

Atomic connectivity group contribution (ACGC) method for the phase transition properties part 1: critical properties

Fangyou Yan¹, Dongdong Cao¹, Xiaojie Feng¹, Jialiang Xiong¹, Qiang Wang¹, Qingzhu Jia¹, and Xia Shuqian²

¹Tianjin University of Science and Technology

²Tianjin University

November 18, 2022

Abstract

In this work, atomic connectivity group contribution (ACGC) method is developed for predicting critical properties of organic compounds. Herein, a new group defining method, namely atomic adjacent group (AAG) method, is proposed to describe the relationship between core atom and its adjacent atoms. For distinguishing isomers effectively, the shape factor (SF) is used to describe the effect of molecular shape on group, and atomic connectivity factors (ACF) are defined for describing the position of each group in a molecule. The external and internal verification methods were utilized during the modelling process. Compared with AAG model, ARE decreased by 6.82-42.57 % when SF was considered and, ARE decreased by 24.19-62.25 % when both SF and ACF were applied as using the ACGC method. Accordingly, SF and ACF are effective in improving the group contribution method and ACGC method is accurate in calculating the properties of organic compounds.

Atomic connectivity group contribution (ACGC) method for the phase transition properties of organic compounds part 1:critical properties

Fangyou Yan^a, Dongdong Cao^a, Xiaojie Feng^a, Jialiang Xiong^a, Qiang Wang^a, Qingzhu Jia^{, b}, Shuqian Xia^{*, c}*

^a School of Material Science and Chemical Engineering, Tianjin University of Science and Technology, Tianjin 300457, P.R.China

^bSchool of Marine and Environmental Science, Tianjin University of Science and Technology, Tianjin300457, P.R. China

^c Key Laboratory for Green Chemical Technology of the State Education Ministry, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China

* To whom correspondence should be addressed.

Email:jiaqingzhu@tust.edu.cn, Tel.: 0086-22-60602695. Fax: 0086-22-60602695;

Email:shuqianxia@tju.edu.cn, Tel.: 0086-22-27405929; Fax: 0086-22-27405929

Abstract:

Group contribution method is one of the main methods for predicting the physicochemical properties of organic compounds, the position of each group is not taken into consideration by classical group contribution method yet. In this work, atomic connectivity group contribution (ACGC) method is developed for predicting critical properties of organic compounds for the first time. Herein, a new group defining method, namely atomic adjacent group (AAG) method, is proposed to describe the relationship between core atom and its

adjacent atoms. For distinguishing isomers effectively, the shape factor (SF) is used to describe the effect of molecular shape on group, and atomic connectivity factors (ACF) are defined for describing the position of each group in a molecule. ACGC models using a general formula are developed for predicting three properties of 710 organic compounds. The external verification and internal verification methods, usually ignored by the traditional group contribution method, were further utilized during the modelling process. Compared with AAG model, ARE decreased by 6.82-42.57 % when SF was considered and, ARE decreased by 24.19-62.25 % when both SF and ACF were applied as using the ACGC method. Accordingly, SF and ACF are effective in improving the group contribution method and ACGC method is accurate in calculating the properties of organic compounds.

KEYWORDS: Atomic connectivity group contribution (ACGC) method; Group contribution (GC) method; Atomic connectivity factors (ACF); Critical properties; Shape factor (SF)

Introduction

The physical and transport properties of compounds play an indispensable role in many chemical engineering applications. Critical properties (critical temperature T_c , critical pressure P_c , critical volume V_c) are important properties of substances in chemistry and chemical engineering. For example, when calculating thermodynamic properties and transfer parameters of matter by corresponding state method, critical properties of matter need to be used. When the corresponding state method or state equation method is used to solve the PVT relation of pure matter and mixtures^{1, 2}, the knowledge of critical properties is needed. Critical properties are essential property parameters for understanding the phase behavior of pure components and mixtures³. Critical properties are also needed in situations where critical phenomena are studied for high-pressure operations such as supercritical extraction^{4, 5} and oil drilling⁶. In view of the importance of critical properties, a great deal of effort has been expended on collecting, sorting and evaluating them bit by bit. Unfortunately, data on experimental critical properties of organic compounds are limited. Experimental measurements of critical properties are sometimes laborious, expensive, and even difficult to measure². On the other hand, with the increase of organic compound data, some experimental data cannot be found from the database, so it becomes crucial to develop mathematical models to provide reasonable estimates of these properties.

