Iranian J. Math. Chem. 10 (4) December (2019) 331 − 341

Iranian Journal of Mathematical Chemistry

Journal homepage: ijmc.kashanu.ac.ir

QSPR Analysis with Curvilinear Regression Modeling and Topological Indices

OZGE COLAKOGLU HAVARE•

Department of Mathematics, Faculty of Science and Arts, Mersin University, 33343, Mersin, Turkey

ARTICLE INFO

Article History:

Received: 28 June 2019
Accepted: 30 November 2019
Published online: 30 December 2019
Academic Editor: Saeed Masoum

Keywords:

Gutman index
Product connectivity
Banhatti index
Variance of degree index
Sigma index
QSPR

ABSTRACT

Topological indices are the real number of a molecular structure obtained using molecular graph G. Topological indices are used for QSPR, QSAR and structural design in chemistry, nanotechnology, and pharmacology. Moreover, physicochemical properties such as the boiling point, the enthalpy of vaporization, and stability can be estimated by QSAR/QSPR models. In this study, the QSPR (Quantitative Structure-Property Relationship) models were designed using the Gutman index, the product connectivity Banhatti index, the Variance of degree index, and the Sigma index to predict the thermodynamic properties of monocarboxylic acids. The relationship analyses between the thermodynamic properties and the topological indices were done by using the curvilinear regression method. It was used the linear, quadratic and cubic equations of the curvilinear regression model. These regression models were then compared.

1. INTRODUCTION

The graph theory, known as a branch of discrete mathematics, was first presented by Euler to find the solution of Königsberg Bridges problem in 1736, and has been studied in Physics, Biology, Computer Sciences, Chemistry and so forth.

Chemical graph theory is concerned with finding topological indices which are well correlated with the properties of chemical molecules. Topological indices are widely used in analyzing the properties of molecules such as the structure-property relationship (QSPR), the structure-activity relationship (QSAR) modelling, and structural design in chemistry,

•Corresponding Author (Email address: ozgecolakoglu@mersin.edu.tr)
DOI: 10.22052/ijmc.2019.191865.1448

nanotechnology, and pharmacology. 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]. In the last seventy years, many topological indices have been defined and used in many applications. Let G be a simple connected graph with a vertex set $V(G)$ and edge set$E(G)$. The number of vertex sets and edge sets are defined by n and$m$, respectively. An edge of G is e, which connects the vertices u and$v$. The degree of a vertex u is defined by $d(u)$. The degree of an edge is defined by, $d(e) = d(u) + d(v) - 2$. A simple graph is an unweighted, undirected graph without any loops or multiple edges. A connected graph is a graph which is a path between all the pairs of vertices. We refer to this for an undefined term and a notation [2].

A molecular graph is a simple graph. Its vertices and edges represent the atoms and the bonds, respectively. Note that hydrogen atoms are omitted. A graph invariant is a real number related to a graph G. The topological index, mainly known as a graph-based molecular descriptor or graph invariant, is a numerical descriptor of the topological structure of a molecule [5].

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

In 2018, Gutman et al. introduced the sigma topological index. This recently introduced index is an extension of the Albertson irregularity index [6]. The sigma index is defined as

$$\sigma(G) = \sum_{uv \in E(G)} (d(u) - d(v))^2, \tag{1}$$

where $d(u)$ is denoted as the degree of vertex u [6].

Collatz and Sinogowitz introduced $E(G) = A(G) - \lambda(G)$ topological index based on eigenvalue as a measure of the "irregularity" of G. By Bell in 1992, the variance of the vertex degrees as an alternative index to this index was defined [1] as follows

$$Var(G) = \frac{1}{n} \sum_{u \in V(G)} d(u)^2 - \left(\frac{2m}{n}\right)^2. \tag{2}$$

In 2019, Kulli et. al. introduced a new topological index. The product connectivity Banhatti index of G which is a Randić type index which is defined as

$$PB(G) = \sum_{ue} \frac{1}{\sqrt{d(u)d(e)}}, \tag{3}$$

where ue means that the vertex u and edge e are incident in G [12].

