

Research Article

On the Analysis and Assessment of First-Order Group Contribution Models for the Calculation of Normal Boiling Point and Critical Properties of Pure Compounds

Vanessa Villazón-León , Adrián Bonilla-Petriciolet , and Juan Carlos Tapia-Picazo

Instituto Tecnológico de Aguascalientes, Aguascalientes 20256, Mexico

Correspondence should be addressed to Adrián Bonilla-Petriciolet; adrian.bp@aguascalientes.tecnm.mx

Received 5 May 2022; Revised 11 July 2022; Accepted 14 July 2022; Published 27 August 2022

Academic Editor: Pratima Meshram

Copyright © 2022 Vanessa Villazón-León et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The thermodynamic properties of pure compounds are relevant data for process systems engineering. Different first-order group contribution models have been reported in the literature to calculate these properties and they are also widely employed in commercial process simulators. However, they may have some limitations and, consequently, a reliable comparison of these models is required to analyze their performance and to determine the best alternative for the calculation of pure compound properties. This paper reports the implementation and evaluation of several first-order group contribution models to calculate the normal boiling point and critical properties (temperature, pressure, and volume) of pure compounds. The performance of these models was characterized and compared for several compound families using a standardized approach to determine their group contributions and parameters. An artificial neural network model was also applied and assessed to improve the estimations obtained with the best group contribution models. Results showed that the calculation of critical temperature was challenging for several compound families where *AARD* values ranged from 0.05 to 56.28%, while the group contribution models were more accurate to estimate the critical volume with *AARD* values ranging from 0.48 to 35.99%. This study allows us to identify the limitations and gaps of this type of thermodynamic models with the objective of improving its performance for the calculation of pure compound thermodynamic properties. The findings of this study can help to enhance the capabilities of thermodynamic models for the calculation of the normal boiling point and critical properties of pure compounds, which are relevant for the process systems engineering of new operations and products.

1. Introduction

Process design, simulation, control, and optimization require the knowledge of the thermodynamic properties of pure compounds and their mixtures. For example, the critical properties are employed as input parameters to predict the volumetric and phase equilibria behavior of pure compounds and mixtures using cubic EoS in process simulators [1]. These properties establish limits to the operating conditions of equipment and processes. The characterization of the critical properties for the synthesis and application of new substances is a relevant issue from the perspective of process systems engineering. However, several authors have indicated that the experimental quantification of critical

properties is time-consuming, potentially expensive for new compounds, and may show other limitations (e.g., some compounds with high molecular weight can degrade before reaching the critical point) [2–5]. Therefore, the development of thermodynamic models to calculate the pure compound critical properties is fundamental for process design [6–10].

The calculation of critical points with group contribution models (GCMs) is usually straightforward and offers additional advantages [1, 5, 7–9, 11–21]. These models are versatile, easy to implement, and do not require a substantial computational effort. GCMs can estimate the property of a compound from the individual contributions of functional groups that conform its molecular structure [16]. They

assume that the contribution of each functional group will be the same for any molecule thus allowing to expand its application for the prediction of almost any type of chemical substance [22].

Different GCMs have been proposed to calculate the normal boiling point and critical properties of pure compounds [1,9,11,12,14,16,23–27]. An extensive review and description of these models was reported by Su et al. [9]. One of the first GCMs for the estimation of critical properties of pure compounds was proposed by Lydersen [23]. Then, Joback and Reid [11] proposed different first-order GCMs to calculate several properties of pure compounds including critical properties and normal boiling point. Another set of GCMs which considered two levels of analysis for the estimation of group contribution values were proposed by Constantinou and Gani [12]. In these models, the basic level used first-order group contributions constituted by simple functional groups and the higher level included second-order groups. This approach allows for distinction between isomers; however, its implementation is more difficult in comparison to the first-order models as those proposed by Joback and Reid [11]. Other GCM that considered group-interaction contributions (i.e., the contributions of interactions between bonding groups in the molecule) for the calculation of critical properties and normal boiling point of pure organic compounds were reported by Marrero-Morejón and Pardillo-Fontdevila [14]. The experimental database used to develop this model included 507 organic compounds where 39 groups were considered. Marrero and Gani [1] proposed another GCM that determined the contributions at three levels. Specifically, the first level utilized contributions from simple groups, and the second and third levels included polyfunctional and structural groups that supply more information about the molecular fragments.

Retzekas et al. [28] proposed a GCM to calculate the normal boiling point, critical temperature, and pressure of hydrocarbons. This model involved compound density as an additional parameter. A second-order GCM for the prediction of the critical temperature of pure compounds was introduced by Dalmazzone et al. [15]. They used an experimental database constituted of 381 organic compounds including hydrocarbons and their derivatives containing oxygen, chlorine, nitrogen, bromine, and sulfur. Wang et al. [25], Wang et al. [26], and Jia et al. [27] published a set of GCMs for the estimation of critical properties (temperature, pressure, and volume) of organic compounds. But these models were only applicable to compounds with relatively small molecular weights and for those compounds with carbon chains from C2 to C18 that contain oxygen, nitrogen, sulfur, chlorine, and bromine groups. These authors used an experimental database with 467, 232, and 219 values of critical temperature, pressure, and volume, respectively, of organic compounds to determine the group contributions. Recently, Ghasemitabar and Movagharnejad [29] developed a second-order GCM for the normal boiling point of pure organic compounds including hydrocarbons with nitrogen, oxygen, sulfur, fluorine, chlorine, bromine, and iodine atoms. This model applied the first-order groups

reported by Joback and Reid [11] but included additional second-order groups that allowed to distinguish between isomers. A similar approach to obtain the critical properties of pure organic compounds was proposed by Tahami et al. [5]. They used an experimental database with 969, 715, and 539 data points of critical temperature, pressure, and volume, respectively.

On the other hand, some artificial intelligence tools such as artificial neural networks (ANN) have been used for the modeling of difficult nonlinear systems thus proving their numerical capabilities and reliable performance [30–33]. For the case of thermodynamic property modeling, ANN has been also applied with GCMs to obtain better estimations of the normal boiling point and critical properties of pure compounds [6, 8, 34].

Although a significant advance has been achieved in the prediction of these relevant thermodynamic properties, the major failure of these models still relies on their high prediction errors for different molecules (especially those with limited experimental data). Note that the lack of numerical values of some group contributions affects the property calculation via GCMs for some molecules, especially with complex chemical structures [35]. In this direction, it is important to highlight that a direct and reliable comparison of the performance of available GCMs is usually not feasible due to these models have been developed using different databases and numerical frameworks to establish their group contributions for the estimation of the thermodynamic property at hand. To date, few studies have reported the performance comparison of some GCMs for the calculation of pure compound critical properties but without recalculating their group contributions using the same basic line.

In this research, several first-order GCMs have been assessed and compared in the calculation of the normal boiling point and critical properties of pure compounds. This study has focused on the evaluation of first-order GCMs because they are a straightforward option to estimate these thermodynamic properties and are already available in commercial process simulators. In contrast to higher-order GCMs, they do not require additional information on molecules and are easy to implement by the users. Therefore, the novelty of this study relies on a reliable comparison of GCMs based on the application of the same analytical framework. First, this comparison was carried out via a standardized approach to establish the group contributions with the same thermodynamic data (correlation and validation sets) and an identical numerical procedure to solve the corresponding parameter estimation problems. A set of statistical metrics was applied to characterize the limitations and capabilities of tested models including a detailed performance analysis for different compound families where the best model was identified. Then, an ANN-based approach was assessed to improve the estimations obtained with the GCMs. Finally, a comparative analysis of these models was carried out using the GCMs implemented in the Aspen Plus® simulator. Therefore, this study highlights some gaps to be resolved for GCMs with the aim of developing reliable thermodynamic models for the calculation of pure compound properties.

2. Approach to Analyze and Compare the Performance of First-Order Group Contribution Models to Calculate the Critical Properties and Normal Boiling Point of Pure Compounds

2.1. Experimental Database. A database with experimental information on critical volume (V_c), critical pressure (P_c), critical temperature (T_c), and normal boiling point (T_b) of different compound families was integrated for this study. This information was extracted from the National Institute of Standards and Technology (NIST), a compilation of articles from the American Chemical Society [36–46], and other sources of physicochemical properties were also consulted to integrate this database [47–49]. A preliminary analysis of the available experimental information on these properties was performed to identify inconsistencies in the values reported in different sources for the same compound. The experimental values that displayed inconsistencies between two or more sources were discarded to reduce the uncertainty in the determination of the contributions of functional groups for the calculation of the corresponding thermodynamic properties. Table 1 shows the number of experimental values that were employed in the database of this study. In short, this database included the next experimental values of pure compounds: 397 for T_b , 514 for T_c , 445 for P_c , and 312 for V_c . This experimental database included organic compounds with carbon numbers from C1 to C60 and molecular weights ranging from 26.04 to 843.6 g/mol, organic compounds with metals (e.g., iron and titanium) as well as silicon and boron groups. These groups are usually not included in other GCMs reported in the literature such as the models proposed by Marrero and Gani [1], Marrero-Morejón and Pardillo-Fontdevila [14], Constantinou and Gani [12], Joback and Reid [11], Wen and Qiang [50], Dalmazzone et al. [15], Valderrama and Alvarez [51], Wang et al. [25], Wang et al. [26], Jia et al. [27], Deng et al. [34], Ghasemitarab and Movagharnjad [29], Mondejar et al. [8] and Tahami et al. [5]. Therefore, this database was used to define the numerical values of group contributions for the set of first-order GCMs reported in this study.

2.2. Definition of Functional Groups for GCMs. Table 2 shows the functional groups of first-order GCMs to estimate each thermodynamic property. The criteria applied to determine these functional groups included:

- The molecular groups were defined to represent a wide range of compounds including multifunctional oxygen, nitrogen, sulfur, halogenated, organometallic, silicon, and boron substances.
- The molecular groups were defined as simple as possible but representative of all the compounds included in the database.
- The molecular groups contained in chemical structures with and without rings (e.g., benzene and

phenol) were differentiated. This aspect was already evaluated by Marrero and Gani [1] showing that the performance of GCMs was better when there was a differentiation of the groups inside and outside the cyclic molecular structures.

2.3. Description of GCMs Used and Compared for the Calculation of V_c , P_c , T_c , and T_b of Pure Compounds. A set of first-order GCMs that do not consider group interactions were applied and compared in the calculation of the thermodynamic properties. Different equations (i.e., functionalities) reported for other GCMs were used in this study (see Table S1 of Supporting Information). A comparison and evaluation of these functionalities to estimate V_c , T_c , P_c , and T_b of pure compounds was performed. The group contributions and parameters of these functionalities were calculated using the same numerical framework (i.e., database, correlation and validation sets, objective function, and optimization method). These GCMs are described below for each thermodynamic property.

2.3.1. Normal Boiling Point. Three GCMs (i.e., *MTB1*, *MTB2*, and *MTB3*) were used and assessed to estimate T_b of pure compounds. In particular, the *MTB1* model was based on the equation proposed by Joback and Reid [11] for the calculation of the normal boiling point of pure compounds. The equation proposed is given by

$$T_b = A_1 + \sum_k N_k t_{bk}, \quad (1)$$

where N_k is the number of repetitions of the functional group k in the molecule of the tested compound, t_{bk} is the contribution of the functional group k to estimate the normal boiling point, and A_1 is a constant value for all compounds.

MTB2 model corresponded to the equation proposed by Nannoolal et al. [24]. The expression of this model is defined by

$$T_b = \frac{\sum_k N_k t_{bk}}{n_t^{A_2} + A_3} + A_4, \quad (2)$$

where n_t is the atoms number in the molecule (except hydrogen) and A_2 , A_3 , and A_4 are constant values for all the compounds.

MTB3 model was a modification of the model proposed by Joback and Reid [11] where a relationship between the molecular weights of the functional group (MM_k , g/mol) and the pure compound (MM , g/mol) was incorporated as follows:

$$T_b = A_5 + \sum_k (MM_k/MM) N_k t_{bk}, \quad (3)$$

where A_5 is a constant value for all the compounds.

2.3.2. Critical Temperature. Four GCMs were implemented to calculate the critical temperature. *MTC1* model was based on the equation proposed by Lydersen [23].

TABLE 1: Description of experimental database of thermodynamic properties used in this study.