As a simple pen-and-paper structure-property relationship (SPR) method⁷⁻⁹, group contribution (GC) method is based on the principle of group addition¹⁰. The group contribution method is widely used to predict various thermodynamic properties of organic compounds, such as critical temperature¹¹⁻²⁰, critical pressure^{13, 21-24}, critical volume^{17, 25} and normal boiling point^{10, 26}. Among GC approaches, Joback²⁷, C-G (Constantinou and Gani)²⁸ and M-G (Marrero and Gani)²⁹ methods are classical GC methods. Joback²⁷ is the first-order group technique and the principal advantages of this method are the simplicity and generality. And yet the combination of first-order groups in one molecule may result in some isomeric molecules and thus the properties of these molecules calculated by first-order group contribution method is the same. The first-order groups are insufficient in describing the portions of the molecular structure and in distinguishing isomers. Based on the first-order group technique, Constantinou and Gani²⁸ developed the second-order group technique, which further considers the influence of the first and second nearest neighbors of the group under consideration. In C-G method, 12 first order level groups and 41 second order level groups were used for about 300 compounds. Average relative error (ARE) is in the range of 0.85-2.89 % for critical properties and decreased by 22-48 % compared with first order level group models. In order to better describe the molecular structure part and distinguish isomers, Marrero and Gani²⁹ developed a group contribution (M-G) method at three levels. In M-G method, 124 first order level groups, 79 second order level groups and 32 third order level groups were used for about 800 compounds with critical properties. The use of third level provided more structural information about compounds and improved the accuracy and applicability of GC method. For example, ARE for critical properties decreased by 5-12 % compared with second order level group models. Hukkerikar et. al.³⁰ presented revised model parameters for M-G method (M-G⁺) at three levels for 18 pure component properties with about 12000 compounds. 130 first order level groups, 90 second order level groups and 31 third order level groups were used for about 900 compounds with critical

properties. The critical properties of ARE are higher than MG method.

The hypothesis of classical GC method is that a group has an equal contribution value at any molecule. The position of each group plays an important role in the properties of compounds, while it is also not considered by the traditional group contribution methods. In addition, no uniform rule for group division was developed in the above classical GC methods, which also brought inconvenience to the division of complex compounds. Also, traditional group contribution methods usually do not carry out external verification and internal verification, so the stability of the model is not verified. Sun and Sahinidis¹³ proposed a method to identify functional groups in molecular structures and to verify the stability of the model by dividing the training and test sets. A total of 74 functional groups were used to describe the critical properties of about 860 organic compounds. The R^2 of critical temperature, critical pressure and critical volume are 0.96, 0.95 and 0.99 respectively, and the ARE are 2.02 %, 4.09 % and 3.10 % respectively.

The aim of proposed atomic connectivity group contribution (ACGC) method is to establish a general and simple group contribution method for calculating phase transition properties of organic compounds. A general formula containing group, shape factor and localization factor is used to predict the different properties of organic compounds. The present work is completed for the calculation of the critical properties. Based on the relationship between core atom and its adjacent atoms, atomic adjacent group (AAG) with unified dividing rules is defined. In order to describe the influence of the molecular shape of compounds such as branched chains and rings on the calculation of properties, the shape factor (SF) is defined. The atomic connectivity factors (ACF) are defined for describing the position of each group. The predictive ability and stability of the model are verified by external and internal validation.

Methodology

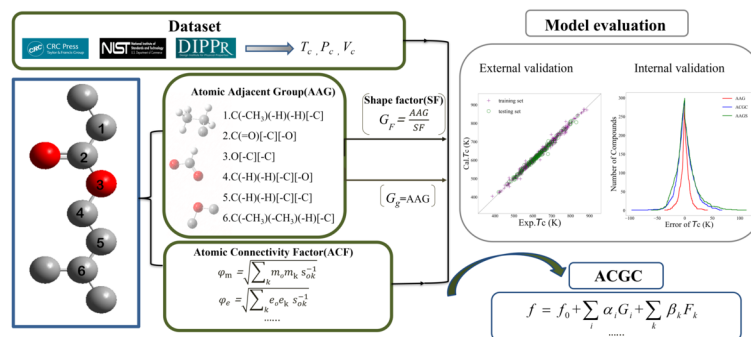


Figure 1 Flow chart of ACGC.

ACGC is composed of atomic adjacent group (AAG), shape factor (SF) and atomic connectivity factor (ACF). The working process of ACGC is described in Figure 1. AAG is a systematic group definition approach that explicitly decomposes each molecular structure into a set of non-overlapping functional groups based on the relationship between core and adjacent atoms. SF is used to calculate the effect of molecular shape. ACF is used to calculate topological position factors by atomic properties to describe the positions of groups. We also analyze and evaluate the model by external verification, internal verification and Y-randomization test³¹.

Atomic adjacent group (AAG)

The traditional group contribution method requires a higher level of groups to make a more accurate division, and the division method is complicated. The AAG is proposed by the atomic adjacent relationship. Atoms

are classified into two types: endpoint atom and connection atom. The atom connected with only one non-hydrogen atom is defined as endpoint atom. The atom connected with two or more non-hydrogen atoms is defined as connection atom. A group consists of core atom, adjacent atoms and bond types, which are shown in Figure 2. Core atom is a connection atom, which neighbors with two or more non-hydrogen atoms. Core atoms include carbon, oxygen, nitrogen, silicon, sulfur, phosphorus and so on. Atoms connected to core atom include endpoint atoms and connection atoms. Endpoint atom is described in the parenthesis, '()'. Connection atom is described in the bracket, '[]'. Bond types between the core atom and adjacent atom are presented before the adjacent atoms. The single bond, double bond, triple bond of linear structure is described as -, =, [=], respectively. The single bond, double bond, triple bond of cyclic structure and aromatic bond are described by ~, [=], [=] and [=], respectively. Four examples are used to describe the group definition rules in Table S1 of Supporting Information (contribution-coefficient.docx).

