}$) to characterize the position of the group in the molecule. The module $_j$ for any point j vector

$$\alpha_j = \left(\sum_k m_{jk}^2 \right)^{1/2} \tag{8}$$

The topological index $_j$ of any group j in the graph is defined by the average square root of the module of the matrix ($k \times$). They assigned the group topological matrix to a topological index $_j$, which was called as the module index in the original papers.

$$\nu_j = \alpha_j \left/ \left(\sum_j \alpha_j^2 \right)^{1/2} \right. \tag{9}$$

Using the boiling points for 402 hydrocarbons including 92 alkanes, 120 alkenes, 110 cyclics and 80 aromatics and the melting points for 339 hydrocarbons including 67 alkanes, 102 alkenes, 92 cyclics and 78 aromatics, Xu and Yang [10] obtained the equations for estimating normal boiling points T_b and melting points T_m of hydrocarbons as follows

$$T_b^{2.0} = -10157.64 + \sum_i \left(N_i C_{bgi} + \sum_j^{n_i} \nu_j C_{bti} + C_{bi} \right) \tag{10}$$

$$T_m = 72.42 + \sum_i \left(N_i C_{mgi} + \sum_j^{n_i} \nu_j C_{mti} + C_{mi} \right) \tag{11}$$

where, C_{bgi}, C_{bti}, and C_{bi} and C_{mgi}, C_{mti}, and C_{mi} are respectively the group-independent contribution, group-topological contribution, and group constant of normal boiling point and melting point for groups of type i; N_i is the number of groups of type i. Remarkably they introduced the power index into T_b to improve the boiling point estimation. The different form of the equations is used to the estimation for the boiling points and melting points. On the contrary, we propose in this research that the estimated temperatures of T_b and T_m can be expressed by a linear function of combining the group-independent and group-topological contributions

$$T_b = T_b(group\text{-}independent) + T_b(group\text{-}topolog\,ical) + C_{b0}$$
$$= \sum_i N_i C_{bgi} + \sum_i \sum_j^{n_i} \nu_j C_{bti} + C_{b0} \tag{12}$$

$$T_m = T_m(group\text{-}independent) + T_m(group\text{-}topo\log ical) + C_{m0}$$

$$= \sum_i N_i C_{mgi} + \sum_i \sum_j^{n_i} v_j C_{mti} + C_{m0} \tag{13}$$

where C_{b0} and C_{m0} are the correlation constants for boiling and melting points.

3. Calculation Procedure

3.1 Database

The experimental data for 1115 hydrocarbons including alkanes, alkenes, alkynes, cyclic and aromatic hydrocarbons were used in this study. The number of the experimental data for boiling temperature is 1041 and 622 for melting temperature. Most of them were taken from the CRC handbook [12] which compiles basic physical constants for over 12000 organic compounds most frequently encountered in a chemical factory and laboratory. The experimental values of heavy hydrocarbons were taken from the literature [13] and the others were from the reference [14]. They include alkanes over the range of carbon number from C_2 to C_{100}, alkenes and alkynes from C_2 to C_{22}, cyclic hydrocarbons from C_3 to C_{42}, and aromatic hydrocarbons from C_6 to C_{28}. They include 236 n-alkyl hydrocarbons and their 879 structural isomers.

3.2 Determination of the group parameters

Multiple linear regression techniques were used to determine the group parameters for group-independent and group-topological contributions. To obtain a set of the optimum values in the regression procedure, we minimized the squared deviations between the experimental and estimated values, T_{exp} and T_{est}, i.e.,

Minimize the objective function: $F = \sum (T_{est} - T_{exp})^2$ (14)

The regression results (*rms* and *r*) between the experimental and calculated values for the boiling and melting points of the hydrocarbons are given in Table 1. The root mean square error *rms* between the experimental and estimated values and multiple correlation coefficient *r* of the regression are defined by

$$rms = \sqrt{F/N} \tag{15}$$

$$r = \frac{N\left(\sum T_{est}T_{exp}\right) - \left(\sum T_{est}\right)\left(\sum T_{exp}\right)}{\left\{\left[N\sum T_{est}^2 - \left(\sum T_{est}\right)^2\right]\left[N\sum T_{exp}^2 - \left(\sum T_{exp}\right)^2\right]\right\}^{1/2}} \tag{16}$$

where N is the number of experimental data points. From the data reduction, we obtained the following expressions for boiling and melting point estimation.

$$T_b = \sum_i N_i C_{bgi} + \sum_i \sum_j^{n_i} v_j C_{bti} - 67.68 \qquad (17)$$

$$T_m = \sum_i N_i C_{mgi} + \sum_i \sum_j^{n_i} v_j C_{mti} - 107.53 \qquad (18)$$

Table 2 lists the group-independent and group-topological contributions to T_b and T_m, along with the 16 groups divided for hydrocarbons studied in this work. The 14 groups listed in Table 2 are the same as those used by Joback and Reid, except for the ring group =C< in the Joback and Reid method. The group =C< that stands for double bond on a ring in cycloalkenes or aromatic hydrocarbons has possibilities to connect to a chain group, another ring group having a aromatic-ring group and a fused-ring group including a fused-aromatic-ring group, and so on. In order to improve the estimation accuracy and distinguish the temperature differences of hydrocarbon isomers in boiling and melting point estimation, we can divide the ring group =C< into group types with parameters as a few as possible [3, 6]. For this reason, the ring group =C< was classified by two group types =C<(c) and =C<(r) in this research. In the melting point estimation, the number N_{CH3} of group –CH$_3$ of n-alkanes was corrected as (N_{CH3}+2) empirically for accurate representation. To explain the calculation of boiling point T_b and melting point T_m of hydrocarbons in terms of the group topological indices, we present an illustrative procedure for 2-methylpentane in Table 3 using the topological matrix shown in Fig. 1.