Compound family	Thermodynamic property [†]											
	T_b , K			T_c , K			P_c , bar			V_c , cm ³ /mol		
	No. of data	Min	Max	No. of data	Min	Max	No. of data	Min	Max	No. of data	Min	Max
Alkane	28	184.6	722.9	32	305.3	974.0	32	3.6	48.7	19	145.5	189.0
Alkene	15	169.0	614.2	17	282.3	772.0	17	11.4	50.4	8	131.1	584.0
Alkyne	3	189.0	310.0	3	308.3	488.0	3	40.6	61.4	3	112.2	275.3
Cyclic hydrocarbon	10	240.0	466.0	12	398.0	707.0	11	25.3	55.4	11	162.0	566.5
Aromatic hydrocarbon	28	353.3	711.2	28	562.1	1013.2	27	15.3	49.0	22	256.0	1078.7
Alcohol	16	337.8	634.6	24	512.5	850.0	24	10.7	80.8	13	117.0	1281.5
Phenol	4	455.0	550.0	4	694.2	836.0	3	43.6	62.4	2	229.0	312.0
Carboxylic acid	16	391.2	672.9	28	590.7	862.0	27	11.1	65.9	13	171.0	1130.0
Aldehyde	3	441.7	481.7	3	639.0	695.0	3	26.0	47.0	2	488.0	598.8
Ketone	14	329.3	550.0	16	508.1	830.0	11	23.2	47.0	15	213.0	896.0
Ether	11	285.0	560.2	12	466.8	824.0	10	18.2	72.0	9	142.0	720.0
Ester	16	330.0	659.0	27	506.5	886.0	24	8.9	47.5	10	228.0	568.2
Nitrogen	61	252.0	628.2	68	400.0	901.0	63	20.0	82.4	41	141.0	584.0
Multifunctional oxygen	16	223.3	557.2	23	380.0	864.0	19	26.6	65.0	9	145.0	549.5
Chlorine	21	247.0	456.4	21	416.3	685.7	11	43.0	66.8	8	140.0	308.6
Fluorine	43	145.1	479.0	50	227.5	673.0	39	14.5	58.7	36	109.0	781.0
Bromine	4	311.5	429.1	4	504.0	670.0	4	41.3	58.0	3	214.0	324.1
Iodine	3	362.4	473.2	3	574.6	749.9	3	43.3	47.3	3	285.5	353.9
Other halogenated	57	191.8	442.0	84	302.0	684.8	64	11.6	58.7	55	147.0	754.0
Sulfur	17	279.1	605.7	16	381.2	897.0	12	18.1	72.3	11	147.0	512.0
Silicon	10	299.7	446.2	31	352.4	893.0	30	4.7	35.3	19	344.0	2524.0
Iron	1	—	—	4	784.0	886.0	4	15.9	36.1	—	—	—
Titanium	—	—	—	4	641.0	751.0	4	9.1	14.0	—	—	—

[†]Symbols Min and Max represent the minimum and maximum experimental values of the thermodynamic property in the database.

$$T_c = \frac{T_b}{B_1 + \sum_k (N_k t_{ck}) - (\sum_k N_k t_{ck})^2}, \quad (4)$$

where t_{ck} is the contribution of the functional group k to estimate the critical temperature and B_1 is a constant value of this GCM.

MTC2 model was an expression proposed by Joback and Reid [11] and corresponded to a modified version of equation (4) that was given by

$$T_c = \frac{T_b}{B_2 + B_3 \sum_k (N_k t_{ck}) - (\sum_k N_k t_{ck})^2}, \quad (5)$$

where B_2 and B_3 are constant parameters.

The third model (*MTC3*) was a modified version of the Joback and Reid [11] model. This model included a relationship of MM_k and MM where it was defined as

$$T_c = \frac{T_b}{B_4 + B_5 \sum_k (MM_k/MM) N_k t_{ck} - (\sum_k (MM_k/MM) N_k t_{ck})^2}, \quad (6)$$

where B_4 and B_5 are constant parameters.

MTC4 model was also a modified expression of Joback and Reid [11] that included the atoms number (except hydrogen) and rings contained in the molecule under analysis

$$T_c = \frac{T_b}{(B_6 \cdot Nr) + (B_7 \cdot n_t^{B_8}) + B_9 \sum_k (N_k t_{ck}) - (\sum_k N_k t_{ck})^2}, \quad (7)$$

where Nr is the rings number in the molecule and $B_6, B_7, B_8,$ and B_9 are constant parameters of this GCM.

2.3.3. Critical Pressure. Four GCMs were utilized to estimate the critical pressure. *MPC1* model was proposed by Lydersen [23] to estimate P_c via the following equation:

$$P_c = \frac{MM}{(C_1 + \sum_k (N_k p_{ck}))^2}, \quad (8)$$

where p_{ck} is the contribution of the functional group k to estimate P_c and C_1 is a constant parameter.

The second model (*MPC2*) was based on the equation proposed by Joback and Reid [11].

$$P_c = \frac{1}{(C_2 - (C_3 \cdot n_t) - \sum_k (N_k p_{ck}))^2}, \quad (9)$$

where C_2 and C_3 are constant parameters.

MPC3 model was a modification of equation (9) and is described as

$$P_c = \frac{MM}{(C_4 - (C_5 \cdot n_t) - \sum_k (N_k p_{ck}))^2}, \quad (10)$$

where C_4 and C_5 are constant parameters of the GCM.

The fourth model (*MPC4*) was another modified equation that included the number of rings of the molecule

TABLE 2: Functional groups of the first-order GCMs used to calculate critical properties and normal boiling point of pure compounds.

No.	
Groups without rings	
1	-CH ₃
2	-CH ₂ -
3	>CH-
4	>C<
5	=CH ₂
6	=CH-
7	=C<
8	=C=
9	≡C-
10	≡CH
Groups with rings	
11	-CH ₂ -
12	>CH-
13	=CH-
14	>C<
15	=C<
Groups with oxygen	
16	-OH (alcohol)
17	-OH (phenol)
18	-O- (without rings)
19	-O- (with rings)
20	>C=O (without rings)
21	>C=O (with rings)
22	-CHO (aldehyde)
23	-COOH (acid)
24	-COO- (ester)
25	=O (others)
Groups with nitrogen	
26	>NH (without rings)
27	>NH (with rings)
28	>N- (without rings)
29	>N- (with rings)
30	-N= (without rings)
31	-N= (with rings)
32	-NH ₂ (without rings)
33	-NO ₂
34	-CN
Groups with halogen	
39	-S- (without rings)
40	-S- (with rings)
41	=S<
42	≥S≤
43	-SH
Groups with silicon	
44	>Si< (without rings)
45	>Si< (with rings)
46	>SiH-
47	-SiH ₃
Groups with metals	
48	-Fe-
49	>Ti<
Groups with boron	
50	>B-

$$P_c = \frac{MM}{(C_6 - (C_7 \cdot n_t) + (C_8 \cdot Nr) - \sum_k (N_k p_{ck}))^2}, \quad (11)$$

where C_6 , C_7 , and C_8 are constant parameters.

2.3.4. *Critical Volume.* *MVC1*, *MVC2*, *MVC3*, and *MVC4* models were compared in the calculation of critical volume of pure compounds. *MVC1* was proposed by Lydersen [23] and defined as

$$V_c = D_1 + \sum_k (N_k v_{ck}), \quad (12)$$

where v_{ck} is the contribution of the functional group k to estimate the critical volume and D_1 is a constant parameter.

MVC2 model was a modification of equation (12) that included the molecular weights of functional groups and pure compound

$$V_c = D_2 + \sum_k \left(\frac{MM_k}{MM} N_k v_{ck} \right), \quad (13)$$

where D_2 is a constant parameter of this GCM.

MVC3 and *MVC4* were obtained from the modification of equation (12) where the atoms number in the molecule (except hydrogen) was included and they were given by the next equations

$$V_c = D_3 + (D_4 \cdot n_t) + \sum_k (N_k v_{ck}), \quad (14)$$

$$V_c = D_5 + (D_6 \cdot n_t) + (D_7 \cdot MM) + \sum_k (N_k v_{ck}), \quad (15)$$

where D_3 , D_4 , D_5 , D_6 , and D_7 are constant parameters for all the pure compounds.

2.4. *Determination of the Contributions of Functional Groups of GCMs for the Thermodynamic Property Calculation.* Group contributions and constant parameters of GCMs (see equations (1)–(15)) were determined via a nonlinear regression of the experimental database. 337, 404, 378, and 265 experimental data of T_b , T_c , P_c , and V_c , respectively, were used in this stage (i.e., 79–85% of the thermodynamic database). The frequency of each functional group in the molecules employed in the correlation stage can be found in an Excel file (Groups_correlation.xlsx) that is provided in the Supporting Information. The objective function utilized for the correlation of the experimental data was defined as follows:

$$FO = \min \sum_{i=1}^m \left(\frac{Y_i^{\text{exp}} - Y_i^{\text{cal}}}{Y_i^{\text{exp}}} \right)^2, \quad (16)$$

where m is the number of experimental data and Y_i^{exp} and Y_i^{cal} are the experimental and calculated values of tested

thermodynamic properties (i.e., V_c , P_c , T_c , and T_b). This objective function was minimized with a combination of a local optimization method and an evolutionary algorithm. This optimization tool was used to determine the parameters and group contributions of the set of GCMs. Several initial values were used to resolve the parameter identification problems because it was assumed as a global optimization problem and the best solution found was reported in this study.

The predictive capabilities of tested GCMs were evaluated using pure compounds (i.e., molecules) not considered in the data correlation stage. These compounds were included in the validation set and represented approximately 15% of the experimental thermodynamic property database (i.e., 60, 110, 67, and 47 experimental data of T_b , T_c , P_c , and V_c , respectively). For illustration, Figure 1 shows the flow-chart of the procedure utilized to determine GCMs parameters and to perform their evaluation. To reduce the deviations obtained in the determination of the functional group contributions and the model parameters, the next criteria were applied to establish the correlation and validation sets:

- The validation set included compounds with experimental data reported by different sources where their experimental deviations were higher than 1% for the normal boiling point and 5% for the critical properties.
- The validation set included compounds whose experimental properties could not be confirmed by at least two literature sources.
- It was established that the validation set did not include molecules with functional groups that were not present in the correlation set.
- Functional groups with repetitions ≤ 5 were considered only in the correlation set.

2.5. Application of ANN to Improve the Predictions of GCMs.

An ANN was applied as a corrector to improve the estimations obtained with GCMs. ANN input variables were the thermodynamic property calculated with GCMs and molecular weight (MM) of the pure compound, while the output variable was the improved (corrected) thermodynamic property value (see Figures S1 and S2). The ANN toolbox of Matlab[®] was used in these calculations. A feed-forward backpropagation ANN was employed. The input data set was randomly divided into three subgroups: training set (70% of the database), validation set (15% of the database), and test set (15% of the database). ANN training was used to determine the best values of weights and biases where the Levenberg–Marquardt training method and a logarithmic sigmoid transfer function were employed. The performance function corresponded to the mean square error (MSE), which is given by

$$MSE = \frac{1}{m} \sum_{i=1}^m (Y_i^{\text{exp}} - Y_i^{\text{cal}})^2. \quad (17)$$

Different ANN configurations (i.e., hidden neurons ranging from 1 to 10 and layers ranging from 1–2) were evaluated to identify the simplest one with the best performance (without overfitting) and to establish their impact on the calculation of thermodynamic properties with the GCM-ANN approach.

2.6. Statistical Analysis to Compare GCMs Performance.

Statistical metrics were utilized to assess the performance of all GCMs for the thermodynamic property calculation. These metrics were the absolute relative deviation (ARD_i), average absolute relative deviation ($AARD$), and standard deviation (σ), which are defined by

$$ARD_i = 100 \cdot \left| \frac{Y_i^{\text{exp}} - Y_i^{\text{cal}}}{Y_i^{\text{exp}}} \right|, \quad (18)$$

$$AARD = \frac{1}{m} \sum_{i=1}^m 100 \cdot \left| \frac{Y_i^{\text{exp}} - Y_i^{\text{cal}}}{Y_i^{\text{exp}}} \right|, \quad (19)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^m (ARD_i - AARD)^2}{m}}. \quad (20)$$

Additionally, a comparative analysis of these GCMs was also carried out using the GCMs implemented in Aspen Plus[®]. These GCMs were proposed by Riedel [52], Lydersen [23], Ambrose [53,54], Fedors [55,56], Joback and Reid [11], and Constantinou and Gani [12]. In particular, the thermodynamic properties of some molecules not included in the training set were calculated with Aspen Plus[®] and these results were compared with those estimations obtained with tested first-order GCMs.

