

QSPR Analysis with Curvilinear Regression Modeling and Temperature-based Topological Indices

H. M. Nagesh

Department of Science and Humanities, PES University, Bangalore, India.
E-mail: nageshnm@pes.edu

Abstract

Establishing quantitative correlations between various molecular properties and chemical structures is of great technological importance for environmental and medical aspects. These approaches are referred to as Quantitative Structure-Property Relationships (QSPR), which relate the physicochemical or thermodynamic properties of compounds to their structures. The main goal of QSPR studies is to find a mathematical relationship between the property of interest and several molecular descriptors derived from the structure of the molecule. Topological indices are the molecular descriptors that characterize the formation of chemical compounds and predict certain physicochemical properties. In this study, the QSPR models are designed using certain temperature-based topological indices such as the sum connectivity temperature index, product connectivity temperature index, F-temperature index, and symmetric division temperature index to predict the thermodynamic properties, such as enthalpies of formation (ΔH_f^0 liquid), enthalpies of combustion (ΔH_c^0 liquid), and enthalpies of vaporization (ΔH_{vap}^0 gas) of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$). The relationship analysis between thermodynamic properties and topological indices is done using linear, quadratic, and cubic equations of a curvilinear regression model. These regression models are then compared.

Keywords: Temperature of a vertex, sum connectivity temperature index, product connectivity temperature index, F-temperature index, symmetric division temperature index.

1 Introduction

The properties of a molecule are indeed closely tied to its structural characteristics and composition. This concept is fundamental to understanding how molecules behave and interact in various chemical reactions and physical processes. In this connection, graph theory has been successfully applied and some thermodynamic properties [1, 2, 6, 13, 17].

Chemical graph theory is a specialized field within mathematical chemistry that focuses on the study of molecules and chemical systems through the lens of graph theory.

Graph theory provides a mathematical framework for analyzing the relationships between objects connected by edges, and in the context of chemical graph theory, these objects are atoms and the edges represent chemical bonds. Chemical graph theory has practical applications in fields such as drug discovery, materials science, computational chemistry, and chemical informatics. It provides a powerful approach to understanding the relationships between molecular structure and properties, which is crucial for designing new molecules with desired characteristics and predicting how molecules will behave under different conditions.

Chemical graph theory plays a crucial role in developing QSPR models. These models correlate graph-based molecular descriptors with various properties such as boiling points, melting points, solubility, etc. For more details, the reader can refer to [3, 7, 9, 10, 11, 12, 15]. Numerous studies have been made relating to QSPR models by using what are called topological indices (TI). The first topological index was the Wiener index, which was introduced by Harold Wiener in 1947. It was used to determine the physical properties of paraffin [17]. Since then, many topological indices have been defined and used in many applications.

A topological index can be classified according to the structural characteristics of the graph such as the degree of a vertex, the distance between vertices, the matching, and the spectrum. The best-known topological indices are the Wiener index which is based on distance, the Zagreb and the Randić indices which are based on degree, the Estrada index which is based on the spectrum of a graph, the Hosoya index based on the matching.

Shafiei in [14] designed the QSPR models using topological indices such as the connectivity index, Szeged index, Balaban index, and Harary number to predict the thermodynamic properties such as enthalpies of formation of liquid, enthalpies of combustion of liquid, enthalpies of vaporization, and enthalpies of sublimation of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$). Later in [5], Havare designed the QSPR models using topological indices such as the Gutman index, variance of degree index, product connectivity Banhatti index, and Sigma index to predict these thermodynamic properties of monocarboxylic acids.

Motivated by these, the structure-property relationship between certain temperature-based topological indices such as sum connectivity temperature index, product connectivity temperature index, F-temperature index, and symmetric division temperature index to the enthalpies of formation (ΔH_f^0 liquid), enthalpies of combustion (ΔH_C^0 liquid), and enthalpies of vaporization (ΔH_{vap}^0 gas) of monocarboxylic acids and their quantitative structure-property relationship are presented in this paper.

2 Basic definitions

Let $G = (V, E)$ be a simple connected graph with vertex set $|V(G)| = n$ and edge set $E(G)$. The number of edges incident to a vertex v is called the *degree* of the vertex v and is denoted by $deg(v)$. Fajtolowicz [4] introduced the notion of the temperature of a vertex v as follows.