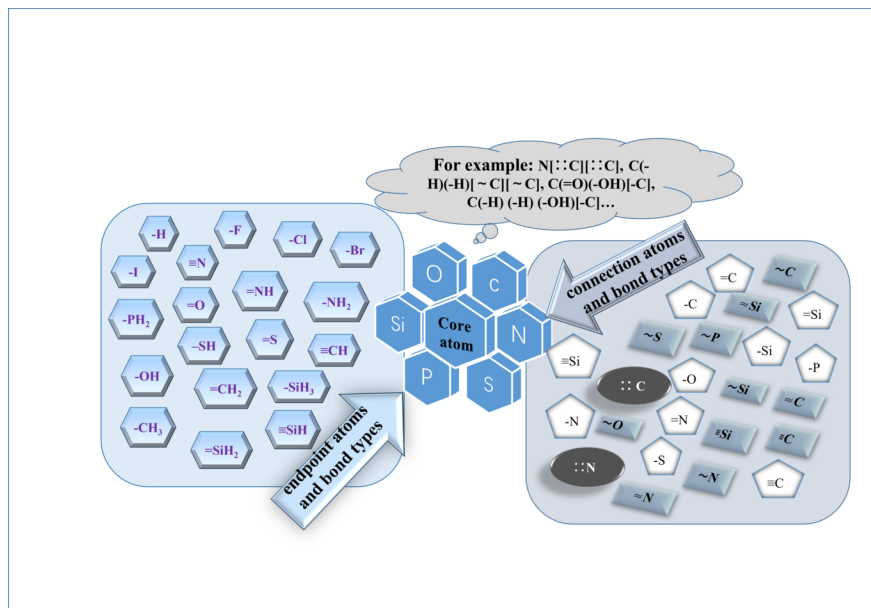


Figure 2 The core atom, adjacent atoms and bond type in atomic adjacent group.

2.2 Shape factor (SF)

In order to make the contribution value of the group more accurate, we introduce the shape factor to describe the influence of the molecular shape of branched chains, straight chains, and rings on the properties. The formula of the shape factor is shown in Equation (1) and (2).

where n_H is the number of non-hydrogen atoms; s_m is the maximum step; s_{ok} is the steps between core atom o and other non-hydrogen atom k .

2.3 Atomic connectivity factor (ACF)

The hypothesis of group contribution method is that a group has an equal contribution value at any molecule. The contribution of group position is not considered by the traditional group contribution methods. Therefore, isomers resulted from group position cannot be distinguished even if multiple level groups were used. Atomic connectivity factor (ACF) is defined as φ_k to describe the contribution of group position on the properties. ACF of a group is defined by the connection of core atom (o) with other non-hydrogen atoms (k) and the properties of each atom such as atomic weight (m), electronegativity (e), ionization energy (i), atomic radius (r), number of outermost electrons (o), number of electrons shell (s), atom on aromatic ring (a) and branched degree (b). The location factor was calculated from ACF. The formula for ACF is defined

as Eq.(3). Examples are used to describe ACF calculation in Supporting Information (Example.xlsx).

where s_{ok} is the steps between core atom o and other non-hydrogen atom k ; x stands for m, e, i, r, o, s, a , b (b is 1 if the atom is in the aromatic ring and 0 if it is not.).

2.4 ACGC method

The ACGC model is combined of group contributions and position factors contributions. Formula (4) is used for T_c, V_c . Formula (5) is used for P_c .

a_i is the contribution coefficient of group; β_κ is the contribution coefficient of the atomic connectivity factor. a_i and β_κ are shown in Table S2~S4 of Supporting Information (contribution-coefficient.docx). G_i is G_f or G_g . G_g is the group number and G_f is the group number divided by SF. The value of the localization factor (φ_κ) for each group is expressed by the calculated value of one of the formulas for the eight atomic connectivity factors.

2.5 Data set

The quality of a method depends largely on the quantity and quality of experimental data used in its development. The data for developing atomic connectivity group contribution method are from National Institute of Standards and Technology (NIST)³², CRC Handbook of Chemistry and Physics³³ and DIPPR801³⁴, and only the experimental data are used for ACGC modelling work. We take the median value of repeated experimental test. And compounds with more than three carbon and silicon atoms combined were selected, because small molecules may affect the stability of the model, even though they can improve the accuracy of the model. If a group appears only in one or two compounds, the error for these compounds is small, while these groups are unreliable for predicting. In order to develop reliable model, the compounds with a group appearing in less than five compounds will be deleted. The data set is composed of 680 critical temperature (T_c), 481 critical pressure (P_c), 471 critical volume (V_c).