3. Results and Discussion

3.1. Normal Boiling Point. Figure 1 shows the comparison of the experimental and calculated T_b values using different GCMs. The determination coefficient (R^2) was included for the correlation and validation stages of each model. Tables S2 and S3 of Supporting Information provide the group contributions and parameters of $MTB1$, $MTB2$, and $MTB3$ models for T_b estimation. Overall, these GCMs showed R^2 values of 0.83–0.97 where the highest ARD was obtained for compounds with T_b higher than 500 K. The performance of these GCMs showed the next trend: $MTB2 > MTB1 > MTB3$. In particular, the $MTB2$ model was the best to calculate the T_b of tested compounds. Note that this model included the number of atoms in the molecule as an additional parameter. Also, it should be highlighted that the nonlinear functionality of the $MTB2$ model outperformed those of the $MTB1$ and $MTB3$ models. Table 3 summarizes the statistical metrics used to evaluate the performances of these models. ARD values ranged from 8.34E-08 to 60.48% in T_b calculation.

For illustration, Figure 2 shows $AARD$ obtained in T_b estimation for different compound families. This analysis allowed us to identify some compound families with the highest deviations between experimental and predicted

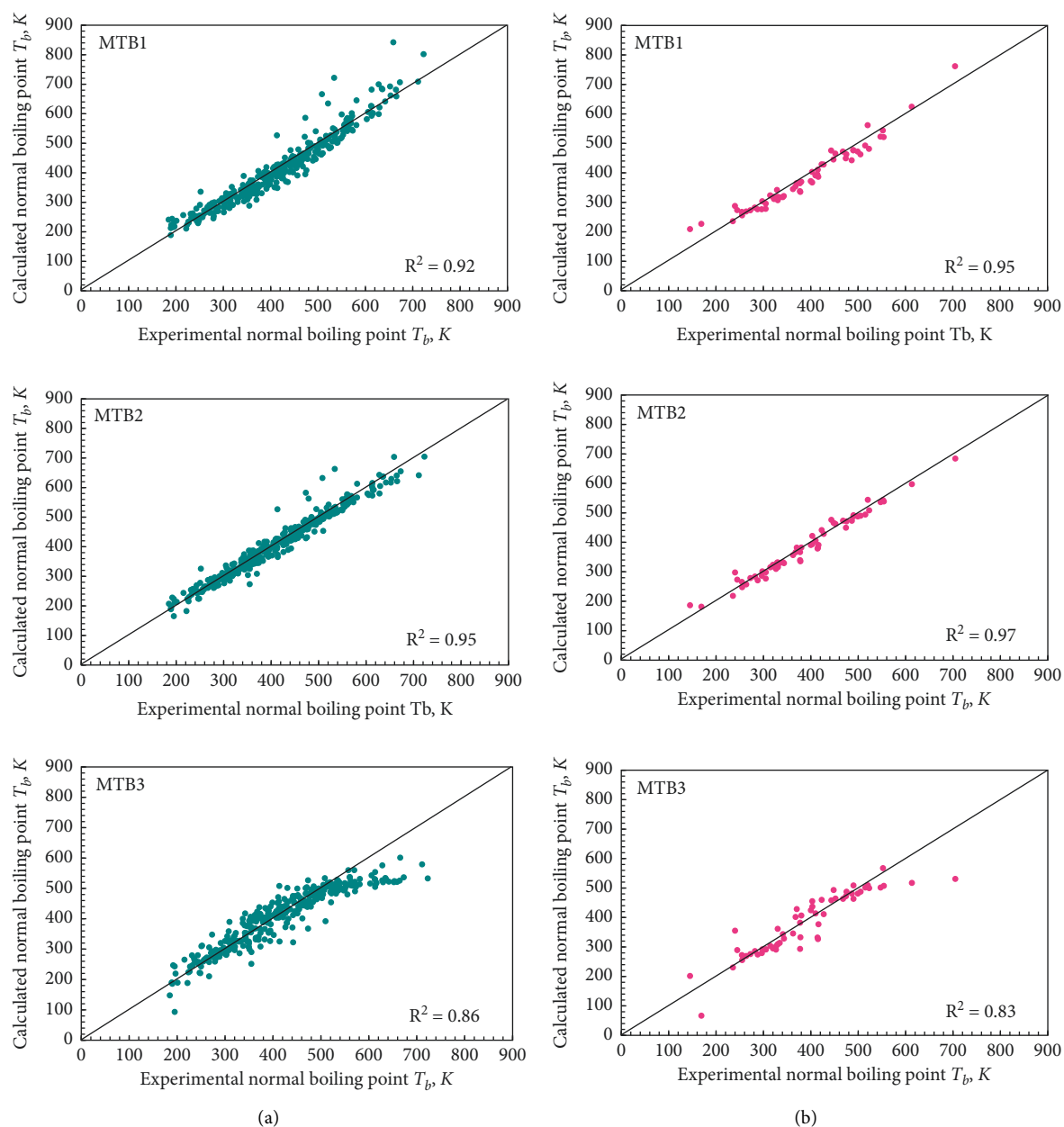


FIGURE 1: Results of the calculation of the normal boiling point of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

values. GCMs were able to calculate the T_b of alkynes, phenols, multifunctional oxygen, sulfur, and silicon compounds with acceptable accuracy. AARD in these families ranged from 0.28 to 6.37%, while this statistical parameter increased for alkanes, carboxylic acids, nitrogen, and halogenated compounds. In general, these GCMs showed the highest AARD for alkanes, alkenes, cyclic hydrocarbons, nitrogen, chlorine, and fluorine compounds. These results can be associated with the relatively high uncertainty of the experimental data reported for different compounds with a complex chemical structure (e.g., some nitrogen, and halogenated compounds). Figure S3 reports the ARD values of the MTB2 model as a function of MM and T_b for different compound families where the magnitude of this statistical

metric is indicated by the size of the symbol "O". Although the MTB2 model was the best model, it presented the highest ARD for those compounds with low and high molecular weights within a homologous series. For example, high deviations were obtained for alkanes and alkenes with low molecular weight. This behavior was identified and highlighted by Nannoolal et al. [16]. Low molecular weight compounds usually do not follow the general trend of the homologous series of a given chemical family. However, the experimental data of T_b for these compounds are commonly available in the literature and there is no need to estimate them using GCMs. For some halogenated compounds with molecular weights <100 g/mol, high deviations were obtained. Ethanedinitrile and 1,1,1,3,5,5,5-heptamethyltrisiloxane were

TABLE 3: Results of statistical metrics used to compare the performance of first-order GCMs for the calculation of critical properties and normal boiling point of pure compounds.

Property	Model	No. of data	Correlation			No. of data	Validation		
			Statistical metric, %				Statistical metric, %		
			Max ARD _i	Min ARD _i	σ		Max ARDi	Min ARDi	σ
T_b	<i>MTB1</i>	337	35.16	4.90E-07	5.58	60	44.27	0.05	7.29
	<i>MTB2</i>		29.31	8.34E-08	4.37		28.53	0.29	5.00
	<i>MTB3</i>		52.14	1.05E-03	6.81		60.48	0.22	11.10
T_c	<i>MTC1</i>	404	21.85	6.36E-08	2.57	110	72.39	0.01	16.83
	<i>MTC2</i>		21.82	1.22E-07	2.47		65.52	0.05	14.98
	<i>MTC3</i>		23.42	4.95E-04	2.71		107.69	0.10	22.97
	<i>MTC4</i>		21.44	3.43E-04	2.43		70.23	0.22	15.06
P_c	<i>MPC1</i>	378	36.08	3.43E-07	5.46	67	46.17	0.23	8.71
	<i>MPC2</i>		35.87	4.34E-07	6.26		40.85	0.17	9.67
	<i>MPC3</i>		36.08	3.27E-05	5.46		46.17	0.23	8.71
	<i>MPC4</i>		36.08	2.35E-04	5.47		46.17	0.15	8.72
V_c	<i>MVC1</i>	265	29.98	4.41E-04	3.47	47	17.82	0.10	3.36
	<i>MVC2</i>		73.94	4.05E-07	16.50		52.07	0.56	12.97
	<i>MVC3</i>		29.98	1.98E-03	3.47		17.82	0.10	3.37
	<i>MVC4</i>		29.97	1.01E-03	3.47		17.82	0.09	3.37

the compounds with the maximum (29.31%) and minimum (8.34E-08%) ARD to calculate their T_b using *MTB2*. It was identified that high deviations (i.e., ARD >20%) were obtained for some molecules containing functional groups with halogen and nitrogen atoms (e.g., -F, -CN, and -NH₂).

3.2. Critical Temperature. Figure 3 and Table 3 show the results of T_c calculation with GCMs and the corresponding model parameters are reported in Tables S4 and S5 of Supporting Information. It was identified that the highest ARD was obtained for pure compounds with $T_c > 600$ K. It should be highlighted that for these compounds, T_c experimental data reported in the literature showed high deviations. *MTC4* model displayed the best performance for the calculation of T_c . This model considered the number of atoms and rings in the molecule as parameters. Overall, ARD ranged from 6.36E-08 to 72.39% for *MTC1*, from 1.22E-07 to 65.52% for *MTC2*, from 4.95E-04 to 107.69% for *MTC3* and from 3.43E-04 to 70.23% for *MTC4*, respectively. AARD values for T_c estimation of different compound families are reported in Figure 4. All GCMs showed high AARD for alkanes, alcohols, carboxylic acids, ethers, esters, nitrogen, halogenated, and silicon compounds. On the other hand, the calculation performance of these GCMs was acceptable (i.e., AARD < 2.5%) for families of alkenes, cyclic and aromatic hydrocarbons, ketones, and multifunctional oxygen compounds. All GCMs presented the highest ARD for organometallic compounds. The highest deviations in the calculation of T_c could be associated with the uncertainty of the experimental values used to determine the group contributions, especially for complex molecules with high molecular weight. *MTC4* outperformed other models to calculate this thermodynamic property where *MTC3* was the worst GCM.

Figure S4 reports the ARD values of the *MTC4* model. This model showed the highest deviations for calculating T_c for those compounds with high molecular weight within a homologous series. This trend was also observed for alkanes

where the deviations were higher for compounds with molecular weights >500 g/mol, alkenes with molecular weights >200 g/mol, esters with molecular weights >300 g/mol, and silicon compounds with molecular weights >400 g/mol. Also, the modeling errors were higher for high molecular weight aromatic hydrocarbon and alcohol compounds. (2R, 4R, 6R)-2,4,6-trimethyl-2,4,6-triphenylcyclotrisiloxane was the compound with the maximum ARD (i.e., 70.23%) to calculate T_c using *MTC4*. This compound was a silicon compound with a high molecular weight (408.67 g/mol). It could be anticipated that GCMs could show high modeling errors for the calculation of critical properties of high molecular weight compounds. As indicated, the procedures utilized in the experimental determination of critical properties may have some limitations, mainly for the high molecular weight compounds where these type of chemicals can degrade before reaching the critical point. These deviations in the experimental determination of the critical point introduced uncertainties in the determination of the group contribution values of tested GCMs. Higher modeling errors (ARD >30%) were obtained in the T_c prediction of some molecules containing silicon, nitrogen, and halogenated functional groups (e.g., >Si<, -F, -CN, -NO₂).

3.3. Critical Pressure. Results of the calculation of P_c with tested GCMs are reported in Figure 5 and Table 3, while the estimated group contributions and additional model parameters are given in Tables S6 and S7 of Supporting Information. Overall, ARD values ranged from 3.43E-07 to 46.17% where GCMs followed the next performance in terms of the deviations for the estimation of P_c : *MPC2* > *MPC1* > *MPC3* > *MPC4*. In particular, *MPC2* showed high ARD (i.e., >10%) to estimate P_c at >60 bars. On the other hand, the *MPC4* model showed the best performance for the prediction of this thermodynamic property. This model also included the number of atoms and rings in the molecule as additional parameters. Table 3 summarizes the statistical metrics used in the comparison of these GCMs.