Definition 2.1 The *temperature* of a vertex $v \in V(G)$ is defined by $\mathcal{T}(v) = \frac{deg(v)}{n-deg(v)}$

Later in 2019, Kulli [8] introduced the concept of sum connectivity temperature index, product connectivity temperature index, F-temperature index, and symmetric division temperature index of a graph G as follows:

Sum connectivity temperature index:

$$ST(G) = \sum_{uv \in E(G)} \frac{1}{\sqrt{\mathcal{T}(u) + \mathcal{T}(v)}} \quad (1)$$

Product connectivity temperature index:

$$PT(G) = \sum_{uv \in E(G)} \frac{1}{\sqrt{\mathcal{T}(u) \times \mathcal{T}(v)}} \quad (2)$$

F-temperature index:

$$FT(G) = \sum_{uv \in E(G)} (\mathcal{T}(u)^2 + \mathcal{T}(v)^2) \quad (3)$$

Symmetric division temperature index:

$$SDT(G) = \sum_{uv \in E(G)} \left(\frac{\mathcal{T}(u)}{\mathcal{T}(v)} + \frac{\mathcal{T}(v)}{\mathcal{T}(u)} \right) \quad (4)$$

3 Methods and techniques

The method used in this article includes finding the temperature of vertices, division of temperature of vertices, and partition of edges based on the temperature of end vertices. The temperature of vertices and the edge separation methods are used for the computation of temperature-based topological indices. The correlation coefficients are calculated using JMP statistical software. 2D graphs are drawn using JMP statistical software/Microsoft Excel.

4 2D molecular structures and computations of monocarboxylic acids

With the help of these four temperature-based topological indices, the molecular structures of 19 monocarboxylic acids are explored. There are three thermodynamic properties of monocarboxylic acids under observation. These properties are enthalpies of formation (ΔH_f^0 liquid), enthalpies of combustion (ΔH_C^0 liquid), and enthalpies of vaporization (ΔH_{vap}^0 gas). Thermodynamic properties of monocarboxylic acids as given in Table 1 are taken from [5, 14].

Let α =Enthalpies of formation of liquids; β =Enthalpies of combustion of liquids; and Γ =Enthalpies of vaporization. Then,

Name of compounds	Formula	α	β	Γ
Acetic acid	$C_2H_4O_2$	483.50	875.16	46.3
Propanoic acid	$C_3H_6O_2$	510.8	1527.3	50
Butanoic acid	$C_4H_8O_2$	533.9	2183.5	54.9
Pentanoic acid	$C_5H_{10}O_2$	558.9	2837.8	58.2
Hexanoic acid	$C_6H_{12}O_2$	581.8	3494.3	63
Heptanoic acid	$C_7H_{14}O_2$	608.5	4146.9	64.8
Octanoic acid	$C_8H_{16}O_2$	634.8	4799.9	69.4
Nonanoic acid	$C_9H_{18}O_2$	658	5456.1	72.3
Decanoic acid	$C_{10}H_{20}O_2$	713.7	6079.3	76.6
Undecanoic acid	$C_{11}H_{22}O_2$	736.2	6376.5	78.9
Dodecanoic acid	$C_{12}H_{24}O_2$	775.1	7377	82.2
Tridecanoic acid	$C_{13}H_{26}O_2$	807.2	8024.2	84.9
Tetradecanoic acid	$C_{14}H_{28}O_2$	834.1	8676.7	87.7
Pentadecanoic acid	$C_{15}H_{30}O_2$	862.4	9327.7	91.4
Hexadecanoic acid	$C_{16}H_{32}O_2$	892.2	9977.2	94.5
Heptadecanoic acid	$C_{17}H_{34}O_2$	924.4	10624.4	100.7
Octadecanoic acid	$C_{18}H_{36}O_2$	947.2	11280.1	102.8
Nonadecanoic acid	$C_{19}H_{38}O_2$	984.1	11923.4	105
Eicosanoic acid	$C_{20}H_{40}O_2$	1012.6	12574.2	109.9

Table 1. The values of enthalpies of formation of liquid ($\Delta H_f^0 kJ/mol$), enthalpies of combustion of liquid ($\Delta H_C^0 kJ/mol$), and enthalpies of vaporization ($\Delta H_{vap}^0 kJ/mol$) of monocarboxylic acids at conditions, normally at 298.15 K, 1 atm.

5 Computation of temperature-based topological indices

Table 2 shows the temperature-based topological indices of 19 monocarboxylic acids calculated using the formulas (1-4).