Results and discussions

The results for T_c, P_c and V_c by ACGC methods are shown in Table 1. Statistical parameters such as square correlation coefficient (R^2), average absolute error (AAE), average relative error (ARE) were used to evaluate the ACGC model. R^2 is 0.9891, 0.9887, 0.9983 for T_c, P_c and V_c respectively. Leave-one-out cross-validation (LOO-CV) was applied for evaluating the predicting ability of ACGC models. The quality of ACGC method was then decided by the squared correlation coefficient of LOO-CV (Q^2). In AAG method, only the atomic adjacent group is used as traditional GC method. Compared with AAG and AAGS, the R^2 of ACGC model was improved. The R^2 and ARE for the AAG models are shown in Table 1. The ARE of T_c, P_c and V_c decreased by 42.57 %, 13.23 % and 6.82 %, respectively, when SF is applied. When ACF is applied to ACGC model, compared with AAGS model, it decreases by 34.27 %, 12.64 % and 34.63 %, respectively. Therefore, ACF is also effective in improving ACGC. ARE decreased by 24.19-62.25 % and AAE decreased by 19.30-61.54 % when SF and ACF are applied, therefore, SF and ACF are effective in improving ACGC method.

Table 1 The results for critical properties by ACGC method.

| Method | Parameter | T_c (K) | P_c (MPa) | V_c (L/kmol) |
|--------------------|-------------------------|-----------|-------------|----------------|
| N _{O.} of | compounds | 680 | 481 | 471 |
| | G_i | 48 | 40 | 39 |
| | φ_κ | 22 | 8 | 9 |
| AAG | R^2 | 0.9319 | 0.9830 | 0.9958 |
| | AAE | 15.2451 | 0.0824 | 0.0104 |
| | ARE % | 2.49 | 3.10 | 2.20 |
| | Q^2 | 0.9132 | 0.9762 | 0.9948 |
| | R^2_{training} | 0.9310 | 0.9841 | 0.9959 |

| Method | Parameter | T_c (K) | P_c (MPa) | V_c (L/kmol) |
|--------|-------------------------|-----------|-------------|----------------|
| AAGS | AAE_{training} | 15.3473 | 0.0810 | 0.0103 |
| | R^2_{test} | 0.9354 | 0.9797 | 0.9955 |
| | AAE_{test} | 14.8366 | 0.0879 | 0.0109 |
| | R^2 | 0.9766 | 0.9860 | 0.9963 |
| | AAE | 8.8661 | 0.0745 | 0.0097 |
| | $ARE\%$ | 1.43 | 2.69 | 2.05 |
| | Q^2 | 0.9670 | 0.9787 | 0.9954 |
| | R^2_{training} | 0.9762 | 0.9860 | 0.9963 |
| ACGC | AAE_{training} | 8.8698 | 0.0752 | 0.0098 |
| | R^2_{test} | 0.9797 | 0.9866 | 0.9963 |
| | AAE_{test} | 8.8515 | 0.0717 | 0.0091 |
| | R^2 | 0.9891 | 0.9887 | 0.9983 |
| | AAE | 5.8634 | 0.0665 | 0.0063 |
| | $ARE\%$ | 0.94 | 2.35 | 1.34 |
| | Q^2 | 0.9826 | 0.9820 | 0.9978 |
| | R^2_{training} | 0.9891 | 0.9887 | 0.9983 |
| | AAE_{training} | 5.8520 | 0.0665 | 0.0064 |
| | R^2_{test} | 0.9898 | 0.9886 | 0.9983 |
| | AAE_{test} | 5.9090 | 0.0665 | 0.0061 |

3.1 External validation

External validation could well reflect the predictive performance of ACGC model. Figure 3 expressed the scatter diagram of experimental and calculated values for the training set and the testing set. As was shown in Figure 3, most data points are closely distributed along the diagonal. The data points were well fitted both for the training set and the testing set for the model. In addition, the R^2_{training} and R^2_{test} values were 0.9891 and 0.9898 for T_c , 0.9887 and 0.9886 for P_c , 0.9983 and 0.9983 for V_c , respectively. The small difference between these two metrics clearly demonstrates the reliability of the current model. The high correlation coefficient of the testing set demonstrated the good prediction ability of the present models.

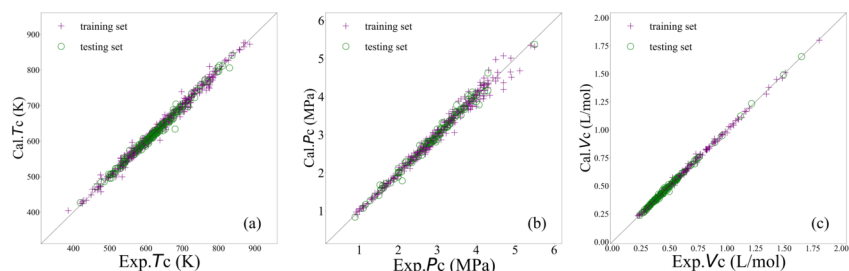


Figure 3 External validation results of the experimental vs. the calculated values for ACGC models of critical properties: (a) T_c , (b) P_c and (c) V_c .