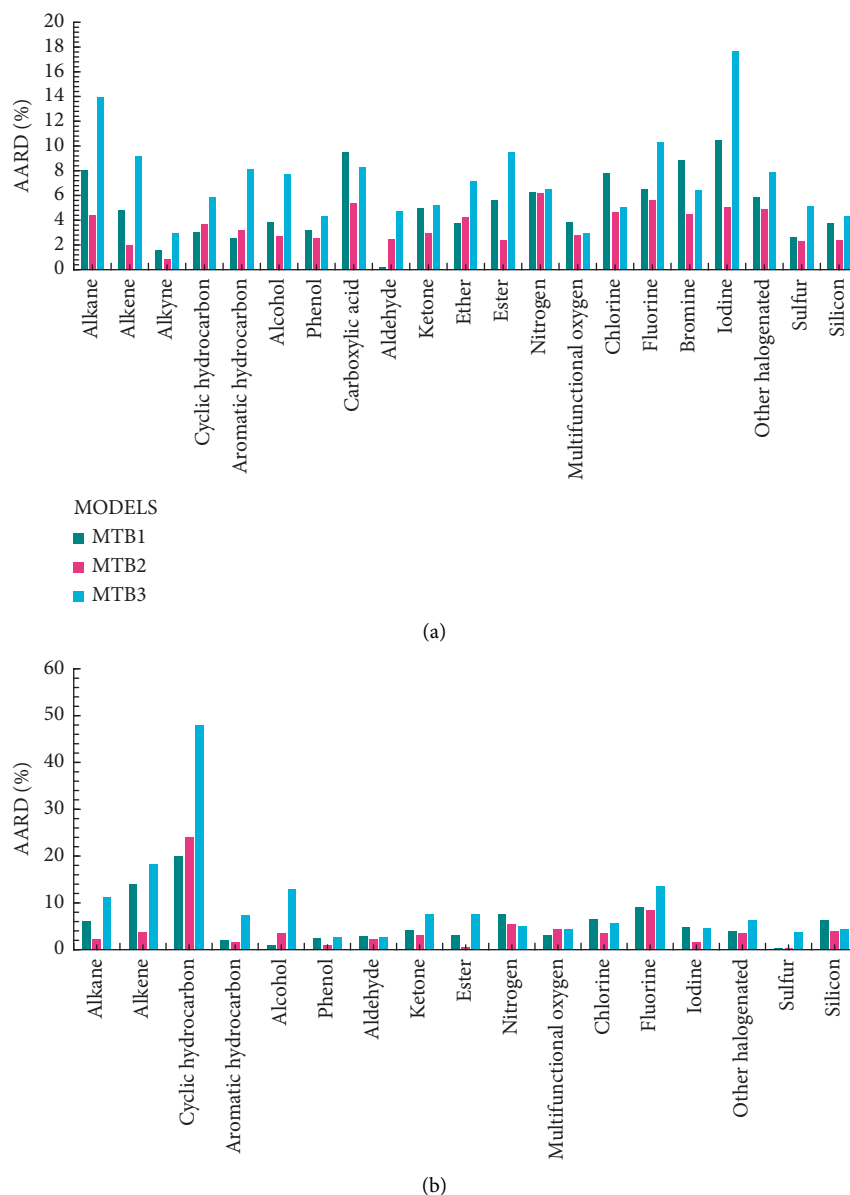


FIGURE 2: AARD for the calculation of the normal boiling point of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

Figure 6 shows the values of AARD in P_c estimation for different compound families. These results indicated that the P_c prediction from these GCMs showed the highest ARD for nitrogen, multifunctional oxygen, halogenated, and silicon compounds. All models showed high modeling errors to calculate P_c of iron compounds. For illustration, ARD values of the *MPC4* model for different families of chemicals are reported in Figure S5. In general, there were no clear trends in model performance with respect to the molecular weight and compound families. However, ARD >10% was obtained for alkanes with molecular weights higher than 400 g/mol. For some halogenated compounds (e.g., fluorine compounds), the performance of this GCM was better for substances with intermediate molecular weights within the homologous series. Trifluoroethanoic acid was the compound with the maximum ARD (46.17%) to calculate P_c using *MPC4*. The low

precision in the experimental information reported for the halogenated compounds could be the source of these modeling errors. It was also identified that the highest deviations (i.e., ARD >20%) were obtained for molecules constituted by functional groups containing iron, nitrogen, silicon, and halogen atoms (e.g., >Si<, -F, >N-, -Fe, and -NO₂).

3.4. Critical Volume. Results of the calculation of V_c using *MVC1*, *MVC2*, *MVC3*, and *MVC4* models are reported in Figure 7, while the corresponding statistical metrics for model comparison are given in Table 3. It was clear that *MVC1*, *MVC3*, and *MVC4* displayed a similar performance with ARD ranging from 4.41E-04 to 29.98%, while *MVC2* was the worst option to calculate this critical property. Tables S8 and S9 of Supporting Information contain the group contributions, and parameters of these GCMs. In

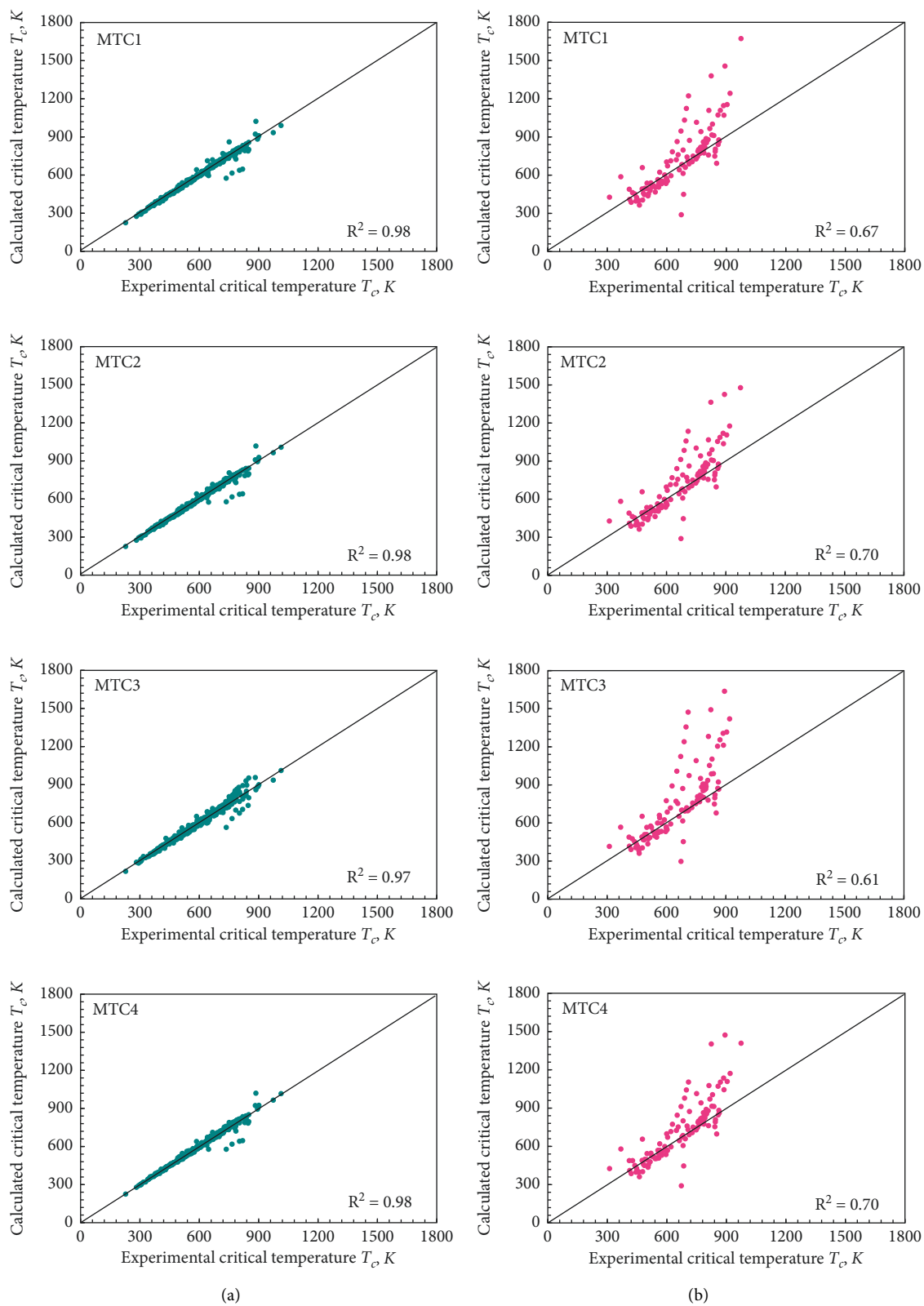


FIGURE 3: Results of the calculation of the critical temperature of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

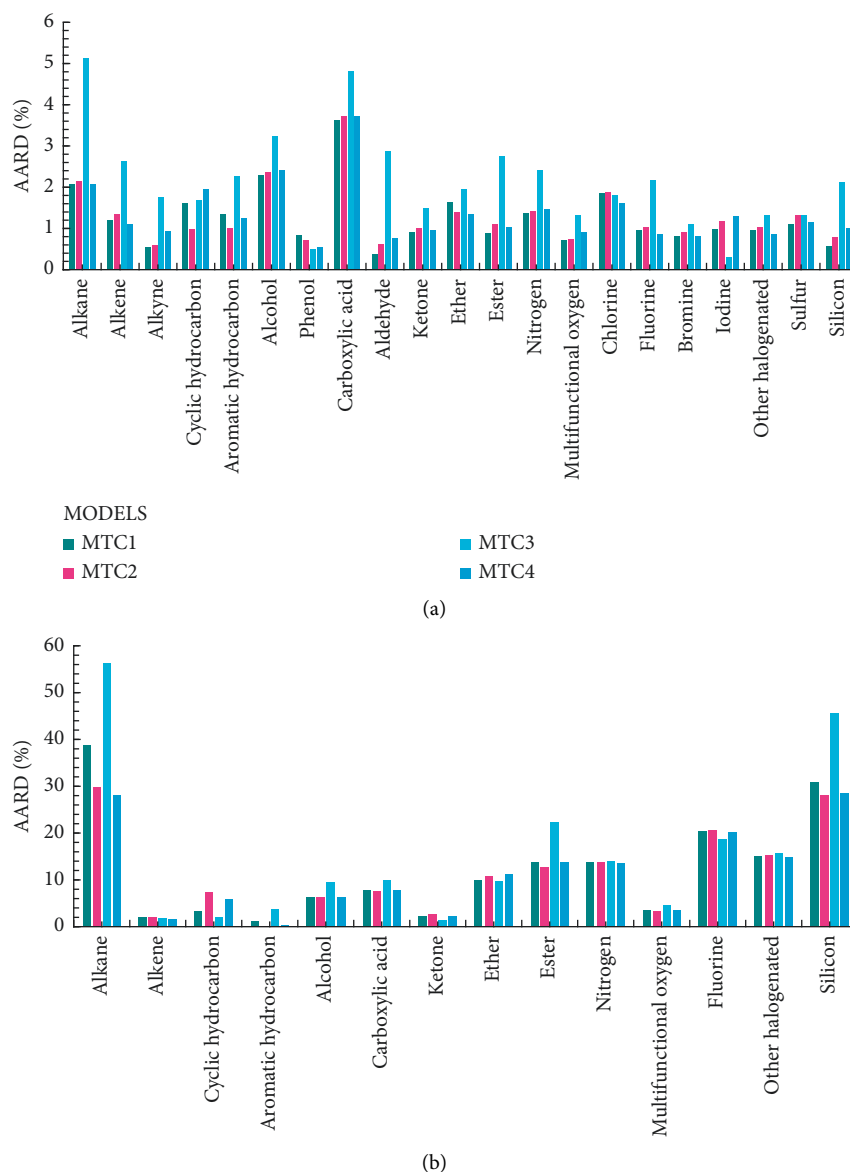


FIGURE 4: AARD for the calculation of the critical temperature of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

general, these GCMs had the highest AARD to predict V_c of nitrogen, multifunctional oxygen, fluorine, and silicon compounds, see Figure 8. Statistical metrics indicated that the *MVC4* model was the best to estimate critical volumes of pure compounds. This model also included the number of atoms and the molecular weight as additional parameters. Figure S6 illustrates the *MVC4* model performance for the V_c calculation of different compound families. AARD increased with respect to the molecular weight for alkanes, carboxylic acids, ketones, and fluorine compounds. These results also confirmed that the calculation of critical properties of high molecular weight compounds is challenging via GCMs. This GCM showed the highest AARD values for alcohols, ketones, nitrogen, and silicon compounds with low molecular weight. 1,1,1,5,5,5-hexafluoro-2,4-pentanedione was the compound with the maximum AARD (29.97%) for V_c calculation using *MVC4*. Different molecules containing some functional

groups with halogen and nitrogen (e.g., -F, >NH, -NH₂) showed the highest deviations (i.e., AARD >10%).