In 1994, the Gutman index which is an extension of the Wiener index was introduced by Gutman. This index also is known as the modified Schultz index and is defined as

$$Gut(G) = \sum_{\{u,v\} \subseteq V(G)} d(u)d(v)d(u,v), \tag{4}$$

where $d(u)$ is denoted as the degree of vertex u and $d(u,v)$ is denoted as the distance between vertices of u and v [7].

Various regression methods using topological indices were implemented for the QSPR and QSAR models, see [8−11] and [13−15] for details. In this study, some topological indices of monocarboxylic acids were calculated by using the graph theory. Later, the most appropriate topological indices and curvilinear regressions analyses were obtained to predict the thermodynamic properties of monocarboxylic acids.

1.1. MATERIALS AND METHODS

The benchmark sets of the thermodynamic properties of the monocarboxylic acids were taken from [14]. Table 1 shows the thermodynamic properties of the monocarboxylic acids.

Table 1. The values of Enthalpies of formation of liquid (ΔH^{o}_{f}kJ/mol), Enthalpies of combustion of liquid (ΔH^{o}_{C} kJ/mol) and Enthalpies of vaporization (ΔH^{o}_{vap} kJ/mol) of monocarboxylic acids ($C_2H_4O_2$-$C_{20}H_{40}O_2$) at conditions (normally 298.15 K, 1 atm).

Name of compounds	Formula	Enthalpies of formation of liquid	Enthalpies of combustion of liquid	Enthalpies of vaporization
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	76,3
Undecanoic acid	$C_{11}H_{22}O_2$	736,2	6736,5	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

It is employed some models that use regression to fit curves instead of straight lines. This method is known as curvilinear regression analysis. In this study, we tested the following equations;

$$Y = a + b_1 X_1 \; ; \qquad n, R^2, s, F \quad \text{(Linear Equation)}$$
$$Y = a + b_1 X_1 + b_2 {X_1}^2 ; \qquad n, R^2, s, F \quad \text{(Quadratic Equation)}$$
$$Y = a + b_1 X_1 + b_2 {X_1}^2 + b_3 {X_1}^3 \; ; \qquad n, R^2, s, F \quad \text{(Cubic Equation)}$$

where Y is the response or dependent variable, a is the regression model constant, b_i $(i = 1, 2, 3)$ are the coefficients for the individual descriptor, X_i $(i = 1,2,3)$ are 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 deviation, and F is the calculated value of the F-ration test. For more detailed information see [4]. Note that the correlation coefficient is nearest to 1 and 1 when the experimental and theoretical results are closest. The observed values and model predictions must be compared to measure the predictive quality of the model, see [3] and [16] for details. So, we will deal with the RMSE (Root Mean Square Error) metric for the predictive power of the model. It is clear that the best predictive model is the minimum error, i.e. the minimum RMSE. R^2 and F parameters will be considered for the goodness of fit of the model.

Todeschini [16] selected the best goodness of fit in models by using any of the parameters $\max(R^2)$, $\max(F)$. The curvilinear regression analyses were obtained by using the SPSS statistical software. The independent variables in the curvilinear regression models are the Gutman index, the product connectivity Banhatti index, the variance of degree index, and the sigma index of 19 monocarboxylic acids.

2. MAIN RESULTS

By using (1.1−1.4) formulas, the topological indices of 19 monocarboxylic acids were computed to design the QSPR. Table 2 shows the value of the topological indices of 19 monocarboxylic acids.