3.2 Internal validation

Figure 4 shows the error distribution of application model and LOO-CV method. The results show that the error distribution of LOO-CV model is consistent and completely similar to the corresponding T_c , P_c , V_c model, which further demonstrates the stability and robustness of these models based on groups, shape factor and topological location factors. In addition, Figure 5 showed the relationship between the

experimental values and LOO-CV predicted values. Results showed that most of data points were distributed compactly around the diagonal. High Q^2 values for the model indicated the stability of our models, which were 0.9826, 0.9820 and 0.9978 for T_c , P_c and V_c , respectively.

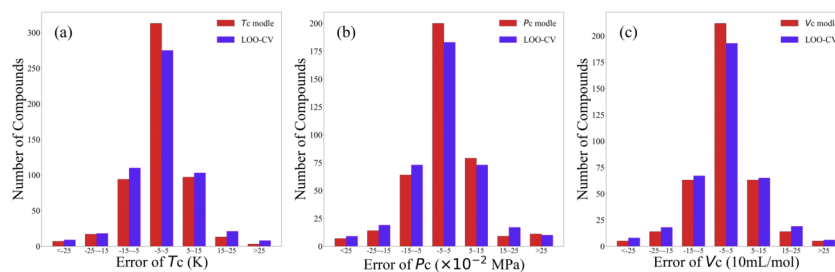


Figure 4 The error distributions of the models and LOO-CV for ACGC models of critical properties: (a) T_c , (b) P_c and (c) V_c .

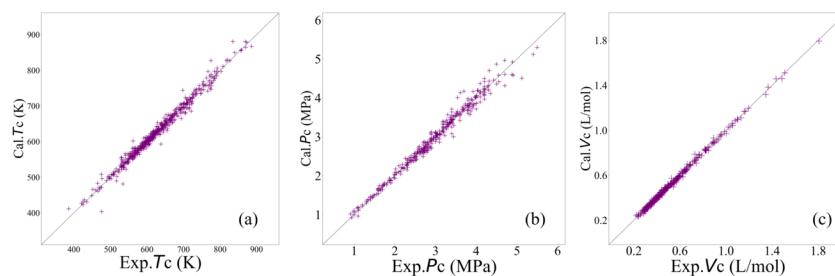


Figure 5 The experimental vs. the calculated (LOO-CV) values for ACGC models of critical properties: (a) T_c , (b) P_c and (c) V_c .

3.3 Y-randomized analysis

We repeated 1000 Y-randomized validation trials to ensure the reliability of the results. As was shown in Figure 6, The mean R^2_Y and Q^2_Y of 1000 Y-random tests of critical properties model were far lower than R^2 and Q^2 of the original model. The comparison of the mean R^2_Y and Q^2_Y of Y-random test is shown in Figure 6. This indicates that our model has no chance correlation in the modeling process.

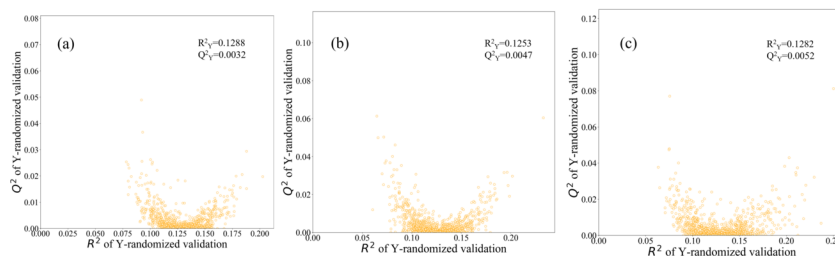


Figure 6 Y-randomized validation of the present models for ACGC models of critical properties: (a) T_c , (b) P_c and (c) V_c .

3.4 Shape factor analysis

The comparison of AAE between AAG and AAGS shown in Table 1 can well reflect that the results of the model have been improved after we introduce the shape factor. The overall AAE of T_c , P_c and V_c were 15.2451 K, 0.0824 MPa and 0.0104 L/kmol before the shape factor was added. The results were reduced by 41.84 %, 9.59 % and 6.73 % respectively after the introduction. The results show that the shape factor can effectively distinguish the error of compound shape calculation. The R^2 of the AAGS model with critical temperature is 0.0477 higher than that of the AAG model. Since R^2 of AAG model with critical volume is very high, the improvement effect of introducing shape factor is not obvious. Moreover, the AAE of the test set was reduced by 16.51-40.34 %, which proves that the introduction of the shape factor is effective compared with the AAG method.