Finally, the Supporting Information contains an Excel file (Properties_calculation.xlsx) that can be utilized to calculate T_b and critical properties of pure compounds using the best GCMs reported in this paper.

3.5. Correction of the Thermodynamic Property Estimations of GCMs with ANN. AARD values for the calculation of V_c , P_c , T_c , and T_b using ANN to correct the GCM estimations are shown in Figure 9. These results corresponded to different ANN configurations in terms of the number of hidden layers and their neurons. The application of ANN allowed improving the estimations of thermodynamic properties obtained with GCMs. For the case of T_b , the AARD value of *MTB1* was 5.82% and decreased to 4.48% with the implementation of ANN with 1

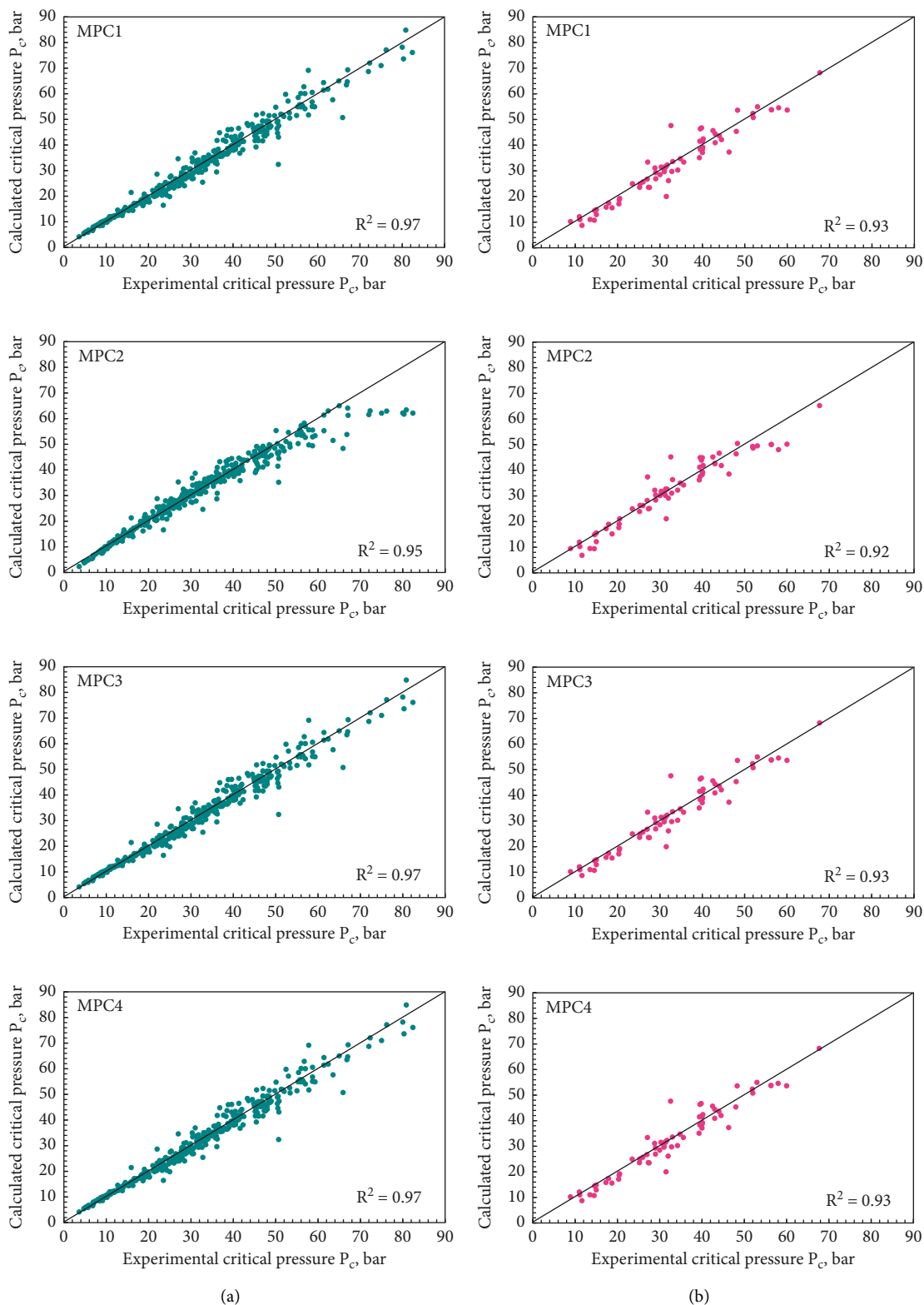


FIGURE 5: Results of the calculation of critical pressure of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

hidden layer and 5 hidden neurons. *MTB2* plus ANN with 2 hidden layers and 3 hidden neurons achieved an AARD of 3.90%, in contrast to 4.25% AARD of the original *MTB2*. AARD values of *MTB3* with and without ANN (1 hidden

layer and 10 hidden neurons) were 8.90 and 5.74%, respectively.

For T_c prediction, the AARD value of *MTC1* decreased from 14.56 to 9.76% using an ANN model with 2 hidden

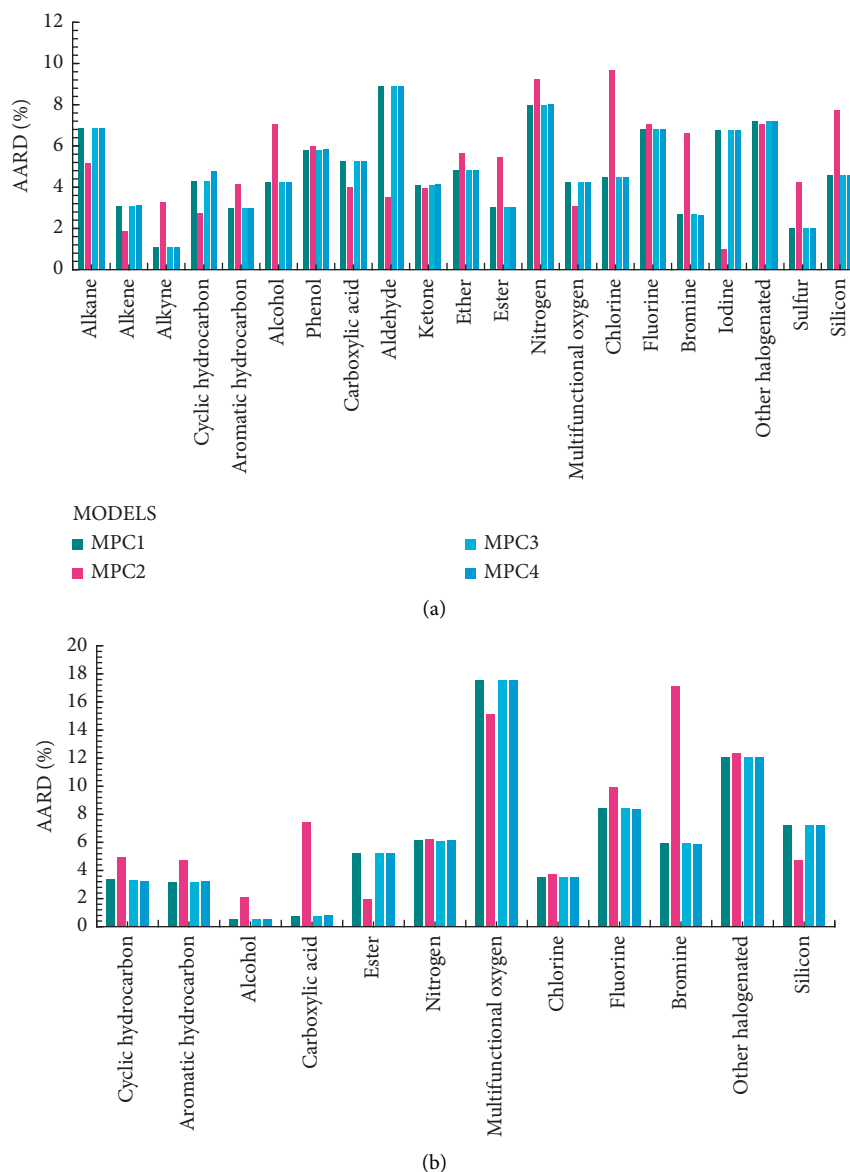


FIGURE 6: AARD for the calculation of critical pressure of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

layers and 5 hidden neurons. The application of ANN with 1 hidden layer and 7 hidden neurons decreased the AARD of *MTC2* from 13.66 to 9.79%. Similar results were obtained for *MTC3* and *MTC4* after applying the ANN as a model corrector where AARD decreased from 19.16 to 10.93% (using ANN with 1 hidden layer and 3 hidden neurons) and from 13.68 to 8.15% (using ANN with 1 hidden layer and 7 hidden neurons), respectively.

The ANN application also improved P_c estimations obtained with all GCMs. AARD decreased from 8.48 (*MPC1*) to 7.75% (*MPC1* plus ANN with 1 hidden layer and 3 hidden neurons), from 9.05 (*MPC2*) to 7.59% (*MPC2* plus ANN with 1 hidden layer and 5 hidden neurons), from 8.48 (*MPC3*) to 7.36% (*MPC3* plus ANN with 1 hidden layer and 5 hidden neurons), and from 8.48 (*MPC4*) to 7.28% (*MPC4* plus ANN with 2 hidden layers and 5 hidden neurons). These

results showed that ANN with several hidden layers and neurons was required to obtain an accurate estimation of this thermodynamic property. However, it is important to indicate that the best ANN structure should be identified to avoid model overfitting.

Finally, the approach GCM + ANN also improved the prediction of V_c although the final result was highly dependent on the model and ANN structure. For example, the AARD value of *MVC2* decreased from 18.24 to 11.13% using ANN with 2 hidden layers and 3 hidden neurons. In contrast, V_c calculations performed with other GCMs (*MVC1*, *MVC3*, *MVC4*) plus ANN did not show a significant reduction of AARD for different configurations of hidden layers and neurons (see Figure 9).

In summary, these results were consistent with the study of Mondejar et al. [8] where GCMs and ANN models

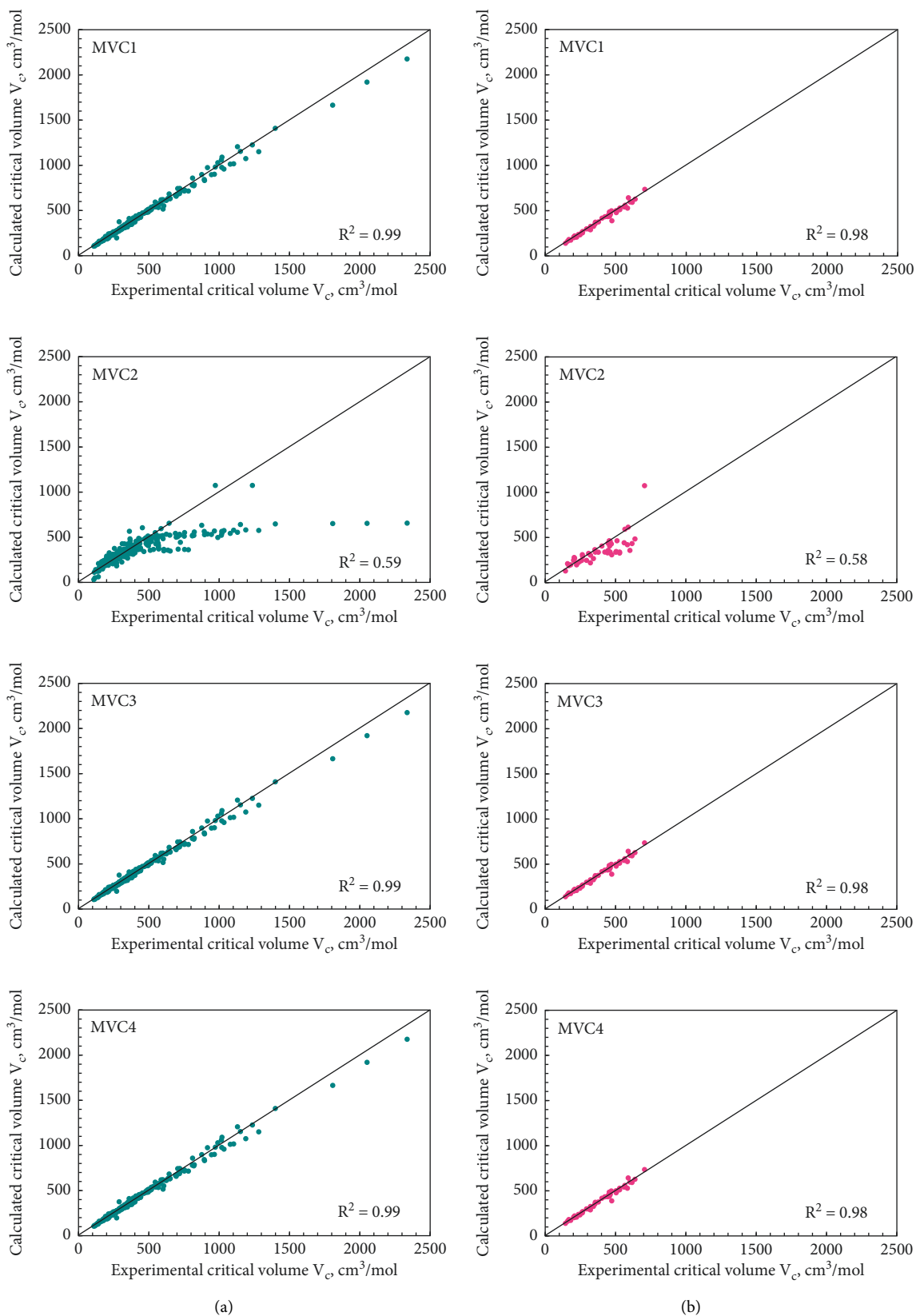


FIGURE 7: Results of the calculation of critical volume of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