3.3 Predictive capability of the proposed model

For the 1041 experimental T_b, we divided them randomly into a training set of 841 compounds for the correlation of the model and a test set of 200 compounds. Every 50 components for alkanes, alkenes and alkynes, and cyclic and aromatic hydrocarbons were selected arbitrably as the test set of the 200 compounds. In order to test the predictive capability of the models, we compared the experimental boiling points for the 200 compounds selected as the test set with the results predicted by the JR and GVS models and the present method. The average absolute errors in the boiling temperatures are 4.83K for the present method, 8.75K for the GVS method, and 23.15K for the JR method.

4. Calculated Results and Discussion

The proposed method for the boiling point and melting point estimations of hydrocarbons includes less parameter than those of other GC^{II}, GIC and GVS methods. Table 4 compares the number of parameters for the estimation of boiling and melting points of alkanes used in several estimation methods.

4.1 Boiling point estimation

For the results of estimated T_b of the hydrocarbons, Table 5 shows the average absolute error (*AAE*) and absolute percent error (*APE*) and compares with those of JR method and GVS method. The modified GVS method gives more accurate results whose the average absolute deviations for the alkanes, alkenes and alkynes, and cyclic and aromatic hydrocarbons were 0.57%, 0.76%, 0.91%, and 1.23%, respectively and 0.86% for over-all the data of 1041 points examined and shows an accurate representation for light hydrocarbons as well as heavy and complex hydrocarbons. For 247 alkanes, the *AAE* and *APE* for T_b estimated by the present method were 2.78K and 0.57%. The numbers of alkanes that lies in the *AAE* ranges 0~5K, 5~10K and 10K~ of 247 alkanes were 236, 5 and 2 respectively. The boiling points of most of alkanes were estimated successfully within a temperature of 5K. As shown in Table 5, the T_b of C_{11} to C_{100} calculated by the present method was improved noticeably in comparison with the previous methods. Table 6 shows the present model can represent the T_b for n-alkyl hydrocarbons better than the previous models.

The estimated T_b for 18 different isomers of n-octane with the experimental values is compared in Table 7. The conventional methods can not represent the boiling points of the isomers distinguishably, but the GVS method and the proposed method the conventional methods could distinguish the temperature differences between the n-octane isomers. This indicates that the boiling points of these isomers can be represented by not only the groups but the topological factors between the groups. Fig. 2 compares the $T_{b,exp}$ with $T_{b,est}$ of n-alkanes from C_2 to C_{100} calculated by the several methods. The proposed method shows good agreement with the experimental results and the conventional methods show larger deviations as increasing the carbon atom number, C>30.

For 296 unsaturated hydrocarbons including alkenes, dienes and alkynes, the AAE and APE of T_b calculated from the proposed method were 2.82K and 0.76 %, and $T_{b,est}$ of 1-alkenes from C_2 to C_{25} obtained by the estimation methods is compared with the $T_{b,exp}$ in Fig. 3.

For 235 cycloalkanes and cycloalkenes, the *AAE* and *APE* between the experimental and estimated T_b were 3.79K and 0.91%. The $T_{b,est}$ of n-alkylcyclohexanes from C_6 to C_{45} for the estimation methods is compared with $T_{b,exp}$ in Fig. 4. The estimated values for n-alkylcyclopentanes were similar as those of n-alkylcyclohexanes. The proposed method shows good agreement with the experimental values and the other methods show larger deviations for heavy hydrocarbons.

For 263 aromatic hydrocarbons including benzenes, naphthalenes, polyaromatics and poly-rings, the *AAE* and *APE* of T_b of 263 aromatic hydrocarbons calculated from the proposed method were 6.89K and 1.23%. The $T_{b,est}$ of n-alkylbenzenes from C_6 to C_{35} obtained by the estimation methods is compared with the $T_{b,exp}$ in Fig. 5. As shown in Table 5, the results of estimated T_b for the 263 aromatic hydrocarbons by the present method were better than the other methods. However the deviations for poly-ring aromatic compounds were slightly lager than those of benzenes, aliphatic compounds or cyclic compounds.

8

4.2 Melting point estimation

For the results of estimated T_m, The average absolute error (AAE) and absolute percent error (APE) of hydrocarbons are listed in Table 8 and compared with those of the JR and GVS methods. The average absolute error and absolute percent error were 20.33K and 8.56% for all the 622 data examined. The calculated results show the present model is superior to the GVS method. As depicted in Table 9, the results for n-alkyl hydrocarbons estimated by the proposed method show better agreement with the experimental values than the previous methods. Table 10 compares the melting points with those estimated by the JR, GC^{II}, and GVS methods. The deviations of T_m in the literatures [6, 10] were about 7~8% and comparable with those obtained in the present work.

For 110 alkanes given in Table 8, the AAE and APE of T_m estimated results by the present method were 15.70K and 7.82%. The $T_{m,est}$ of n-alkanes from C_2 to C_{45} calculated by the estimation methods is compared with $T_{m,exp}$ in Fig. 6.

For 153 alkenes, the AAE and APE of T_m were 14.18K and 8.43%. The $T_{m,est}$ of 1-alkenes from C_3 to C_{25} obtained by the estimation methods is compared with the $T_{m,exp}$ in Fig. 7.

For 133 cyclic hydrocarbons, the AAE and APE of T_m obtained by the present method were 14.49K and 7.50%, and the estimated results for n-alkylcyclohexanes from C_7 to C_{45} are compared with $T_{m,exp}$ in Fig. 8.

For 226 aromatic hydrocarbons, the AAE and APE of T_m were 30.19K and 9.62%. The $T_{m,est}$ of n-alkylbenzenes from C_7 to C_{35} obtained by the estimation methods is compared with the $T_{m,exp}$ in Fig. 9.