The following structure-property models were obtained the linear, quadratic and cubic regression models for the Variance of the degree topological index, the Gutman index, the product connectivity Banhatti index and the sigma index. Tables 3–6 show linear, quadratic and cubic regression models of $VAR(G)$, $Gut(G)$, $BP(G)$ and $\sigma(G)$ indices for ΔH^{o}_{f}, respectively. Tables 7–10 show linear, quadratic and cubic regression models of $VAR(G)$, $Gut(G)$, $BP(G)$ and $\sigma(G)$ indices for ΔH^{o}_{C}, respectively. Table 11–14 show linear, quadratic and cubic regression models $VAR(G)$, $Gut(G)$, $BP(G)$ and $\sigma(G)$ indices for $\Delta H^{o}_{\text{vap}}$, respectively.

Table 2. The value of topological indices of 19 monocarboxylic acids.

Formula	$Var(G)$	$Gut(G)$	$PB(G)$	$\sigma(G)$
$C_2H_4O_2$	$\frac{3}{4}$	15	$\frac{3\sqrt{2}+\sqrt{6}}{2}$	12
$C_3H_6O_2$	$\frac{16}{25}$	36	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{4}{3}$	10
$C_4H_8O_2$	$\frac{5}{9}$	73	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{7}{3}$	10
$C_5H_{10}O_2$	$\frac{24}{49}$	130	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{10}{3}$	10
$C_6H_{12}O_2$	$\frac{28}{64}$	223	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{13}{3}$	10
$C_7H_{14}O_2$	$\frac{32}{81}$	332	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{16}{3}$	10
$C_8H_{16}O_2$	$\frac{36}{100}$	473	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{19}{3}$	10
$C_9H_{18}O_2$	$\frac{40}{121}$	650	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{22}{3}$	10
$C_{10}H_{20}O_2$	$\frac{44}{144}$	855	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{25}{3}$	10
$C_{11}H_{22}O_2$	$\frac{48}{169}$	1116	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{28}{3}$	10
$C_{12}H_{24}O_2$	$\frac{52}{196}$	1425	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{31}{3}$	10
$C_{13}H_{26}O_2$	$\frac{56}{225}$	1786	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{34}{3}$	10
$C_{14}H_{28}O_2$	$\frac{60}{256}$	2203	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{37}{3}$	10
$C_{15}H_{30}O_2$	$\frac{64}{289}$	2654	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{40}{3}$	10
$C_{16}H_{32}O_2$	$\frac{68}{324}$	3195	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{43}{3}$	10
$C_{17}H_{34}O_2$	$\frac{72}{361}$	3804	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{46}{3}$	10
$C_{18}H_{36}O_2$	$\frac{76}{400}$	4485	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{49}{3}$	10
$C_{19}H_{38}O_2$	$\frac{80}{441}$	5242	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{52}{3}$	10
$C_{20}H_{40}O_2$	$\frac{84}{484}$	6079	$\frac{3\sqrt{2}+\sqrt{6}}{2}+\frac{55}{3}$	10

Table 3. The curvilinear regressions models of $VAR(G)$ index for $\Delta H^o_{f.}$

ΔH^o_f is equality for the following equations	R^2	F	RMSE
$1057.692 - 932.993VAR(G)$	0.834	85.147	67.36
$1416.353 - 2980.633VAR(G) + 2383.560VAR(G)^2$	0.979	381.625	23.65
$1774{,}574 - 6022.273VAR(G) + 9977.986(VAR(G))^2 - 5668.638\,VAR(G)^3$	0.998	3206.667	6.51

From the above analysis, it appears that the best goodness of fitamong obtained curvilinear equations using$VAR(G)$ topological index for ΔH^o_f is $\Delta H^o_f = 1774.574 - 6022.273VAR(G) + 9977.986(VAR(G))^2 - 5668.638\,VAR(G)^3$.

Table 4. The curvilinear regressions models of $Gut(G)$ index for $\Delta H^o_{f.}$

ΔH^o_f is equality for the following equations	R^2	F	RMSE
$586.343 + 0.084Gut(G)$	0.891	138.695	54.55
$538.854 + 0.169Gut(G) - 1.583E-5Gut(G)^2$	0.975	312.742	26.07
$516.208 + 0.251Gut(G) - 5.573E-5\,(Gut(G))^2 + 4.638E-9Gut(G)^3$	0.993	718.690	13.72

The best ideal model from the above analysis for ΔH^o_f is $\Delta H^o_f = 516.208 + 0.251\,Gut(G) - 5.573E-5\,(Gut(G))^2 + 4.638E-9Gut(G)^3$.