Formula	$ST(G)$	$PT(G)$	$FT(G)$	$SDT(G)$
$C_2H_4O_2$	1.643167	3	27.3333333	27.3333333
$C_3H_6O_2$	3.23569	6.715476	7.826389	18.069445
$C_4H_8O_2$	4.837467	9.048627	4.12	17.8
$C_5H_{10}O_2$	6.585927	16.355579	2.730833	18.669444
$C_6H_{12}O_2$	8.478948	22.649944	2.030113	19.993658
$C_7H_{14}O_2$	10.97437	32.485281	1.421875	21.5
$C_8H_{16}O_2$	12.671578	38.220202	1.338057	23.224868
$C_9H_{18}O_2$	14.957559	47.500261	1.143533	24.985648
$C_{10}H_{20}O_2$	17.700444	59.821403	0.998127	26.8
$C_{11}H_{22}O_2$	19.879457	69.054894	0.885875	28.651768
$C_{12}H_{24}O_2$	22.505505	81.330388	0.796447	30.530691
$C_{13}H_{26}O_2$	25.235916	94.605041	0.723516	32.429945
$C_{14}H_{28}O_2$	28.066954	106.017218	0.662893	34.344811
$C_{15}H_{30}O_2$	30.995246	124.1525061	0.611696	36.271925
$C_{16}H_{32}O_2$	34.017723	140.425548	0.567881	38.208823
$C_{17}H_{34}O_2$	37.131587	157.698239	0.529953	39.581903
$C_{18}H_{36}O_2$	40.334265	175.9706351	0.496798	42.105034
$C_{19}H_{38}O_2$	43.62339	195.2427855	0.467565	44.061842
$C_{20}H_{40}O_2$	46.99677	215.514726	0.441595	47.229573

Table 2. The values of temperature-based topological indices of monocarboxylic acids.

6 Regression models

Regression analysis is a statistical method that shows the relationship between two or more variables.

To study the relationship between thermodynamic properties of monocarboxylic acids and the temperature-based topological indices, the following equations from [14] are used.

Linear equation:

$$Y = a + b_1X_1; \quad n, R^2, s, F$$

Quadratic equation:

$$Y = a + b_1X_1 + b_2X_1^2; \quad n, R^2, s, F$$

Cubic equation:

$$Y = a + b_1X_1 + b_2X_1^2 + b_3X_1^3; \quad n, R^2, s, F$$

Here, Y is the dependent variable, a is the regression model, b_i ($1 \leq i \leq 3$) are the coefficients for the individual descriptor, X_i ($1 \leq i \leq 3$) are the independent variables, n is the number of samples used for building the regression equation, R^2 is the correlation coefficient, s is the standard error of deviation, and F is the calculated value of the F-ration test.

Note that the quality of a QSPR model can be conveniently measured by the correlation coefficient (R^2). A good QSPR model must have $R^2 > 0.99$. The observed values and model predictions must be compared to measure the predictive quality of the model. So, we deal with the RMSE (Root Mean Square Error) metric for the predictive power of the model. The best predictive model is the minimum error, i.e. the minimum RMSE.

Furthermore, R^2 and F parameters will be considered for the goodness of fit of the model. As given in [16], the best goodness of fit in models is selected by using the parameters either $\max(R^2)$ or $\max(F)$.

We now determine the linear, quadratics, and cubic curvilinear models of the sum connectivity temperature index, product connectivity temperature index, F-temperature index, and symmetric division temperature index for ΔH_f^0 , ΔH_C^0 , and ΔH_{vap}^0 .

The significance of bold numbers in each table denote highest correlation value.

6.1 Curvilinear regression models of $ST(G)$, $PT(G)$, $FT(G)$, $SDT(G)$ for ΔH_f^0

Table 3. The curvilinear regressions models of $ST(G)$ index for ΔH_f^0 .

ΔH_f^0	R^2	F	RMSE
$485.412+11.8001(ST(G))$	0.9918	2056.329	15.8062
$456.784+15.606(ST(G))-0.08110(ST(G)^2)$	0.9988	6658.513	6.2330
$460.745+14.626(ST(G))-0.0296(ST(G)^2) -0.000717(ST(G)^3)$	0.9988	4449.708	6.2257

The analysis as mentioned in Table 3 indicates that the best goodness of fit among obtained curvilinear equations using $ST(G)$ topological index for ΔH_f^0 are:

$$\Delta H_f^0 = 456.784 + 15.606(ST(G)) - 0.08110(ST(G)^2); \text{ and}$$

$$\Delta H_f^0 = 460.745 + 14.626(ST(G)) - 0.0296(ST(G)^2) - 0.000717(ST(G)^3).$$

Table 4. The curvilinear regressions models of $PT(G)$ index for ΔH_f^0 .