The absolute percent error of T_b between the estimated and experimental values was 0.86% for 1041 hydrocarbons and that of T_m 8.56% for 622 hydrocarbons. The accuracy estimated for melting points was much lower than that of boiling points. As shown in Tables 8 and 9, the absolute percent error of melting points between the estimated and experimental values were 7.82% for alkanes and 6.27% for n-alkanes, but those were 10.16% and 13.67% for the same alkanes and n-alkanes without N_{CH3} number correction. The results estimated for n-alkanes were improved greatly when N_i of Eq. (18) used for group –CH$_3$ of n-alkanes in the melting point estimation was corrected as (N_{CH3}+2) empirically. To explain this problem, Fig. 10 illustrates the experimental T_b and T_m of n-alkanes, 1-alkenes, n-alkylcyclopentanes and n-alkylbenzenes against carbon atom number within the range of C_2 to C_{45}. The difference of the maximum and minimum values for the four types of n-alkyl hydrocarbons with the same carbon atom number is within 5.5~19.4K for boiling points but scattered at 2.8~148K for melting points within the range of C_2 to C_{45}. Furthermore, the difference between the maximum and minimum temperatures of melting points among the structural isomers is much larger than that of boiling points. For example, the temperature difference between the maximum and minimum temperatures of boiling points of n-octane (125.6 ^{o}C) and 2,2,4-trimethylpentane (99.20 ^{o}C) is 26.4 ^{o}C and that of the melting points of 2, 2, 3, 3-tetramethylbutane (100.7 ^{o}C) and 3,3-dimethylhexane (-126.1 ^{o}C) is 226.8 ^{o}C. The changes of melting points are more complex

9

than those of boiling point due to different structures of the isomers. These experimental evidences indicate the melting point estimation may be relatively more difficult than the estimation of boiling point.

5. Conclusion

The modified GVS method has greatly improved the conventional group-contribution methods through combining group contribution with group topological index approach, and has been successfully applied to estimate the normal boiling points and melting points for a wide range and variety of hydrocarbons. The overall average absolute error and absolute percent error between the estimated and experimental boiling points were 4.06K and 0.86 % for 1041 hydrocarbons examined. For the melting points, the average absolute error and absolute percent error for 622 hydrocarbons were 20.33K and 8.56%. Compared with the calculated results of melting points, the boiling points could be estimated more accurately. Finally, the represent model shows a good capability of both in representing the boiling and melting temperature of the hydrocarbons and in distinguishing the temperature differences among the structural isomers of the hydrocarbons in comparison of the previous methods.

List of symbols

a, b, c	constants in Eqs. (1) and (4)
C_{bgi}, C_{bti}	group-independent contribution and group-topological contribution of T_b
C_{mgi}, C_{mti}	group-independent contribution and group-topological contribution of T_m
C_{bi}, C_{mi}	group contribution of T_b and T_m of i-type group in Eqs. (10), (11)
C_{b0}, C_{m0}	T_b correlation constant and T_m correlation constant in Eqs. (12), (13)
C_{b1i}, C_{b2j}, C_{b3k}	first, second and third-order group contributions of T_b in Eqs. (5)
C_{m1i}, C_{m2j}, C_{m3k}	first, second and third -order group contributions of T_m in Eq. (6)
F	objective function value
r	multiple regression coefficient
M	molecular weight
m_{kj}	component of matrix ($k\times$)
N_i, M_j, O_k	number of first-order, second-order and third-order group in Eq. (6)
N	number of data points
n	number of atoms in a molecule (except for hydrogen) in Eq. (4)
T_b, T_m	normal boiling point and melting point
T_{est}, T_{exp}	estimated and experimental value

Greek Letter

$_j$	module of point j vector
v_j	topological index of group j in a molecule

θ a property such as T_b in Eq. (1)

θ limiting value of θ as M

References

[1] M.R. Riazi, T.A. Al-Sahhaf, Fluid Phase Equilibria 117 (1996) 217-224

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

[3] W. Cordes, J. Rarey, Fluid Phase Equilibria 201 (2002) 409-433

[4] Y. Iwai, S. Yamanaga, Y. Arai, Fluid Phase Equilibria 163 (1999) 1-8

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

[6] J. Marrero, R. Gani, Fluid Phase Equilibria 183-184 (2001) 183-208

[7] J. Marerro-Morejon, E. Pardillo-Fontdevila, AIChE J. 45 (1999) 615-621

[8] A.A. Toropov, A.P. Toropova, N.L. Voropaeva, I.N. Ruban, S.Sh. Rashidova, J. Struct. Chem. 40 (1999) 169-172

[9] C.Z. Cao, H. Yuan, J. Chem. Inf. Comput. Sci. 41 (2001) 867-877

[10] Xu W., Yang Q., J. Chem. Eng. Data 47 (2002) 286-288

[11] Xu W., Yang Q., Ind. Eng. Chem. Res. 41 (2002) 5534-5537

[12] D.R. Lide, Editor-in-Chief, CRC Handbook of Chemistry and Physics, 83rd, CRC Press, Edition 2002-2003

[13] B.J. Zwolinski, R.C. Wilhoit, Handbook of Vapor Pressures and Heats of Vaporization of Hydrocarbons and Related Compounds, 1971

[14] J.A. Riddick, W.B. Bunger, T.K. Sakano, Organic Solvents, 4[th]Edn. Wiley Interscience, New York, 1986

Figure captions

Fig. 1. Chemical structures and corresponding matrices for some compounds

Fig. 2. Comparison of boiling points of n-alkanes estimated by the present model with those obtained by previous models

Fig. 3. Comparison of boiling points of 1-alkenes estimated by the present model with those obtained by previous models

Fig. 4. Comparison of boiling points of n-alkylcyclohexanes estimated by the present model with those obtained by previous models

Fig. 5. Comparison of boiling points of n-alkylbenzenes estimated by the present model with those obtained by previous models

Fig. 6. Comparison of melting points of n-alkanes estimated by the present model with those obtained by previous models

Fig. 7. Comparison of melting points of 1-alkenes estimated by the present model with those obtained by previous models

Fig. 8. Comparison of melting points of n-alkylcyclohexanes estimated by the present model with those obtained by previous models

Fig. 9. Comparison of melting points of n-alkylbenzenes estimated by the present model with those obtained by previous models

Fig. 10. Plots of experimental T_b and T_m against carbon number of n-alkyl hydrocarbons