Table 5. The curvilinear regressions models of $PB(G)$ index for $\Delta H^o_{f.}$

ΔH^o_f is equality for the following equations	R2	F	RMSE
$360.956 + 29.934PB(G)$	0.997	4940.177	9.66
$384.948 + 25.235PB(G) + 0.186PB(G)^2$	0.978	3221.182	8.21
$444.458 + 5.956PB(G) + 1.923\,(PB(G))^2 - 0.046PB(G)^3$	0.999	4629.392	5.42

The best ideal model from the above analysis for $\Delta H^o_{f.}$ is $\Delta H^o_f = 444.458 + 5.956PB(G) + 1.923\,(PB(G))^2 - 0{,}046PB(G)^3$.

Table 6. The curvilinear regressions models of $\sigma(G)$ index for ΔH^o_f

ΔH^o_f is equality for the following equations	R^2	F	RMSE
$2107.800 - 135.358\sigma(G)$	0.134	2.631	153.65
$0.000\sigma(G) + 0.000\sigma(G)^2$	-	-	-
$0.000 + 0.000\sigma(G) + 0.000\,(\sigma(G))^2 + 0.000\sigma(G)^3$	-	-	-

The best ideal model among obtained curvilinear equations using $\sigma(G)$ topological index for $\Delta H^{o}_{f.}$ is ΔH^{o}_{f} = 2107.800 – 135.358$\sigma(G)$. Cubic equations for ΔH^{o}_{f} are the best among the curvilinear regression equations included $VAR(G)$, $Gut(G)$ and $PB(G)$ topological indices. Also, $PB(G)$ for ΔH^{o}_{f} is the best predictive topological index. The predictive power of the curvilinear equations containing the sigma index is very weak.

Table 7. The curvilinear regressions models of $VAR(G)$ index for ΔH^{o}_{C}.

ΔH^{o}_{C} is equality for the following equations	**R^2**	**F**	**RMSE**
13682.378 –20407.636VAR(G)	0.860	104.307	1331.18
20714.656 – 60555.744VAR(G)+46734.499 $VAR(G)^2$	0.981	408.510	492.82
27949.624 – 121987.608$VAR(G)$ + 200118.744$(VAR(G))^2$ – 114489.198$VAR(G)^3$	0.997	1980.500	178.44

The best ideal model from the above analysis for ΔH^{o}_{C} is ΔH^{o}_{C} =27949.624 – 121987.608$VAR(G)$ + 200118.744 $(VAR(G))^2$– 114489.198$VAR(G)^3$.

Table 8. The curvilinear regressions models of $Gut(G)$ index for ΔH^{o}_{C}.

ΔH^{o}_{C} is equality for the following equations	**R^2**	**F**	**RMSE**
3444.074+1.797Gut(G)	0.880	124.829	1231.11
2421.006+ 3.622Gut(G) + 0.000$Gut(G)^2$	0.964	216,978	1193.25
1895.372+ 5.532Gut(G)–0.001 $(Gut(G))^2$ +1.076E–7$Gut(G)^3$	0.985	337.177	429.84

The best ideal model from the above analysis is ΔH^{o}_{C} =1895.372 + 5.532$Gut(G)$ – 0.001 $(Gut(G))^2$ + 1.076E – 7$Gut(G)^3$.

Table 9. The curvilinear regressions models of $PB(G)$ index for ΔH^{o}_{C}.