ΔH_f^0	R^2	F	RMSE
$532.892+2.465(PT(G))$	0.9662	487.167	32.053
$489.678+4.024(PT(G))-0.0077(PT(G)^2)$	0.9972	2950.84	9.355
$479.340+4.754(PT(G))-0.0168(PT(G)^2) -2.995e-5(PT(G)^3)$	0.9986	3753.734	6.777

The analysis as mentioned in Table 4 indicates that the best goodness of fit among obtained curvilinear equations using $PT(G)$ topological index for ΔH_f^0 is:

$$\Delta H_f^0 = 479.340 + 4.754(PT(G)) - 0.0168(PT(G)^2) - 2.995e-5(PT(G)^3)$$

Table 5. The curvilinear regressions models of $FT(G)$ index for ΔH_f^0 .

ΔH_f^0	R^2	F	RMSE
$782.969-14.820(FT(G))$	0.2913	6.9901	146.942
$873.844-87.507(FT(G))+2.692(FT(G)^2)$	0.6220	13.1658	110.619
$1013.011-271.1496(FT(G))+34.177(FT(G)^2)-0.9134(FT(G)^3)$	0.8510	28.574	71.713

The analysis as mentioned in Table 5 indicates that the best goodness of fit among obtained curvilinear equations using $FT(G)$ topological index for ΔH_f^0 is:

$$\Delta H_f^0 = 1013.011 - 271.1496(FT(G)) + 34.177(FT(G)^2) - 0.9134(FT(G)^3).$$

Table 6. The curvilinear regressions models of $SDT(G)$ index for ΔH_f^0 .

ΔH_f^0	R^2	F	RMSE
$224.623+17.044(SDT(G))$	0.8630	107.0981	64.607
$276.589+13.4216(SDT(G))+0.0578(SDT(G)^2)$	0.8636	50.6906	66.429
$653.886-26.121(SDT(G))+1.3603(SDT(G)^2)-0.0135(SDT(G)^3)$	0.8658	32.275	68.059

The analysis as mentioned in Table 6 indicates that the best goodness of fit among obtained curvilinear equations using $SDT(G)$ topological index for ΔH_f^0 is:

$$\Delta H_f^0 = 653.886 - 26.121(SDT(G)) + 1.3603(SDT(G)^2) - 0.0135(SDT(G)^3).$$

From the analysis as mentioned in Table 3 - Table 6, the topological indices $ST(G)$ and $PT(G)$ are the best suitable for predicting the ΔH_f^0 of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$) since $R^2 > 0.99$.

6.2 Curvilinear regression models of $ST(G)$, $PT(G)$, $FT(G)$, $SDT(G)$ for ΔH_C^0

Table 7. The curvilinear regressions models of $ST(G)$ index for ΔH_C^0 .

ΔH_C^0	R^2	F	RMSE
$1258.293+253.772(ST(G))$	0.9889	1526.730	394.502
$512.470+352.946(ST(G))-2.1129(ST(G)^2)$	0.9992	1038.96	107.495
$274.317+411.859(ST(G))-5.2041(ST(G)^2) +0.0431(ST(G)^3)$	0.9998	3037.92	51.341

The analysis as mentioned in Table 7 indicates that the best goodness of fit among obtained curvilinear equations using $ST(G)$ topological index for ΔH_C^0 is:

$$\Delta H_C^0 = 274.317 + 411.859(ST(G)) - 5.2041(ST(G)^2) + 0.0431(ST(G)^3).$$

Table 8. The curvilinear regressions models of $PT(G)$ index for ΔH_C^0 .

ΔH_C^0	R^2	F	RMSE
$2288.883+52.909(PT(G))$	0.9594	402.091	757.149
$1304.786+88.4016(PT(G))-0.1753(PT(G)^2)$	0.9941	1351.173	297.2921
$920.142+115.588(PT(G))-0.5165(PT(G)^2) +0.0010(PT(G)^3)$	0.9982	2795.027	169.119

The analysis as mentioned in Table 8 indicates that the best goodness of fit among obtained curvilinear equations using $PT(G)$ topological index for ΔH_C^0 is:
 $\Delta H_C^0 = 1304.786 + 88.4016(PT(G)) - 0.1753(PT(G)^2)$.