Table 1

Regression results for boiling points and melting points

Compounds	Boiling points			Melting points		
	N	rms [K]	r	N	rms [K]	r
Alkanes	247	3.59	0.9998	110	19.45	0.9641
Alkens / alkynes	296	4.01	0.9979	153	17.50	0.9324
Cyclic	235	4.78	0.9991	133	18.50	0.9589
Aromatic	263	9.51	0.9920	226	39.14	0.8969
Total	1041	5.97	0.9992	622	27.79	0.9499

Table 2

Group-independent and group-topological contributions in estimating boiling and melting points

	Groups	T_b		T_m	
		C_{bgi}	C_{bti}	C_{mgi}	C_{mti}
	-CH$_3$	-16.66	199.02	20.20	88.87
	-CH$_2$-	-9.01	197.54	-4.36	89.22
	>CH-	7.23	142.92	-36.14	109.28
	>C<	60.24	-64.68	60.45	-290.47
Non-ring	=CH$_2$	-15.50	186.34	110.06	-107.38
groups	=CH-	-17.55	221.57	-19.26	158.71
	=C<	-14.36	238.92	-64.20	235.46
	=C=	-52.74	338.99	-457.10	1321.69
	CH	-31.27	219.54	31.96	155.88
	C-	-24.46	256.50	20.75	69.40
	-CH$_2$-	-5.90	187.31	13.93	83.99
	>CH-	1.30	159.18	-40.12	143.62
Ring groups	>C<	21.69	68.84	29.27	-17.64
	=CH-	-8.14	189.02	15.02	55.96
	=C<(c)	-5.48	211.67	-13.48	115.09
	=C<(r)	-5.27	258.29	-17.91	234.74

Table 3

Example calculated for 2-methylpentane whose matrix is obtained from Fig. 1

Group type i	-CH$_3$			-CH$_2$-		>CH-	Note
Group No. j	1	2	3	4	5	6	
Matrix $m(k,j)$	0	2	4	3	2	1	
	2	0	4	3	2	1	
	4	4	0	1	2	3	
$\Sigma_k m^2_{jk}$	20	20	32	19	12	11	$\Sigma = 114$
v_j	0.4189	0.4189	0.5297	0.4082	0.3244	0.3106	
Σv_j	1.3675			0.7327		0.3106	
Group T_b (K)	222.17			126.72		51.62	$n_i C_{bgi} + \Sigma v_j C_{bgi}$
T_b (est, K)	222.17 + 126.72 + 51.62 - 67.68= 332.83						C$_{b0}$ = -67.68
T_b (exp, K)	333.35						Error = -0.52
Group T_m (K)	182.12			56.65		-2.20	$n_i C_{mgi} + \Sigma v_j C_{mgi}$
T_m (est, K)	182.12 + 56.65 - 2.20 - 107.53 = 129.05						C_{m0} = -107.53
T_m (exp, K)	119.45						Error = 9.60

Table 4

Comparison of number of parameters used for T_b estimation of alkanes in several methods

No of parameters	JR	GC$^\text{I}$	GC$^\text{II}$	GIC	GVS	This work
First-groups (or bonding groups)	4	4	4	10	12	8
Second-groups			5			
Model parameters	1	3	1	2	2	1
Total parameters	5	7	10	12	14	9

Table 5

Comparison of boiling points of hydrocarbons estimated by present model with those obtained by previous models

Compounds	Types	Ranges		N	JR		GVS		This work	
					AAE [K]	APE [%]	AAE [K]	APE [%]	AAE [K]	APE [%]
Alkanes	alkanes	C_2	C_{10}	149	8.47	2.28	3.64	0.89	2.90	0.72
	alkanes	C_{11}	C_{100}	98			131.90	14.38	2.59	0.34
	(Total)			247	8.47	2.28	54.53	6.24	2.78	0.57
Unsaturated	alkenes	C_2	C_{22}	199	7.25	1.98	2.44	0.66	2.10	0.55
	dienes	C_3	C_{12}	50	7.93	2.18	4.51	1.21	3.96	1.09
	alkynes	C_2	C_{20}	47	12.28	3.13	3.84	1.03	3.59	1.00
	(Total)			296	8.15	2.19	3.19	0.87	2.82	0.76
Cyclic	cycloalkanes	C_3	C_{42}	179	27.95	4.44	8.32	1.62	3.16	0.73
	cycloalkenes	C_4	C_{10}	56	7.36	1.98	9.07	2.29	5.28	1.35
	(Total)			235	23.05	3.86	8.70	1.82	3.79	0.91
Aromatic	benzenes	C_6	C_{28}	117	17.85	3.21	5.67	1.16	4.39	0.88
	naphthalenes	C_{10}	C_{22}	74	14.79	2.49	9.05	1.58	5.00	0.87
	poly-ring	C_{12}	C_{25}	72	37.44	6.19	14.51	2.37	9.57	1.61
	(Total)			263	21.38	3.67	9.79	1.61	6.89	1.23
Total				1041	15.88	3.08	18.30	2.58	4.06	0.86

Note: $AAE = \sum \left| T_{b,est} - T_{b,\exp} \right| / N$, $APE = \dfrac{\sum \left| T_{b,est} - T_{b,\exp} \right| / T_{b,\exp}}{N} \times 100\%$, where N is the number of data points, $T_{b,est}$ the estimated value of boiling point, $T_{b,exp}$ the experimental value of boiling point

Table 6

Comparison [AAE, %] of boiling points of n-alkyl hydrocarbons estimated by the present model with those obtained by previous models

Compounds	Ranges		N	JR	GC^{I}	GC^{II}	GIC	GVS	This work
n-alkanes	C_2	C_{100}	99		4.50	5.26	19.22	14.15	0.30
1-alkenes	C_3	C_{22}	20	6.02	2.68	2.33	1.14	0.91	0.70
1-alkynes	C_2	C_{20}	14	4.71	2.19	2.43	2.33	1.32	1.25
n-alkylcyclopentanes	C_5	C_{40}	23	7.63	2.10	3.24	2.08	1.41	0.41
n-alkylcyclohexanes	C_6	C_{42}	31	12.86	2.40	4.57	2.79	2.21	0.36
n-alkylbenzenes	C_6	C_{28}	23	6.83	1.92	3.25	1.17	1.24	0.87
Total			210	8.27	3.32	4.25	10.09	7.46	0.49