ΔH^{o}_{C} is equality for the following equations	**R^2**	**F**	**RMSE**
–1443.264+645.737PB(G)	1.000	126072.117	41.28
–1353.309+628.118PB(G)+ 0.698$PB(G)^2$	1.000	75483.883	36.60
–1176.477+ 570.831PB(G)+5.860 $(PB(G))^2$– 0.137$PB(G)^3$	1.000	63008.867	31.67

The best goodness of fitand predictive modelamong obtained curvilinear equations using $PB(G)$ topological index for ΔH^{o}_{C} is ΔH^{o}_{C} =–1176.477+ 570.831$PB(G)$+5.860 $(PB(G))^2$– 0.137$PB(G)^3$.

Table 10. The curvilinear regressions models of $\sigma(G)$ index for ΔH^{o}_{C}.

ΔH^{o}_{C} **is equality for the following equations**	R^2	F	RMSE
37973.033 – 3091.489σ(G)	0.151	3.018	3276.99
0.000+ 0.000σ(G) +0.000$\sigma(G)^2$	-	-	-
0.000+ 0.000 σ(G) +0.000 $(\sigma(G))^2$ + 0.000$\sigma(G)^3$	-	-	-

The best ideal model among obtained curvilinear equations using $\sigma(G)$ topological index for ΔH^{o}_{C} is ΔH^{o}_{C} =37973.033– 3091.489$\sigma(G)$. The best predictive capability among topological indices is $PB(G)$, for ΔH^{o}_{C}. The cubic equations calculate the best solutions among the curvilinear equations containing topological indices for ΔH^{o}_{C}. The sigma index does not enough to make a good prediction for ΔH^{o}_{C}.

Table 11. The curvilinear regressions models of $VAR(G)$ index for ΔH^{o}_{vap}.

ΔH^{o}_{vap} **is equality for the following equations**	R^2	F	RMSE
115.588 – 108.647VAR(G)	0.873	117.085	6.68
150.015 –305.195$VAR(G)$ + 228.792$VAR(G)^2$	0.977	340.762	2.84
188.568 – 632.546$VAR(G)$ +1046.127 $(VAR(G))^2$ – 610.076$VAR(G)^3$	0.994	834.012	1.45

The best ideal model from the above analysis is ΔH^{o}_{vap} =188,568– 632.546$VAR(G)$ + 1046.127 $(VAR(G))^2$ – 610.076$VAR(G)^3$.

Table 12. The curvilinear regressions models of $Gut(G)$ index for ΔH^{o}_{vap}.

ΔH^{o}_{vap} **is equality for the following equations**	R^2	F	RMSE
61.298 + 0.009Gut(G)	0.872	115.648	6.72
56.009+ 0.019Gut(G)–1.763E – 6$Gut(G)^2$	0.953	160.678	4.09
53.089 + 0.029Gut(G) – 6.907E-6 $(Gut(G))^2$ +5.979E – 10$Gut(G)^3$	0.976	200.966	2.92

The best ideal equation from the above analysis is ΔH^{o}_{vap} = 53.089+0.029$Gut(G)$ – 6.907E - 6 $(Gut(G))^2$ + 5.979E – 10$Gut(G)^3$.

Table 13. The curvilinear regressions models of $PB(G)$ index for ΔH^{o}_{vap}.

ΔH^{o}_{vap} **is equality for the following equations**	R^2	F	RMSE
$35.439 + 3.408\mathrm{PB(G)}$	0.998	7960.232	0.86
$34.681 + 3.556\mathrm{PB(G)} - 0.006\mathrm{PB(G)}^2$	0.998	3879.867	0.85
$30.694 + 4.848\ \mathrm{PB(G)} - 0.122\ (\mathrm{PB(G)})^2 + 0.003\mathrm{PB(G)}^3$	0.998	3174.132	0.74

From $\max(R^2)$ and $\max(F)$ and min(RMSE), the best ideal model is ΔH^{o}_{vap} $=30.694 + 4.848\ PB(G) - 0.122\ (PB(G))^2 + 0.003PB(G)^3$.

Table 14. The curvilinear regressions models of $Gut(G)$ index for ΔH^{o}_{C}.