3.5 Distinction of isomers

Figure 7 shows the error distribution of ACGC, AAG and AAGS methods. The results show that the error of ACGC is in a smaller range than that of AAG and AAGS methods for the three critical properties in Figure 7, which further demonstrates that ACF can significantly improve the calculation performance of ACGC method for the critical properties. Isomers owing to group position are analyzed in Figure 8 for critical properties, which shows that the AAE of ACGC are much smaller than these of AAG for most of isomers, especially for these with large AAE. The points in the figure are the average values of AAE obtained for molecules with the same corresponding groups. The average of AAE for all isomers of AAG model and ACGC model with critical temperature, critical pressure and critical volume are 10.2728 and 4.8844 K, 0.0809 and 0.0680 MPa, 0.0058 and 0.0040 L/kmol, respectively. After the introduction of SF and ACF, the AAE of T_c , P_c and V_c of all isomers decreased by 52.45%, 15.95% and 31.03%, respectively. This result shows that SF and ACF can effectively distinguish isomers.

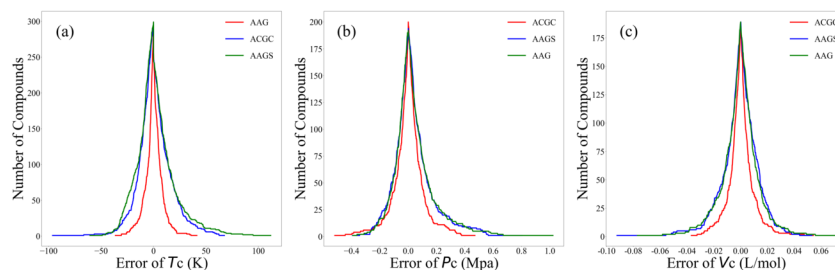


Figure 7 The error distributions of the models and AAG for critical properties: (a) T_c , (b) P_c and (c) V_c .

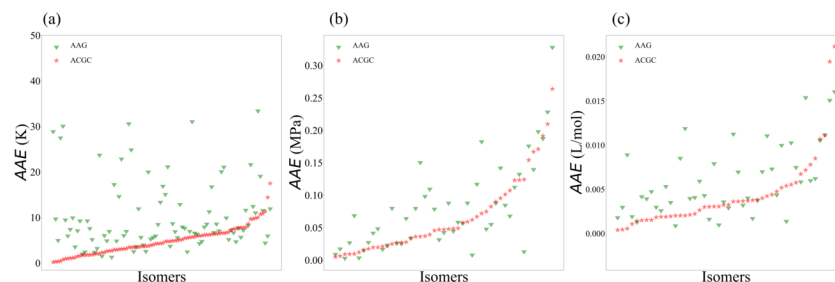


Figure 8 AAE distribution of isomers for critical properties: (a) T_c , (b) P_c and (c) V_c .

3.6 Comparison results with existing methods

In Table 2, the ACGC models are compared with the classical group contribution methods (Joback: first-order group technique, C-G: second-order group technique, M-G: third-order group technique, M-G⁺: third-order group technique) and S-S (Sun and Sahinidis: third-order group technique). The groups in Joback method describe part of atomic adjacent relationship, and the AAG method is more refined than the Joback method, which is theoretically more accurate than the Joback method. And the groups we have divided are guaranteed to appear in more than five compounds, thus the ARE for the critical temperature looks higher than the Joback method. The ARE of V_c in ACGC model was 1.34%, which was lower than that of Joback (2.30 %), C-G (1.79 %), M-G (1.80 %), M-G⁺ (2.05 %) and S-S (3.10 %). These results show that the ACGC models are accurate. The ARE of critical properties calculated by ACGC method is lower than M-G⁺ method and S-S method. Although ARE of the T_c ACGC model is higher than that of the C-G and M-G methods, we screened the data so that the number of occurrences of groups in the data set is not less than 5 times, so the model is more reliable. In addition, this work aims to develop reliable models with good predictive ability, using 20 % of the data set as test sets and the rest as training sets to fit the property model, which is different from the traditional group contribution method.

Table 2 Comparison of ACGC with GC methods.

| Property | Parameter | Joback ²⁷ | C-G ²⁸ | M-G ²⁹ | M-G ⁺ ³⁰ | S-S ¹³ | AAG | AAGS | ACGC |
|----------------|-----------|----------------------|-------------------|-------------------|--------------------------------|-------------------|------|------|------|
| T_c (K) | Samples | 409 | 285 | 783 | 858 | 860 | 680 | 680 | 680 |
| | Variables | 40 | 53 | 232 | 251 | 73 | 48 | 48 | 70 |
| | ARE% | 0.81 | 0.85 | 0.80 | 1.23 | 2.02 | 2.49 | 1.43 | 0.94 |
| P_c (MPa) | Samples | 392 | 269 | 775 | 852 | 852 | 481 | 481 | 481 |
| | Variables | 40 | 52 | 235 | 242 | 73 | 40 | 40 | 48 |
| | ARE% | 5.20 | 2.89 | 2.30 | 3.90 | 4.09 | 3.10 | 2.69 | 2.35 |
| V_c (L/kmol) | Samples | 310 | 251 | 762 | 797 | 852 | 471 | 471 | 471 |
| | Variables | 40 | 51 | 234 | 230 | 74 | 39 | 39 | 48 |
| | ARE% | 2.30 | 1.79 | 1.80 | 2.05 | 3.10 | 2.20 | 2.05 | 1.34 |