JR = Group Contributions, Joback and Reid (1987)

GC^{I} = Group Contributions, Cordes and Rarey (2002)

GC^{II} = Group Contributions, Marrero and Gani (2001)

GIC = Group-Interaction Contributions, Marrero and Pardillo (1999)

GVS = Group Vector Space Method, Xu and Yang (2002)

Table 7

Boiling points of C_8H_{18} isomers estimated and compared with the experimentals

Isomers	exp, K	JR	GC^I	GC^{II}	GIC	GVS	This work
n-Octane	398.75	382.64	401.83	398.10	393.84	399.23	400.27
2-Methyheptane	390.75	382.20[a]	392.57[b]	387.04	386.27	390.20	389.29
3-Methyheptane	392.05	382.20[a]	392.57[b]	387.18[c]	386.16[d]	391.96	390.99
4-Methyheptane	390.85	382.20[a]	392.57[b]	387.18[c]	386.16[d]	392.68	391.69
3-Ethylhexane	391.75	382.20[a]	392.57[b]	387.18[c]	386.04	393.27	392.35
2,2-Dimethylhexane	379.95	379.41[a]	383.65[b]	378.52	379.62	381.52	380.58
2,3-Dimethylhexane	388.75	381.76[a]	383.32[b]	388.33[c]	385.67	387.49	384.01
2,4-Dimethylhexane	382.65	381.76[a]	383.32[b]	375.55[c]	378.59	383.98	383.94
2,5-Dimethylhexane	382.25	381.76[a]	383.32[b]	375.40	378.70	379.44	382.08
3,3-Dimethylhexane	385.05	379.41[a]	383.65[b]	378.23[c]	384.96	383.45	385.93
3,4-Dimethylhexane	390.85	381.76[a]	383.32[b]	388.33[c]	385.55[d]	388.45	386.00
2-Methyl-3-ethylpentane	388.75	381.76[a]	383.32[b]	375.55[c]	385.55[d]	388.24	386.00
3-Methyl-3-ethylpentane	391.35	379.41[a]	383.65[b]	378.23[c]	390.30	384.58	389.06
2,2,3-Trimethylpentane	383.15	378.97[a]	374.39[b]	382.98	382.49	377.02	379.96
2,2,4-Trimethylpentane	372.35	378.97[a]	374.39[b]	366.41	372.05	370.29	374.22
2,3,3-Trimethylpentane	387.95	378.97[a]	374.39[b]	382.70	387.94	377.06	382.74
2,3,4-Trimethylpentane	386.65	381.32	374.06	389.47	385.18	385.16	379.20
2,2,3,3-Tetramethylbutane	379.55	376.18	365.46	376.26	381.62	380.53	376.50
AAE ($N = 18$)		6.88	5.11	4.8	2.82	2.46	2.43

Note: [a,b,c,d]-Isomer compounds cannot be distinguished.

Table 8

Comparison of melting points of hydrocarbons estimated by present model with those obtained by previous models

Compounds	Ranges		N	JR		GVS		This work	
				AAE [K]	APE [%]	AAE [K]	APE [%]	AAE [K]	APE [%]
Alkanes	C_2	C_{43}	110	40.41	16.26	31.92	12.89	15.70[a]	7.82[a]
Alkens / alkynes	C_2	C_{22}	153	29.04	15.99	13.98	7.88	14.18	8.43
Cyclic	C_4	C_{42}	133	39.40	17.08	25.57	10.74	14.49	7.50
Aromatic	C_7	C_{34}	226	48.71	16.48	30.39	9.49	30.19	9.62
Total			622	40.41	16.45	25.60	9.96	20.33	8.56

Note: [a] - Estimated deviations while N_{CH3} number of n-alkanes was corrected as (N_{CH3}+2) empirically; they were 23.15K and 10.16% without N_{CH3} number correction.

Table 9

Comparison [AAE, %] of melting points of n-alkyl hydrocarbons estimated by the present model with those obtained by previous models

Compounds	Ranges		N	JR	GC^{II}	GVS	This work
n-alkanes	C_2	C_{43}	34	17.56	16.45	13.15	6.27 [a]
1-alkenes	C_3	C_{22}	17	11.64	11.11	5.85	6.30
1-alkynes	C_2	C_{20}	14	13.40	6.52	5.82	5.50
n-alkylcyclopentanes	C_6	C_{40}	20	19.09	6.88	8.57	7.35
n-alkylcyclohexanes	C_7	C_{42}	27	20.90	5.77	14.71	4.60
n-alkylbenzenes	C_7	C_{34}	27	25.10	4.03	8.57	2.57
Total			139	18.75	8.93	10.27	5.31

Note: [a] - Estimated result while N_{CH3} number of n-alkanes was corrected as (N_{CH3}+2) empirically; it was 13.67% without correction.

Table 10

Comparison of previous methods available from literatures in melting point estimation

Compounds	N	JR		GC^{II}		GVS		Ref.
		AAE [K]	APE [%]	AAE [K]	APE [%]	AAE [K]	APE [%]	
Organic components	388	22.6	11.2					[2]
Organic components	312			14.03	7.23			[5]
Hydrocarbons	339					16.77	7.97	[10]

Heptane

$$\begin{array}{c} \\ c_1 \\ c_2 \end{array}\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left[\begin{array}{ccccccc} 0 & 6 & 5 & 4 & 3 & 2 & 1 \\ 6 & 0 & 1 & 2 & 3 & 4 & 5 \end{array}\right]\end{array}$$

2-Methyl pentane

$$\begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \end{array}\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \left[\begin{array}{cccccc} 0 & 2 & 4 & 3 & 2 & 1 \\ 2 & 0 & 4 & 3 & 2 & 1 \\ 4 & 4 & 0 & 1 & 2 & 3 \end{array}\right]\end{array}$$