Table 9. The curvilinear regressions models of $FT(G)$ index for ΔH_C^0 .

ΔH_C^0	R^2	F	RMSE
$7702.569-334.266(FT(G))$	0.3195	7.9835	3101.051
$9679.411-1915.462(FT(G))+58.562(FT(G)^2)$	0.6569	15.3177	2269.749
$12543.022-5694.203(FT(G))+706.436(FT(G)^2)-18.795(FT(G)^3)$	0.8659	32.312	1465.029

The analysis as mentioned in Table 9 indicates that the best goodness of fit among obtained curvilinear equations using $FT(G)$ topological index for ΔH_C^0 is:
 $\Delta H_C^0 = 12543.022 - 5694.203(FT(G)) + 706.436(FT(G)^2) - 18.795(FT(G)^3)$.

Table 10. The curvilinear regressions models of $SDT(G)$ index for ΔH_C^0 .

ΔH_C^0	R^2	F	RMSE
$-4241.663+362.973(SDT(G))$	0.8437	91.831	1485.794
$-3123.265+284.995(SDT(G))+1.2453(SDT(G)^2)$	0.8444	43.4386	1528.186
$1084.461-156.0012(SDT(G))+15.7711(SDT(G)^2)-0.1515(SDT(G)^3)$	0.8450	27.269	1575.346

The analysis as mentioned in Table 10 indicates that the best goodness of fit among obtained curvilinear equations using $SDT(G)$ topological index for ΔH_C^0 is:
 $\Delta H_C^0 = 1084.461 - 156.0012(SDT(G)) + 15.7711(SDT(G)^2) - 0.1515(SDT(G)^3)$.

From the analysis as mentioned in Table 7 - Table 10, the topological indices $ST(G)$ and $PT(G)$ are the best suitable for predicting the ΔH_C^0 of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$) since $R^2 > 0.99$.

6.3 Curvilinear regression models of $\mathcal{ST}(G)$, $\mathcal{PT}(G)$, $\mathcal{FT}(G)$, $\mathcal{SDT}(G)$ for ΔH_{vap}^0

Table 11. The curvilinear regressions models of $\mathcal{ST}(G)$ index for ΔH_{vap}^0 .

ΔH_{vap}^0	R^2	F	RMSE
$49.7480+1.3369(\mathcal{ST}(G))$	0.9835	1015.590	2.5482
$45.531+1.897(\mathcal{ST}(G))-0.0119(\mathcal{ST}(G)^2)$	0.9952	1681.431	1.4087
$42.932+2.540(\mathcal{ST}(G))-0.0456(\mathcal{ST}(G)^2) +0.000470(\mathcal{ST}(G)^3)$	0.9978	2317.041	0.9811

The analysis as mentioned in Table 11 indicates that the best goodness of fit among obtained curvilinear equations using $\mathcal{ST}(G)$ topological index for ΔH_{vap}^0 is:
 $\Delta H_{vap}^0 = 42.932 + 2.540(\mathcal{ST}(G)) - 0.0456(\mathcal{ST}(G)^2) + 0.000470(\mathcal{ST}(G)^3)$.

Table 12. The curvilinear regressions models of $\mathcal{PT}(G)$ index for ΔH_{vap}^0 .

ΔH_{vap}^0	R^2	F	RMSE
$55.201+0.278(\mathcal{PT}(G))$	0.9522	338.806	4.341
$49.973+0.4670(\mathcal{PT}(G))-0.00039(\mathcal{PT}(G)^2)$	0.9872	621.4481	2.3079
$47.306+0.655(\mathcal{PT}(G))-0.0032(\mathcal{PT}(G)^2) +7.510(\mathcal{PT}(G)^3)$	0.9943	880.0532	1.5891

The analysis as mentioned in Table 12 indicates that the best goodness of fit among obtained curvilinear equations using $\mathcal{PT}(G)$ topological index for ΔH_{vap}^0 is:
 $\Delta H_{vap}^0 = 47.306 + 0.655(\mathcal{PT}(G)) - 0.0032(\mathcal{PT}(G)^2) + 7.510(\mathcal{PT}(G)^3)$.

Table 13. The curvilinear regressions models of $\mathcal{FT}(G)$ index for ΔH_{vap}^0 .