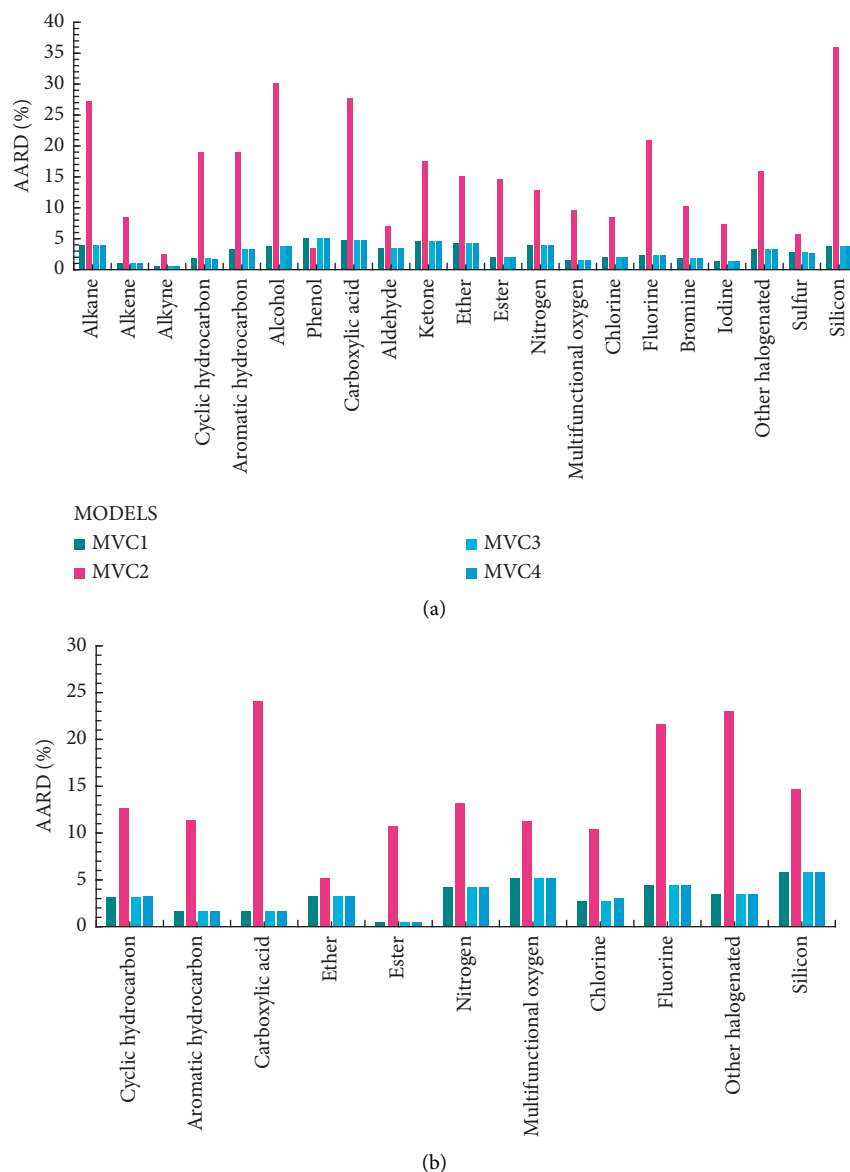


FIGURE 8: AARD for the calculation of critical volume of pure compounds using different first-order GCMs. (a) Correlation and (b) validation.

were also applied. However, these authors only analyzed a specific group of chlorine and fluorine compounds with carbon numbers from C2 to C20 in contrast to the present study where a broad spectrum of chemical families was analyzed.

3.6. Comparative Analysis of GCMs Reported in the Literature.

A comparison of the performances of different GCMs was carried out and the results are given in Table 4. Figure 10 shows the distribution of AARD for the calculation of tested thermodynamic properties of pure compounds using the first-order GCMs and ANN. As stated, the Aspen Plus® simulator was also employed to calculate these thermodynamic properties. This process simulator offers the possibility of using the next GCMs: R–Riedel [52], LS–Lydersen [23], A–Ambrose [53,54], F–Fedors [55,56], JR–Joback and

Reid [11], and CG–Constantinou and Gani [12]. Therefore, this comparison was carried out using a specific group of molecules where all the models were applicable, and the results obtained with Aspen Plus® were compared with the best first-order GCMs identified in this study.

Results reported in Table 4 and Figure 10 showed that the estimations of normal boiling point with *MTB2* and *MTB2* + ANN with 2 hidden layers and 3 hidden neurons were better than those obtained with the models of Constantinou and Gani [12] and Joback and Reid [11]. For T_c calculation, *MTC4* and *MTC4* + ANN with 1 hidden layer and 7 hidden neurons showed higher modeling errors than those of models proposed by Lydersen [23], Ambrose [53], Fedors [56], Joback and Reid [11], and Constantinou and Gani [12]. *MPC4* (with and without ANN) outperformed Lydersen [23], Ambrose [54], and Constantinou and Gani [12] models. Also, *MVC4* (with and without ANN) provided

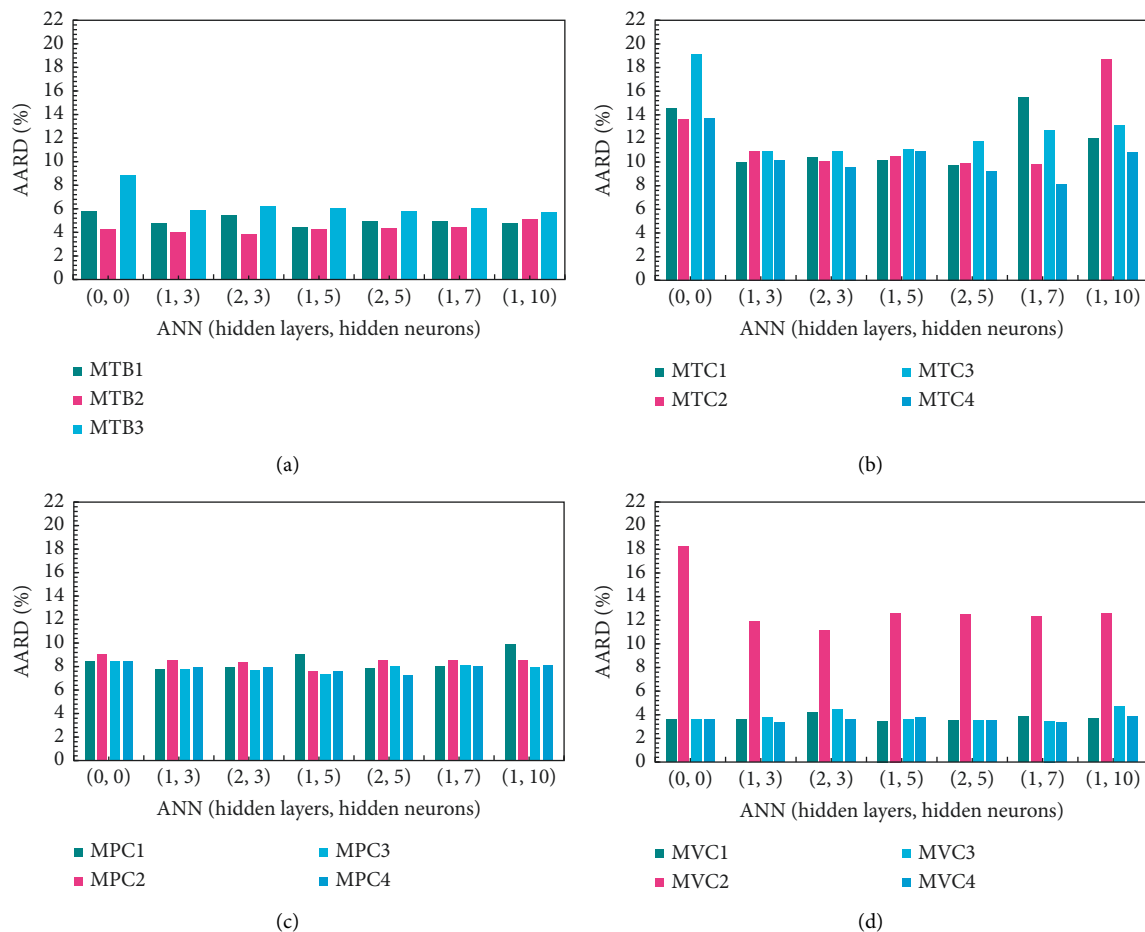


FIGURE 9: AARD for the calculation of normal boiling point, critical temperature, critical pressure, and critical volume of pure compounds using first-order GCMs, and ANN. (a) Normal boiling point, (b) critical temperature, (c) critical pressure, and (d) critical volume.

TABLE 4: Comparison of the performance of different GCMs in the calculation of critical properties and normal boiling point of pure compounds.

Statistical metric, %	MTB2	MTB2 + ANN (2,3)	JR	CG	LS	F	A	R
AARD	4.24	3.90	5.10	7.93	—	—	—	—
MaxARD _i	28.53	32.06	48.95	100.42	—	—	—	—
Min ARD _i	0.29	0.18	0.21	0.07	—	—	—	—
σ	5.24	5.50	7.72	15.80	—	—	—	—
	MTC4	MTC4 + ANN (1,7)						
AARD	9.82	8.15	5.82	6.19	6.42	14.05	3.94	—
MaxARD _i	56.72	60.53	128.59	35.15	182.95	79.90	40.76	—
Min ARD _i	0.25	4.31E-03	0.11	0.14	0.16	0.12	0.01	—
σ	10.38	10.12	15.22	6.50	19.20	16.95	5.95	—
	MPC4	MPC4 + ANN (2,5)						
AARD	8.14	7.28	7.92	8.38	10.44	—	8.93	—
MaxARD _i	46.17	45.70	48.64	36.77	34.79	—	35.70	—
Min ARD _i	0.49	0.31	0.62	0.02	0.22	—	0.06	—
σ	9.23	7.82	9.30	8.09	8.47	—	9.11	—
	MVC4	MVC4 + ANN (1,3)						
AARD	3.12	3.34	4.61	7.03	4.92	4.92	—	7.40
MaxARD _i	10.00	17.08	32.11	33.51	32.60	32.70	—	33.79
Min ARD _i	0.09	0.04	0.08	0.00	0.07	0.02	—	0.31
σ	2.75	3.38	5.72	6.06	6.07	5.66	—	7.66