2-Ethyl-3-methyl-1-pentene

$$\begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{array}\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \left[\begin{array}{cccccccc} 0 & 3 & 5 & 1 & 4 & 2 & 4 & 3 \\ 3 & 0 & 4 & 2 & 3 & 1 & 3 & 2 \\ 5 & 4 & 0 & 4 & 1 & 3 & 3 & 2 \\ 4 & 3 & 3 & 3 & 2 & 2 & 0 & 1 \end{array}\right]\end{array}$$

Methyl cyclopropane

$$\begin{array}{c} \\ c_1 \\ r_1 \end{array}\begin{array}{cccc} 1 & 2 & 3 & 4 \\ \left[\begin{array}{cccc} 0 & 2 & 2 & 1 \\ 4 & 3 & 3 & 3 \end{array}\right]\end{array}$$

2-Ethyl naphthalene

$$\begin{array}{c} \\ c_1 \\ r_1 \\ r_2 \end{array}\begin{array}{cccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \left[\begin{array}{cccccccccccc} 0 & 1 & 3 & 3 & 4 & 6 & 7 & 6 & 5 & 2 & 4 & 5 \\ 10 & 9 & 7 & 8 & 7 & 6 & 6 & 6 & 6 & 8 & 6 & 6 \\ 8 & 7 & 6 & 6 & 6 & 7 & 8 & 8 & 7 & 6 & 6 & 6 \end{array}\right]\end{array}$$

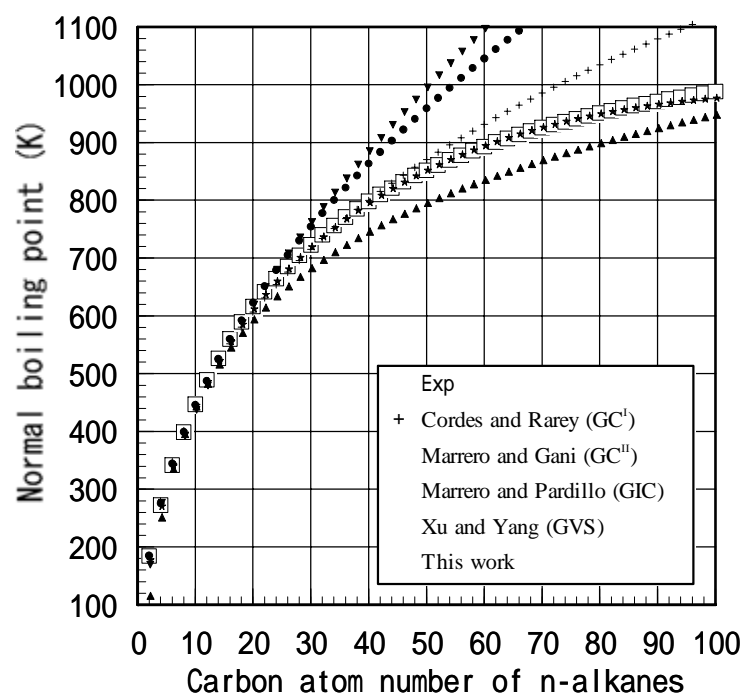**Fig. 1. Chemical structures and corresponding matrices for some compounds**

matrices

Fig. 2. Comparison of boiling points of n-alkanes estimated by the present model with those obtained by previous models
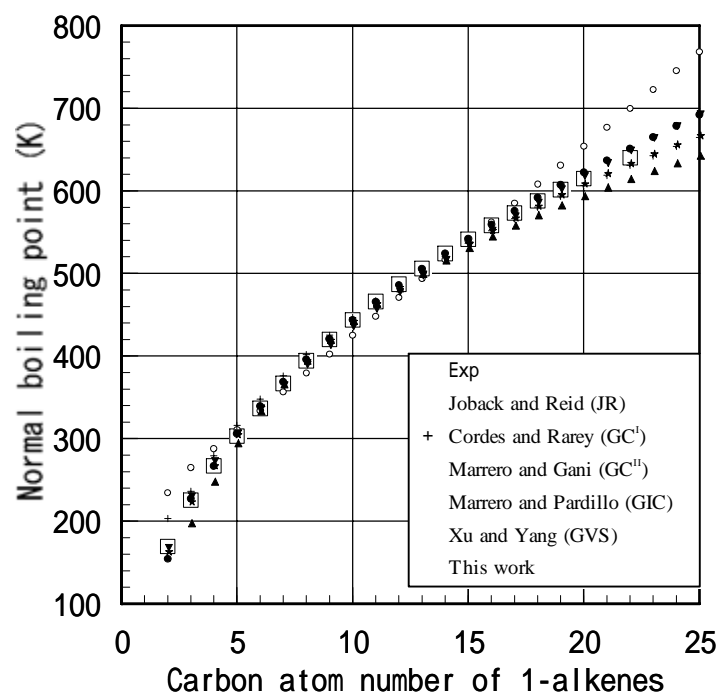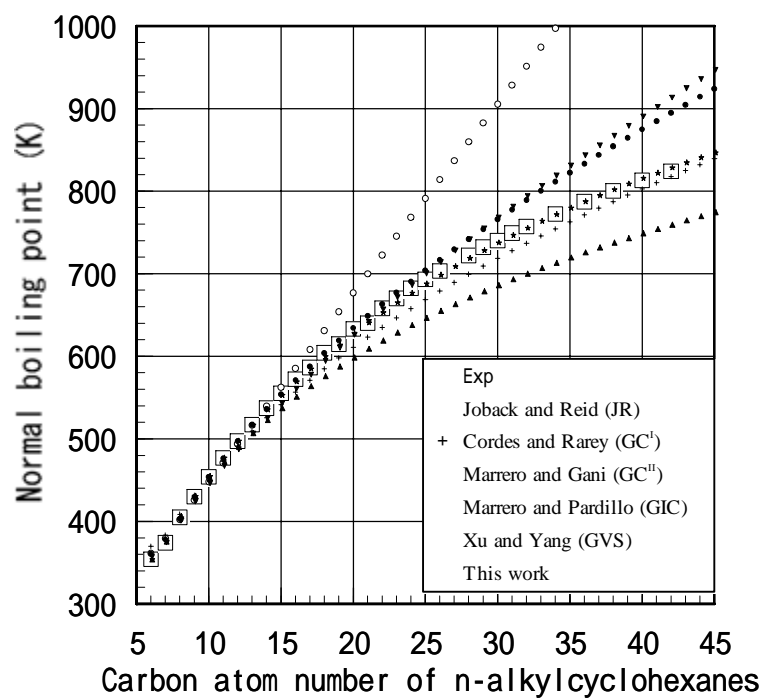
The figure legend contains:
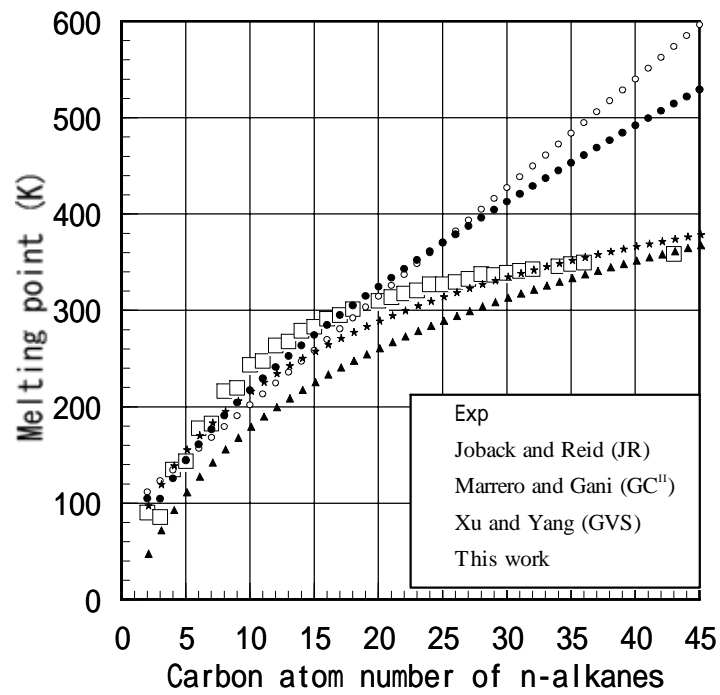
Exp
Cordes and Rarey (GC^I)
Marrero and Gani (GC^{II})
Marrero and Pardillo (GIC)
Xu and Yang (GVS)
This work