Conclusions

Combining the atom connection relationship and group contribution method, atomic connectivity group contribution (ACGC) method is proposed. Based on the relationship between core atom and adjacent atoms, a new group defining method is developed. To make the model more accurate, the shape factor is defined. Atomic connectivity factors are defined to describe the position of each group in a molecule based on the relationship between the core atom and other non-hydrogen atoms. Together with the shape factors, most organic isomers can be distinguished. The R^2 of T_c , P_c and V_c were 0.9891, 0.9887 and 0.9983, respectively. The ARE of T_c , P_c and V_c were 0.94%, 2.35% and 1.34 % respectively. ARE of ACGC models with critical property are smaller than these of AAG and AAGS, which shows that SF and ACF significantly improve the performance of ACGC method. The ACGC model has high accuracy compared to the classical GC model. From the above comparison, we can conclude that ACGC method is accurate and reliable. We screened the data to ensure that each group is present in more than five compounds. SF and ACF can distinguish isomers effectively. Using the same formula, ACGC can accurately calculate the critical properties of organic compounds, which greatly facilitates the prediction and collection of properties of undetectable organic compounds.

Conflict of interest

The authors confirm that this article has no conflicts of interest.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (22278319, 21808167), Tianjin Municipal Science and Technology Bureau (20JCQNJC00090).

Data Availability Statement

Data available on request from the authors. The data that support the findings of this study are available from the corresponding author upon reasonable request. The implementation of the ACGC method can be implemented by relevant mathematical software (matlab and mathematica) as well as computer language (e.g. python, R and C++).

References

1. An, H.; Yang, W. M.; Maghbouli, A.; Chou, S. K.; Chua, K. J., Detailed physical properties prediction of pure methyl esters for biodiesel combustion modeling. *APPLIED ENERGY* **2013**, *102*, 647-656.
2. Evangelista, N. S.; do Carmo, F. R.; de Sant'Ana, H. B., Estimation of Physical Constants of Biodiesel-Related Fatty Acid Alkyl Esters: Normal Boiling Point, Critical Temperature, Critical Pressure, and Acentric Factor. *Industrial & Engineering Chemistry Research* **2018**, *57* (25), 8552-8565.
3. Do Carmo, F. R.; Evangelista, N. S.; Fernandes, F. A. N.; De Sant'Ana, H. B., Evaluation of Optimal Methods for Critical Properties and Acentric Factor of Biodiesel Compounds with Their Application on Soave-Redlich-Kwong and Peng-Robinson Equations of State. *Journal of Chemical and Engineering Data* **2015**, *60* (11), 3358-3381.
4. Jahromi, S. A.; Roosta, A., Estimation of critical point, vapor pressure and heat of sublimation of pharmaceuticals and their solubility in supercritical carbon dioxide. *FLUID PHASE EQUILIBRIA* **2019**, *488*, 1-8.
5. Zhang, J. C.; Wu, X. Y.; Cao, W. L., Study on critical properties for CO₂+cosolvent binary system and ternary system. *CHINESE JOURNAL OF CHEMICAL ENGINEERING* **2002**, *10* (2), 223-227.
6. Street, C. G.; Guigard, S. E., Treatment of Oil-Based Drilling Waste Using Supercritical Carbon Dioxide. *JOURNAL OF CANADIAN PETROLEUM TECHNOLOGY* **2009**, *48* (6), 26-29.
7. Su, Y.; Wang, Z.; Jin, S.; Shen, W.; Ren, J.; Eden, M. R., An architecture of deep learning in QSPR modeling for the prediction of critical properties using molecular signatures. *AIChE Journal* **2019**, *65* (9).
8. Shi, Y.; Yu, M.; Liu, J.; Yan, F.; Luo, Z.-H.; Zhou, Y.-N., Quantitative Structure–Property Relationship Model for Predicting the Propagation Rate Coefficient in Free-Radical Polymerization. *Macromolecules* **2022**, *55* (21), 9397-9410.
9. Jia, Q.; Cui, X.; Li, L.; Wang, Q.; Liu, Y.; Xia, S.; Ma, P., Quantitative Structure-Activity Relationship for High Affinity 5-HT_{1A} Receptor Ligands Based on Norm Indexes. *J Phys Chem B* **2015**, *119* (51), 15561-7.
10. Ghasemitabar, H.; Movagharnejad, K., Estimation of the normal boiling point of organic compounds via a new group contribution method. *Fluid Phase Equilibria* **2016**, *411*, 13-23.
11. Yan, X.; Dong, Q.; Hong, X., Reliability analysis of group-contribution methods in predicting critical temperatures of organic compounds. *Journal of Chemical and Engineering Data* **2003**, *48* (2), 374-380.
12. Qiang, W.; Peisheng, M.; Qingzhu, J.; Shuqian, X., Position group contribution method for the prediction of critical temperatures of organic compounds. *Journal of Chemical and Engineering Data* **2008**, *53* (5), 1103-1109.
13. Sun, Y.; Sahinidis, N. V., A New Functional Group Selection Method for Group Contribution Models and Its Application in the Design of Electronics Cooling Fluids. *Industrial & Engineering Chemistry Research* **2021**, *60* (19), 7291-7300.
14. Li, Z.; Zuo, L.; Wu, W.; Chen, L., The New Method for Correlation and Prediction of Thermophysical Properties of Fluids. Critical Temperature. *Journal of Chemical & Engineering Data* **2017**, *62* (11), 3723-3731.