ΔH_{vap}^0	R^2	F	RMSE
$83.876-1.822(\mathcal{FT}(G))$	0.3402	8.7681	16.1314
$94.306-10.164(\mathcal{FT}(G))+0.308(\mathcal{FT}(G)^2)$	0.6767	16.7478	11.6393
$108.734-29.203(\mathcal{FT}(G))+3.5733(\mathcal{FT}(G)^2)-0.09470(\mathcal{FT}(G)^3)$	0.8669	32.572	7.7128

The analysis as mentioned in Table 13 indicates that the best goodness of fit among obtained curvilinear equations using $\mathcal{FT}(G)$ topological index for ΔH_{vap}^0 is:
 $\Delta H_{vap}^0 = 108.734 - 29.203(\mathcal{FT}(G)) + 3.5733(\mathcal{FT}(G)^2) - 0.09470(\mathcal{FT}(G)^3)$.

Table 14. The curvilinear regressions models of $SDT(G)$ index for ΔH_{vap}^0 .

ΔH_{vap}^0	R^2	F	RMSE
$21.241+1.896(SDT(G))$	0.8255	80.4709	8.2942
$28.6641+1.3792(SDT(G))+0.00826(SDT(G)^2)$	0.8266	38.1525	8.5231
$36.2095+0.5884(SDT(G))+0.03431(SDT(G)^2)-0.08027(SDT(G)^3)$	0.8267	23.8565	80800973

The analysis as mentioned in Table 14 indicates that the best goodness of fit among obtained curvilinear equations using $SDT(G)$ topological index for ΔH_{vap}^0 is:
 $\Delta H_{vap}^0 = 36.2095 + 0.5884(SDT(G)) + 0.03431(SDT(G)^2) - 0.08027(SDT(G)^3)$.

From the analysis as mentioned in Table 11 - Table 14, the topological indices $ST(G)$ and $PT(G)$ are the best suitable for predicting the ΔH_{vap}^0 of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$) since $R^2 > 0.99$.

7 Plots of the cubic regression equation

The analysis as mentioned in Table 3 - Table 14 indicates that the cubic equation gives the best goodness of fit among three regression equations.

Figure 1 - Figure 3 shows the correlation of $ST(G)$ and $PT(G)$ with ΔH_f^0 , ΔH_C^0 , and ΔH_{vap}^0 using cubic regression equation.

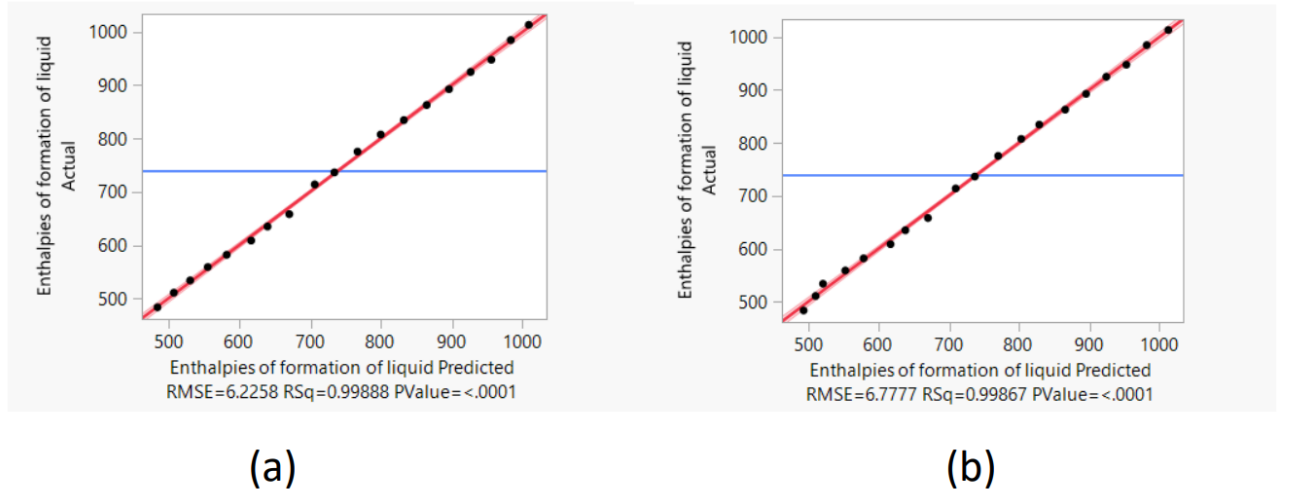


Figure 1: (a) Cubic regression equation of ΔH_f^0 with $ST(G)$. (b) Cubic regression equation of ΔH_f^0 with $PT(G)$.