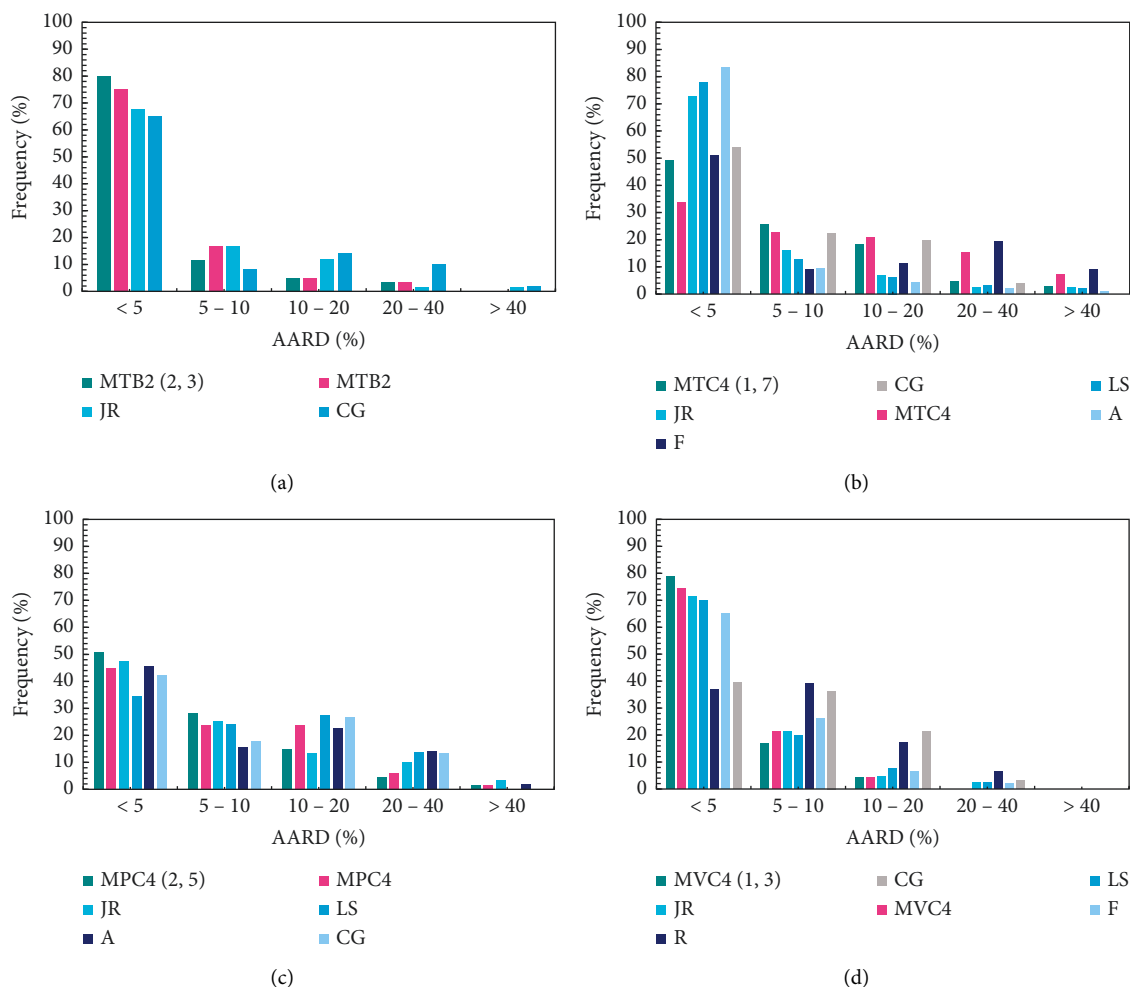


FIGURE 10: Distribution of AARD for the calculation of normal boiling point, critical temperature, critical pressure, and critical volume of pure compounds using first-order GCMs, and ANN. (a) Normal boiling point, (b) critical temperature, (c) critical pressure, and (d) critical volume.

better predictions of V_c than those obtained with the models of Riedel [52], Lydersen [23], Fedors [55], Constantinou, and Gani [12], and Joback and Reid [11].

Finally, the performance of $MTC4$, $MPC4$, and $MVC4$ was compared with the ANN model proposed by Gharagheizi et al. [6]. This ANN-based model has a complex network configuration with a substantial number of weights and biases parameters (i.e., more than 2000 parameters of ANN) and, as expected, its accuracy for the prediction of these thermodynamic properties is high. For this comparison, the critical properties of 200 additional molecules reported by Gharagheizi et al. [6] were estimated using the three GCMs. Figure 11 shows the percentage deviation between the predictions generated by these models taking as a reference the result obtained with the model of Gharagheizi et al. [6]. In general, most of the estimated properties differed by less than 10% from those calculations with the model of Gharagheizi et al. [6]. It was clear that the first-order GCMs (i.e., $MTC4$, $MPC4$, and $MVC4$) can predict the critical properties of pure compounds with acceptable accuracy compared to this complex ANN model. Note that these results also highlighted the importance of

utilizing a reliable approach to develop and implement GCMs for the calculation and prediction of pure compound thermodynamic properties.

3.7. Perspectives and Recommendations for Future Applications of GCMs for the Calculation of Critical Properties and Normal Boiling Point of Pure Compounds. The estimation of critical properties of pure compounds using predictive models is challenging. First-order GCMs are a suitable and straightforward alternative to perform these calculations. However, these models can show high errors to calculate the thermodynamic properties of compounds with high molecular weight besides the contribution values for some functional groups are not reported. These drawbacks are strongly related to the lack of experimental information on the critical properties reported in the literature and to the uncertainties in the databases utilized to obtain the group contribution values.

Some efforts still should be performed to improve the performance of first-order GCMs to obtain reliable estimations of critical properties of pure compounds. These attempts could be focused on the following directions:

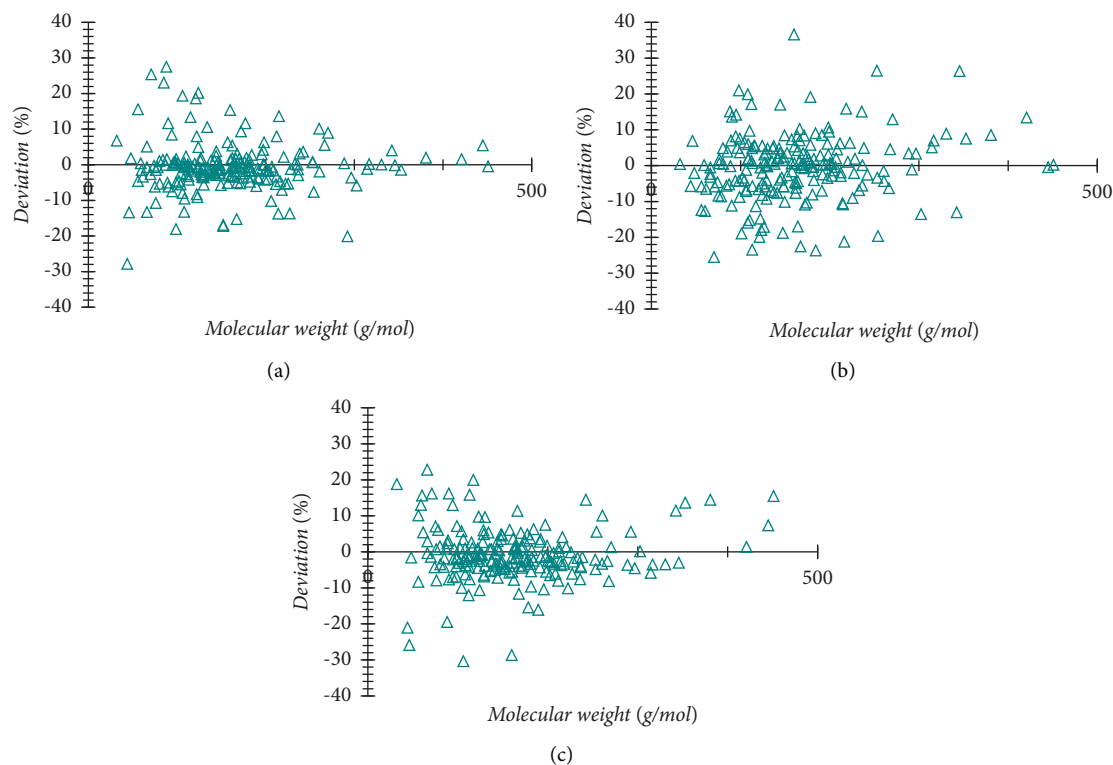


FIGURE 11: Deviations between the calculations of MTC4, MPC4, and MVC4 models and the model proposed by Gharagheizi et al. [6]. (a) Critical temperature, (b) critical pressure, and (c) critical volume.

- (a) It is important to establish proper experimental strategies to quantify the critical properties of novel compounds and those with limited experimental data to create a reliable experimental thermodynamic database. This aspect is essential for the development of reliable GCMs with a wide applicability spectrum to calculate the thermophysical properties of pure compounds. In this direction, it is recommended to recalculate the group contributions of available GCMs after incorporating new experimental information.
- (b) It is also suggested to improve the strategies based on hybrid models GCM + ANN to estimate critical properties of pure compounds. Other artificial intelligence tools could also be tested for these calculations.

4. Summary and Conclusions

This study has covered the next topics:

A comparative analysis of different first-order GCMs for the calculation of normal boiling point and critical properties of pure compounds was performed.

The incorporation of parameters associated with molecular characteristics in these thermodynamic models (e.g., molecular weight, number of atoms, and rings) contributed to improve their performance.

The critical temperature was a challenging property to be calculated, while GCMs were more accurate to estimate the critical volume.

These GCMs showed higher modeling errors in the thermodynamic property calculation of nitrogen, alkanes, halogenated, and silicon compounds. The highest deviations were also obtained in the estimation of critical properties and normal boiling points of molecules with low and high molecular weights within a homologous series, which usually did not follow the trend of property behavior of the chemical family under analysis.

The incorporation of an ANN as a corrector to improve the estimations obtained with GCMs is an interesting approach for the simulation of thermodynamic properties. This artificial intelligence algorithm can be used to enhance the capabilities of models proposed in the literature including those that are available and implemented in commercial simulation software.

Results also showed that the predictive models incorporated in the commercial Aspen Plus® software can provide inaccurate estimations of the critical thermodynamic properties for some compounds.

This study contributes with straightforward first-order GCMs that can be employed to calculate the critical volume and pressure besides the normal boiling point of pure compounds with better accuracy than that obtained with other models reported in the literature even those incorporated in commercial process simulators.

Nomenclature

AARD:	Average absolute relative deviation
$A_i, B_i, C_i,$	Constant parameters of the models
$D_i:$	
ARD _i :	Absolute relative deviation
$m:$	Number of experimental data
MM:	Molecular weight of the pure compound
MM _k :	Molecular weight of the functional group k
$N_k:$	Number of repetitions of the functional group k in the molecule of tested compound
$N_r:$	Number of rings in the molecule
$n_t:$	Number of atoms in the molecule (except hydrogen)
$P_c:$	Critical pressure
$p_{ck}:$	Contribution of the functional group k to estimate the critical pressure
$R^2:$	Determination coefficient
$\sigma:$	Standard deviation
$T_b:$	Normal boiling point
$t_{bk}:$	Contribution of the functional group k to estimate the normal boiling point
$T_c:$	Critical temperature
$t_{ck}:$	Contribution of the functional group k to estimate the critical temperature
$V_c:$	Critical volume
$v_{ck}:$	Contribution of the functional group k to estimate the critical volume
$Y_i^{\text{exp}}:$	Experimental values of tested thermodynamic properties (i.e., V_c, P_c, T_c, T_b)
$Y_i^{\text{cal}}:$	Calculated values of tested thermodynamic properties (i.e., V_c, P_c, T_c, T_b).

Data Availability

The data of this paper are available on request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This article was prepared with funding provided by the Instituto Tecnológico de Aguascalientes, Mexico.