15. Valderrama, J. O.; Forero, L. A.; Rojas, R. E., Extension of a group contribution method to estimate the critical properties of ionic liquids of high molecular mass. *Industrial and Engineering Chemistry Research* **2015**, *54* (13), 3480-3487.
16. Zbogar, A.; Lopes, F. V. D. S.; Kontogeorgis, G. M., Approach suitable for screening estimation methods for critical properties of heavy compounds. *Industrial and Engineering Chemistry Research* **2006**, *45* (1), 476-480.
17. Nannoolal, Y.; Rarey, J.; Ramjugernath, D., Estimation of pure component properties: Part 2. Estimation of critical property data by group contribution. *Fluid Phase Equilibria* **2007**, *252* (1), 1-27.
18. Espinosa, G.; Yaffe, D.; Arenas, A.; Cohen, Y.; Giralt, F., A Fuzzy ARTMAP-Based Quantitative Structure-Property Relationship (QSPR) for Predicting Physical Properties of Organic Compounds. *Industrial & Engineering Chemistry Research* **2001**, *40* (12), 2757-2766.
19. Wen, X.; Qiang, Y., A new group contribution method for estimating critical properties of organic compounds. *Industrial and Engineering Chemistry Research* **2001**, *40* (26), 6245-6250.
20. Daubert, T. E.; Bartakovits, R., Prediction of critical temperature and pressure of organic compounds by group contribution. *Industrial and Engineering Chemistry Research* **1989**, *28* (5), 638-641.
21. Wang, Q.; Jia, Q.; Ma, P., Position group contribution method for the prediction of critical pressure of organic compounds. *Journal of Chemical and Engineering Data* **2008**, *53* (8), 1877-1885.
22. Otobrise, C.; Osondu Monago, K., New group contribution method for predicting the critical pressures of fatty acids and triglycerides. *International Journal of Computational and Theoretical Chemistry* **2015**, *3* (6).
23. Nikitin, E. D.; Popov, A. P.; Bogatishcheva, N. S.; Faizullin, M. Z., Critical Temperatures, Pressures, Heat Capacities, and Thermal Diffusivities of C3 to C6 Dinitriles. *Journal of Chemical & Engineering Data* **2022**, *67* (4), 836-845.
24. Poling, B. E.; Prausnitz, J. M.; O'Connell, J. P., The Properties of Gases and Liquids, 5th ed. *McGraw-Hill Education: New York* **2001**.
25. Nannoolal, Y.; Rarey, J.; Ramjugernath, D., Estimation of pure component properties. *Fluid Phase Equilibria* **2007**, *252* (1-2), 1-27.
26. Sanghvi, R.; Yalkowsky, S. H., Estimation of the normal boiling point of organic compounds. *INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH* **2006**, *45* (8), 2856-2861.
27. Joback, K. G. A unified approach to physical property estimation using multivariate statistical techniques. Massachusetts Institute of Technology, 1984.
28. Constantinou, L.; Gani, R., New group contribution method for estimating properties of pure compounds. *AIChE Journal* **1994**, *40* (10), 1697-1710.
29. Marrero, J.; Gani, R., Group-contribution based estimation of pure component properties. *Fluid Phase Equilib.* **2001**, *183-184*, 183-208.
30. Hukkerikar, A. S.; Sarup, B.; Ten Kate, A.; Abildskov, J.; Sin, G.; Gani, R., Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria* **2012**, *321*, 25-43.
31. Rücker, C.; Rücker, G.; Meringer, M., y-Randomization and Its Variants in QSPR/QSAR. *Journal of Chemical Information and Modeling* **2007**, *47* (6), 2345-2357.
32. National Institute of Standards and Technology (NIST). <https://webbook.nist.gov/chemistry/> (accessed **2022**) .

33. Haynes, W. M., *CRC handbook of chemistry and physics* . CRC press: 2014.
34. Rowley, R.; Wilding, W.; Oscarson, J.; Zundel, N.; Marshall, T.; Daubert, T.; Danner, R., DIPPR® Data Compilation of Pure Compound Properties, Design. *Institute for Physical Properties, AIChE, New York* **2002** .