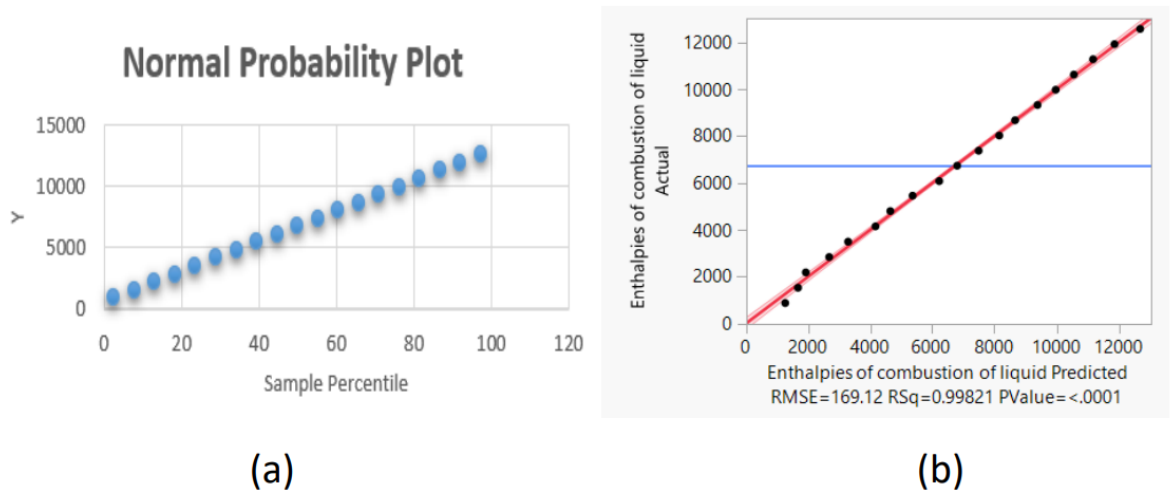


Figure 2: (a) Cubic regression equation of ΔH_C^0 with $ST(G)$ using normal probability plot. (b) Cubic regression equation of ΔH_C^0 with $PT(G)$.

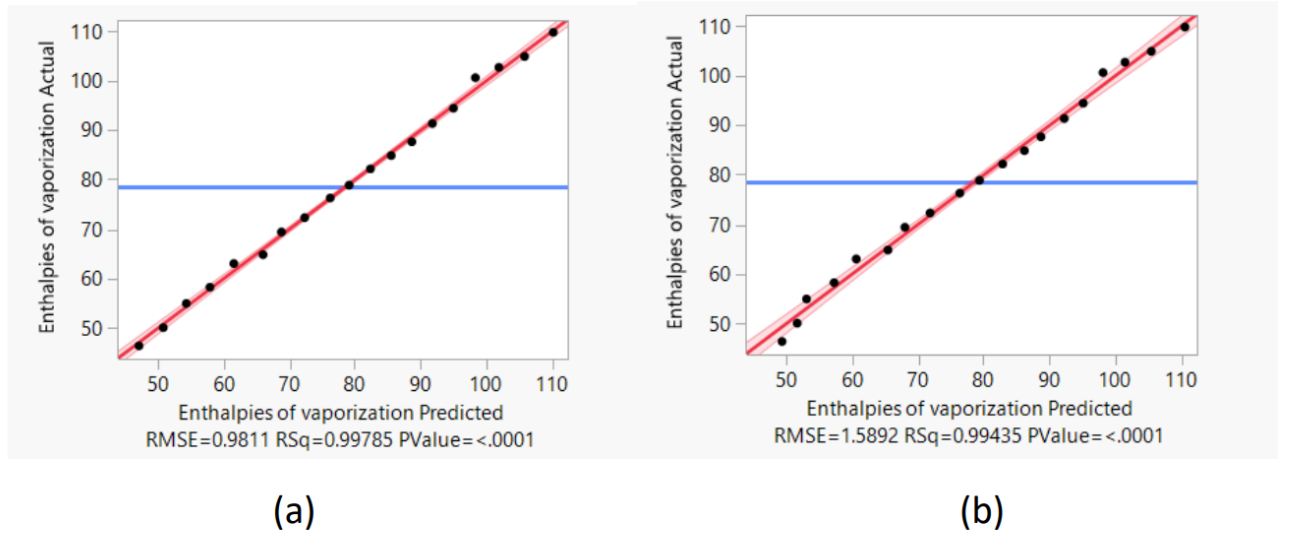


Figure 3: (a) Cubic regression equation of ΔH_{vap}^0 with $ST(G)$. (b) Cubic regression equation of ΔH_{vap}^0 with $PT(G)$.