The y-axis is labeled "Normal boiling point (K)" ranging from 100 to 1100.
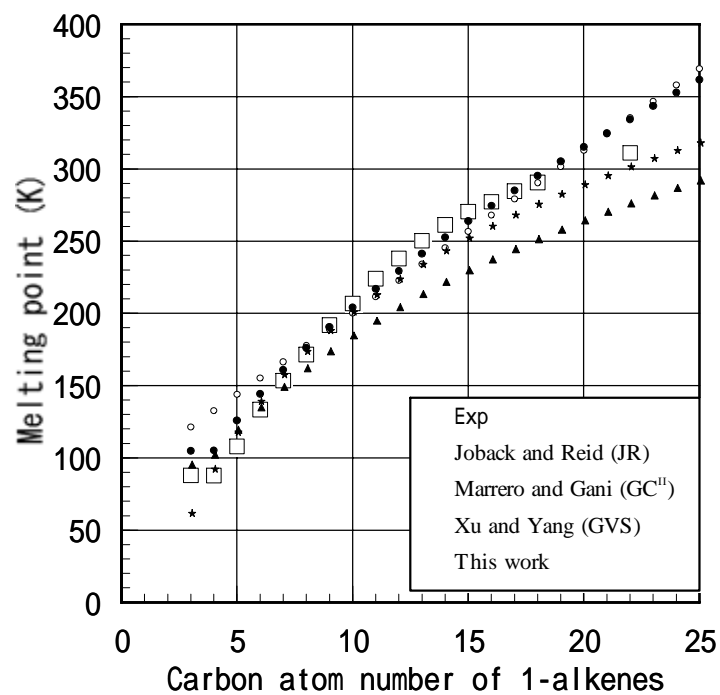The x-axis is labeled "Carbon atom number of n-alkanes" ranging from 0 to 100.

Fig. 3. Comparison of boiling points of 1-alkenes estimated by the present model with those obtained by previous models

Legend (inside plot):

Exp

Joback and Reid (JR)

Cordes and Rarey (GC^I)

Marrero and Gani (GC^{II})

Marrero and Pardillo (GIC)

Xu and Yang (GVS)

This work

Fig. 4. Comparison of boiling points of n-alkylcyclohexanes estimated by the present model with those obtained by previous models
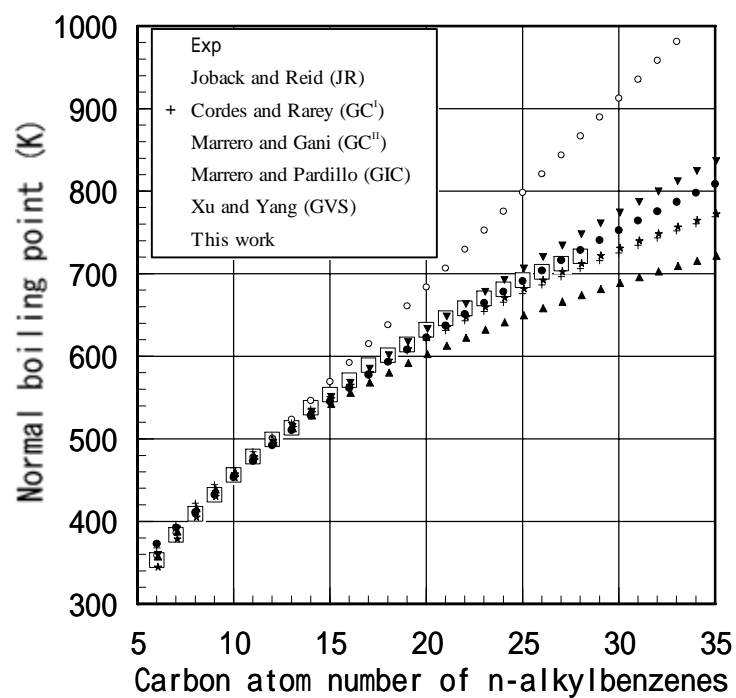
Fig. 5. Comparison of boiling points of n-alkylbenzenes estimated by the present model with those obtained by previous models