Supplementary Materials

Supplementary information of the calculations reported in this paper is provided in the supporting information file. It includes results and data of the implementation and analysis of tested group contribution models. (*Supplementary Materials*)

References

- [1] J. Marrero and R. Gani, "Group-contribution based estimation of pure component properties," *Fluid Phase Equilibria*, vol. 183, pp. 183–208, 2001.
- [2] D. Ambrose and C. L. Young, "Vapor-liquid critical properties of elements and compounds. 1. An introductory survey," *Journal of Chemical and Engineering Data*, vol. 40, pp. 345–357, 1995.
- [3] C.-B. Soo, P. Théveneau, C. Coquelet, D. Ramjugernath, and D. Richon, "Determination of critical properties of pure and multi-component mixtures using a "dynamic-synthetic" apparatus," *The Journal of Supercritical Fluids*, vol. 55, no. 2, pp. 545–553, 2010.
- [4] K. Ghandi, "A review of ionic liquids, their limits and applications," *Green and Sustainable Chemistry*, vol. 04, no. 01, pp. 44–53, 2014.
- [5] S. Tahami, K. Movagharnejad, and H. Ghasemitabar, "Estimation of the critical constants of organic compounds via a new group contribution method," *Fluid Phase Equilibria*, vol. 494, pp. 45–60, 2019.
- [6] F. Gharagheizi, A. Eslamimanesh, A. H. Mohammadi, and D. Richon, "Determination of critical properties and acentric factors of pure compounds using the artificial neural network group contribution algorithm," *Journal of Chemical & Engineering Data*, vol. 56, no. 5, pp. 2460–2476, 2011.
- [7] M. Sattari, A. Kamari, A. H. Mohammadi, and D. Ramjugernath, "On the prediction of critical temperatures of ionic liquids: model development and evaluation," *Fluid Phase Equilibria*, vol. 411, pp. 24–32, 2016.
- [8] M. E. Mondejar, S. Cignitti, J. Abildskov, J. M. Woodley, and F. Haglind, "Prediction of properties of new halogenated olefins using two group contribution approaches," *Fluid Phase Equilibria*, vol. 433, pp. 79–96, 2017.
- [9] W. Su, L. Zhao, and S. Deng, "Group contribution methods in thermodynamic cycles: physical properties estimation of pure working fluids," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 984–1001, 2017.
- [10] S. M. Hosseini, A. Mulero, and M. M. Alavianmehr, "Predictive methods and semi-classical Equations of State for pure ionic liquids: a review," *The Journal of Chemical Thermodynamics*, vol. 130, pp. 47–94, 2019.
- [11] K. G. Joback and R. C. Reid, "Estimation of pure-component properties from group-contributions," *Chemical Engineering Communications*, vol. 57, no. 1-6, pp. 233–243, 1987.
- [12] L. Constantinou and R. Gani, "New group contribution method for estimating properties of pure compounds," *AIChE Journal*, vol. 40, no. 10, pp. 1697–1710, 1994.
- [13] G. M. Kontogeorgis and D. P. Tassios, "Critical constants and acentric factors for long-chain alkanes suitable for corresponding states applications. A critical review," *Chemical Engineering Journal*, vol. 66, no. 1, pp. 35–49, 1997.
- [14] J. Marrero-Morejón and E. Pardillo-Fontdevila, "Estimation of pure compound properties using group-interaction contributions," *AIChE Journal*, vol. 45, no. 3, pp. 615–621, 1999.
- [15] D. Dalmazzone, A. Salmon, and S. Guella, "A second order group contribution method for the prediction of critical temperatures and enthalpies of vaporization of organic compounds," *Fluid Phase Equilibria*, vol. 242, no. 1, pp. 29–42, 2006.
- [16] Y. Nannoolal, J. Rarey, and D. Ramjugernath, "Estimation of pure component properties: Part 2. Estimation of critical property data by group contribution," *Fluid Phase Equilibria*, vol. 252, no. 1-2, pp. 1–27, 2007.
- [17] S. S. Godavarthy, R. L. Robinson, and K. A. M. Gasem, "Improved structure-property relationship models for prediction of critical properties," *Fluid Phase Equilibria*, vol. 264, no. 1-2, pp. 122–136, 2008.

- [18] D. Sola, A. Ferri, M. Banchemo, L. Manna, and S. Sicardi, "QSPR prediction of N-boiling point and critical properties of organic compounds and comparison with a group-contribution method," *Fluid Phase Equilibria*, vol. 263, no. 1, pp. 33–42, 2008.
- [19] A. Kazakov, C. D. Muzny, V. Diky, R. D. Chirico, and M. Frenkel, "Predictive correlations based on large experimental datasets: critical constants for pure compounds," *Fluid Phase Equilibria*, vol. 298, no. 1, pp. 131–142, 2010.
- [20] M. A. Sobati and D. Aboali, "Molecular based models for estimation of critical properties of pure refrigerants: quantitative structure property relationship (QSPR) approach," *Thermochimica Acta*, vol. 602, pp. 53–62, 2015.
- [21] A. Varamesh, A. Hemmati-Sarapardeh, M. K. Moraveji, and A. H. Mohammadi, "Generalized models for predicting the critical properties of pure chemical compounds," *Journal of Molecular Liquids*, vol. 240, pp. 777–793, 2017.
- [22] J. O. Valderrama, L. A. Forero, and R. E. Rojas, "Critical properties and normal boiling temperature of ionic liquids. Update and a New Consistency Test," *Industrial & Engineering Chemistry Research*, vol. 51, no. 22, pp. 7838–7844, 2012.
- [23] A. L. Lydersen, *Estimation of Critical Properties of Organic Compounds by the Method of Group Contributions*, University of Wisconsin, Engineering Experiment Station, Madison, WI, USA, 1955.
- [24] Y. Nannoolal, J. Rarey, D. Ramjugernath, and W. Cordes, "Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions," *Fluid Phase Equilibria*, vol. 226, pp. 45–63, 2004.
- [25] Q. Wang, P. Ma, Q. Jia, and S. Xia, "Position group contribution method for the prediction of critical temperatures of organic compounds," *Journal of Chemical & Engineering Data*, vol. 53, no. 5, pp. 1103–1109, 2008.
- [26] Q. Wang, Q. Jia, and P. Ma, "Position group contribution method for the prediction of critical pressure of organic compounds," *Journal of Chemical & Engineering Data*, vol. 53, no. 8, pp. 1877–1885, 2008.
- [27] Q. Jia, Q. Wang, and P. Ma, "Position group contribution method for the prediction of critical volume of organic compounds," *Journal of Chemical & Engineering Data*, vol. 53, no. 11, pp. 2606–2612, 2008.
- [28] E. Retzekas, E. Voutsas, K. Magoulas, and D. Tassios, "Prediction of physical properties of hydrocarbons, petroleum, and coal liquid fractions," *Industrial & Engineering Chemistry Research*, vol. 41, no. 6, pp. 1695–1702, 2002.
- [29] H. Ghasemtabar and K. Movagharnejad, "Estimation of the normal boiling point of organic compounds via a new group contribution method," *Fluid Phase Equilibria*, vol. 411, pp. 13–23, 2016.
- [30] G. Sadeghi, M. Najafzadeh, and H. Safarzadeh, "Utilizing gene-expression programming in modelling the thermal performance of evacuated tube solar collectors," *Journal of Energy Storage*, vol. 30, Article ID 101546, 2020.
- [31] G. Sadeghi, M. Najafzadeh, and M. Ameri, "Thermal characteristics of evacuated tube solar collectors with coil inside: an experimental study and evolutionary algorithms," *Renewable Energy*, vol. 151, pp. 575–588, 2020.
- [32] S. Nazari, M. Najafzadeh, and R. Daghigh, "Techno-economic estimation of a non-cover box solar still with thermoelectric and antiseptic nanofluid using machine learning models," *Applied Thermal Engineering*, vol. 212, Article ID 118584, 2022.
- [33] G. Sadeghi, M. Najafzadeh, M. Ameri, and M. Jowzi, "A case study on copper-oxide nanofluid in a back pipe vacuum tube solar collector accompanied by data mining techniques," *Case Studies in Thermal Engineering*, vol. 32, Article ID 101842, 2022.
- [34] S. Deng, W. Su, and L. Zhao, "A neural network for predicting normal boiling point of pure refrigerants using molecular groups and a topological index," *International Journal of Refrigeration*, vol. 63, pp. 63–71, 2016.
- [35] J. O. Valderrama, L. A. Forero, and R. E. Rojas, "Extension of a group contribution method to estimate the critical properties of ionic liquids of high molecular mass," *Industrial & Engineering Chemistry Research*, vol. 54, no. 13, pp. 3480–3487, 2015.
- [36] D. Ambrose and C. Tsonopoulos, "Vapor-liquid critical properties of elements and compounds. 2. Normal alkanes," *Journal of Chemical & Engineering Data*, vol. 40, no. 3, pp. 531–546, 1995.
- [37] C. Tsonopoulos and D. Ambrose, "Vapor-liquid critical properties of elements and compounds. 3. Aromatic hydrocarbons," *Journal of Chemical & Engineering Data*, vol. 40, pp. 547–558, 1995.
- [38] M. Gude and A. S. Teja, "Vapor-liquid critical properties of elements and compounds. 4. Aliphatic alkanols," *Journal of Chemical & Engineering Data*, vol. 40, no. 5, pp. 1025–1036, 1995.
- [39] T. E. Daubert, "Vapor-liquid critical properties of elements and compounds. 5. Branched alkanes and cycloalkanes," *Journal of Chemical & Engineering Data*, vol. 41, no. 3, pp. 365–372, 1996.
- [40] C. Tsonopoulos and D. Ambrose, "Vapor-liquid critical properties of elements and compounds. 6. Unsaturated aliphatic hydrocarbons," *Journal of Chemical & Engineering Data*, vol. 41, no. 4, pp. 645–656, 1996.
- [41] A. P. Kudchadker, D. Ambrose, and C. Tsonopoulos, "Vapor-liquid critical properties of elements and compounds. 7. Oxygen compounds other than alkanols and cycloalkanol," *Journal of Chemical & Engineering Data*, vol. 46, no. 3, pp. 457–479, 2001.
- [42] C. Tsonopoulos and D. Ambrose, "Vapor-liquid critical properties of elements and compounds. 8. Organic sulfur, silicon, and tin compounds (C + H + S, Si, and Sn)," *Journal of Chemical & Engineering Data*, vol. 46, no. 3, pp. 480–485, 2001.
- [43] K. N. Marsh, C. L. Young, D. W. Morton, D. Ambrose, and C. Tsonopoulos, "Vapor-liquid critical properties of elements and compounds. 9. Organic compounds containing nitrogen," *Journal of Chemical & Engineering Data*, vol. 55, no. 3, p. 1459, 2010.
- [44] K. N. Marsh, A. Abramson, D. Ambrose et al., "Vapor-liquid critical properties of elements and compounds. 10. Organic compounds containing halogens," *Journal of Chemical & Engineering Data*, vol. 52, no. 5, pp. 1509–1538, 2007.
- [45] D. Ambrose, C. Tsonopoulos, and E. D. Nikitin, "Vapor-liquid critical properties of elements and compounds. 11. Organic compounds containing B + O; halogens +N, +O, +O + S, +S, +Si; N + O; and O + S, +Si," *Journal of Chemical & Engineering Data*, vol. 54, no. 3, pp. 669–689, 2009.
- [46] D. Ambrose, C. Tsonopoulos, E. D. Nikitin, D. W. Morton, and K. N. Marsh, "Vapor-liquid critical properties of elements and compounds. 12. Review of recent data for hydrocarbons and non-hydrocarbons," *Journal of Chemical & Engineering Data*, vol. 60, pp. 3444–3482, 2015.
- [47] "KDB, Korean Data Base," 2003, <https://www.therc.org/kdb/>.

- [48] B. E. Poling, J. M. Prausnitz, and J. P. O'Connell, *The Properties of Gases and Liquids*, McGraw-Hill, New York, NY, USA, 2001.
- [49] C. L. Yaws, *Thermophysical Properties of Chemicals and Hydrocarbons*, Elsevier, Amsterdam, Netherlands, 2014.
- [50] X. Wen and Y. Qiang, "A new group contribution method for estimating critical properties of organic compounds," *Industrial & Engineering Chemistry Research*, vol. 40, no. 26, pp. 6245–6250, 2001.
- [51] J. O. Valderrama and V. H. Alvarez, "A new group contribution method based on equation of state parameters to evaluate the critical properties of simple and complex molecules," *Canadian Journal of Chemical Engineering*, vol. 84, no. 4, pp. 431–446, 2008.
- [52] L. Riedel, "Kritischer Koeffizient, Dichte des gesättigten Dampfes und Verdampfungswärme. Untersuchungen über eine Erweiterung des Theorems der übereinstimmenden Zustände, Teil III," *Chemie Ingenieur Technik*, vol. 26, no. 12, pp. 679–683, 1954.
- [53] D. Ambrose, *Correlation and Estimation of Vapour-Liquid Critical Properties: I. Critical Temperatures of Organic Compounds*, National Physical Laboratory Reports Chemistry, vol. 92, 1978.
- [54] D. Ambrose, "Correlation and estimation of vapor-liquid critical properties: II. Critical pressures and volumes of organic compounds," *National Physical Laboratory Reports Chemistry*, vol. 98, 1979.
- [55] R. F. Fedors, "A method to estimate critical volumes," *American Institute of Chemical Engineers Journal*, vol. 25, p. 202, 1979.
- [56] R. F. Fedors, "A relationship between chemical structure and the critical temperature," *Chemical Engineering Communications*, vol. 16, no. 1-6, pp. 149–151, 1982.