8 Conclusion

The temperature-based topological indices such as the sum connectivity temperature index, product connectivity temperature index, F-temperature index, and symmetric division temperature index of 19 monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$) are calculated. Using these topological indices, curvilinear regression models are designed to predict certain thermodynamic properties such as enthalpies of formation (ΔH_f^0 liquid), enthalpies of combustion (ΔH_c^0 liquid), and enthalpies of vaporization (ΔH_{vap}^0 gas) of monocarboxylic acids. The most accurate results for the prediction of these thermodynamic properties can be calculated by using the sum connectivity temperature index and product connectivity temperature index. Furthermore, these thermodynamic properties also have a good correlation with the symmetric division temperature index, but the F-temperature index is not enough to make a good prediction of thermodynamic properties of monocarboxylic acids ($C_2H_4O_2$ - $C_{20}H_{40}O_2$). The optimum is the cubic equation form among curvilinear equations.

References

- [1] A. T. Balaban, *Topological index based on topological distances in a molecular graph*, Pure Appl. Chem., 1983, 55, 199-206.
- [2] A. T. Balaban and T. S. Balaban, *New vertex invariant and topological indices of chemical graphs based on information on distance*, J. Math. Chem., 1991, 8, 383-397.
- [3] J. Devillers and A. T. Balaban, *Topological indices and related descriptors in QSAR and QSPR*, Gordon and Breach Science Publishers, 1999, Amsterdam, Netherlands.
- [4] S. Fajtolowicz, *On Conjectures of Graffitti*, Discrete Mathematics, 1988, 72, 113-118.
- [5] O. C. Havare, *QSPR Analysis with Curvilinear Regression Modeling and Topological Indices*, Iranian Journal of Mathematical Chemistry, 2019, 10(4), 331-341.
- [6] H. Hosoya, *Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons*, Bull. Chem. Soc. Jpn., 1971, 44, 2332-2339.
- [7] L. B. Kier, L. H. Hall and J. W. Frazer, *Design of molecules from quantitative structure-activity relationship models. 1. Information transfer between path and vertex degree counts*, J. Chem. Inf. Comput. Sci., 1993, 33, 143-147.
- [8] V. R. Kulli, *Computation of Some Temperature Indices of $HC_5C_5[p, q]$ Nanotubes*, Annals of Pure and Applied Mathematics, 2019, 20(2), 69-74.
- [9] B. Lucić and N. Trinajstić, *New Developments in QSPR/QSAR Modeling Based on Topological Indices*, SAR and QSAR in Environmental Research, 1997, 7(1), 45-62.

- [10] M. Randić, *Characterization of Molecular Branching*, Journal of the American Chemical Society, 1975, 23, 6609-6615.
- [11] M. Randić and S. C. Basak, *Multiple Regression Analysis with Optimal Molecular Descriptors*, SAR and QSAR in Environmental Research, 2000, 11(1), 1-23.
- [12] M. Randić, *Quantitative Structure of property relationship: boiling points of planar benzenoids*, New. J. Che., 1996, 20, 1001-1009.
- [13] M. Randić, *On characterization of molecular attributes*, Acta. Chim. Slov., 1998, 45(3), 239-252.
- [14] F. Shafiei, *Relationship between Topological Indices and Thermodynamic Properties of the Monocarboxylic Acids Applications in QSPR*, Iranian Journal of Mathematical Chemistry, 2015, 6(1), 15-28.
- [15] M. Thakur, A. Thakur and P. V. Khadikar, *QSAR study on benzene sulphonamide carbonic anhydrase inhibitors: Topological approach using Balaban index*, Bioorg. Med. Chem., 2004, 12(4), 789-793.
- [16] R. Todeschini, Useful and unuseful summaries of regression Models, 2010. [http://www.molecularDescriptors.eu/tutorials/5-molecular descriptors-models.pdf](http://www.molecularDescriptors.eu/tutorials/5-molecular%20descriptors-models.pdf). (Access date: 23.06.2018).
- [17] H. Wiener, *Structural Determination of paraffin boiling points*, J. Am. Chem. Soc., 1947, 69(1), 17-20.