Fig. 6. Comparison of melting points of n-alkanes estimated by the present model with those obtained by previous models
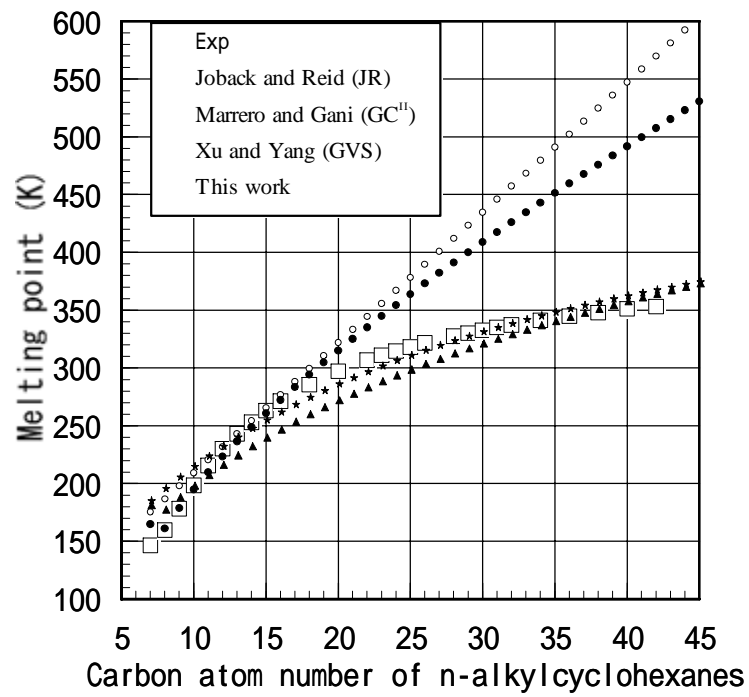
Fig. 7. Comparison of melting points of n-alkenes estimated by the present model with those obtained by previous models

Fig. 8. Comparison of melting points of n-alkylcyclohexanes estimated by the present model with those obtained by previous models
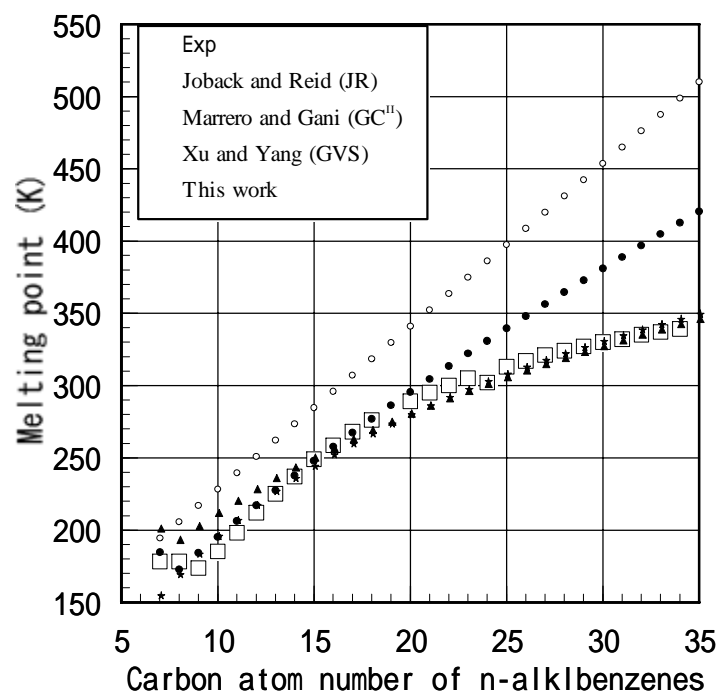
Fig. 9. Comparison of melting points of n-alkylbenzenes estimated by the present model with those obtained by previous models
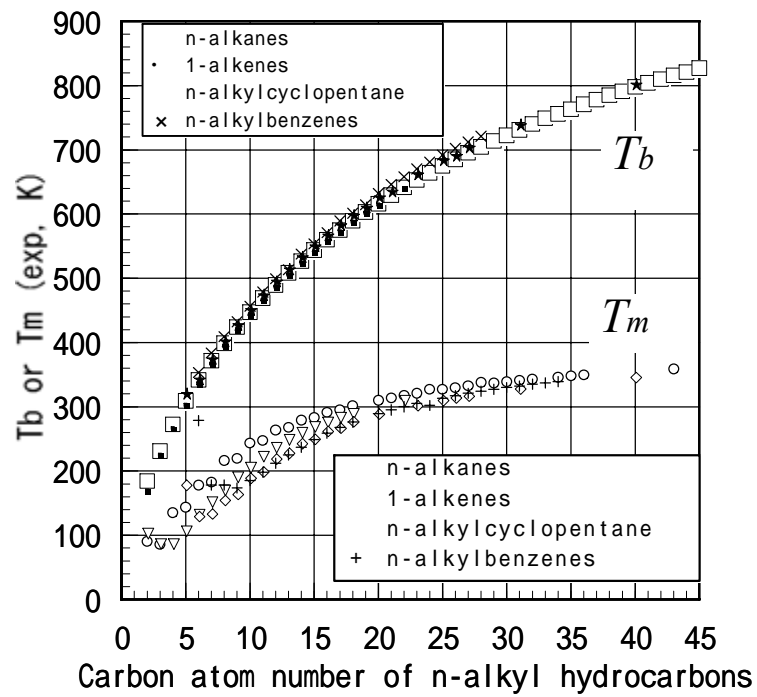
Fig. 10. Plots of experimental T_b and T_m against carbon number of n-alkyl hydrocarbons