ΔH^{o}_{vap} **is equality for the following equations**	R^2	F	RMSE
$250.800 - 17.042\sigma(\mathrm{G})$	0.164	3.338	17.17
$0.000 + 0.000\sigma(\mathrm{G}) + 0.000\sigma(\mathrm{G})^2$	.	.	.
$0.000\sigma(\mathrm{G}) + 0.000\ (\sigma(\mathrm{G}))^2 + 0.000\ \sigma(\mathrm{G})^3$	.	.	.

The best ideal model among curvilinear equations containing $\sigma(G)$ topological index for ΔH^{o}_{vap} is $\Delta H^{o}_{\mathrm{vap}} = 250.800 - 17.042\sigma(G)$, for ΔH^{o}_{vap}; The best predictive among the topological indices is $PB(G)$. The optimum is the cubic equation form among curvilinear equations. The sigma index does not enough to make a prediction.

3. CONCLUSIONS

The product connectivity Banhatti index, the Variance of degree index, the Gutman index, and the sigma index were introduced and studied. Through these indices, curvilinear regression models were created to predict some thermodynamic properties of monoacids.

The most accurate results for the prediction of the enthalpies of the formation of the liquid (ΔH^{o}_{f} kJ/mol) of monocarboxylic acids can be calculated by using product connectivity Banhatti index in cubic equation form.

The most accurate results for the prediction of the enthalpies of the combustion of liquid (ΔH^{o}_{C} kJ/mol) of monocarboxylic acids can be calculated by using the product connectivity Banhatti index in cubic equation form. The most accurate results for the enthalpies of the vaporization ($\Delta \mathrm{H}^{o}_{vap}$ kJ/mol) of monocarboxylic acids can be calculated by using the product connectivity Banhatti index in the cubic equation form.

The Sigma index has not been obtained QSPR models for ΔH^{o}_{f}, $\Delta \mathrm{H}^{o}_{C,}$ and $\Delta \mathrm{H}^{o}_{vap}$ of monocarboxylic acids. So, the sigma index makes not a prediction for ΔH^{o}_{f}, $\Delta \mathrm{H}^{o}_{C,}$ and ΔH^{o}_{vap} of monocarboxylic acids.

REFERENCES

1. F. K. Bell, A note on the irregularity of graphs, *Linear Algebra Appl*. **161** (1992) 45–54.
2. G. Chartrand and L. Lesniak, *Graphs and Digraphs*, CRS Press, 2005.
3. V. Consonni, D. Ballabio and R. Todeschini, Comments on the definition of the Q^2 parameter for QSAR validation, *J. Chem. Inf. Model.* **49** (7) (2009) 1669–1678.
4. J. C. Dearden, *Advances in QSAR Modeling*, Springer International Publishing, Switzerland, 2017.
5. I. Gutman, A property of the simple topological index, *MATCH Commun. Math. Comput. Chem.* **25** (1990) 131–140.
6. I. Gutman, M. Togan, A. Yurttas, A. S. Cevik and I. N. Cangul, Inverse problem for sigma index**,** *MATCH Commun. Math. Comput. Chem.* **79** (2018) 491–508.
7. I. Gutman, Selected properties of Schultz molecular topological index, *J. Chem. Inf. Coumput. Sci.* **34** (1994) 1087–1089.
8. S. Hosamani, D. Perigidad, S. Jamagoud, Y. Maled and S. Gavade, QSPR analysis of certain degree based topological indices, *J. Stat. Appl. Prob.* **6** (2017) 361–371.
9. H. Hosseini and F. Shafiei, Quantitative structure property relationship models for the prediction of gas heat capacity of benzene derivatives using topological indices, *MATCH Commun. Math. Comput. Chem.* **75** (2016) 583–592.
10. H. Hosseini and F. Shafiei, Entropy prediction of benzene derivatives using topological indices, *Studia UBB Chemia* LXII, **2** (2017) 297–310.
11. P. V. Khadikar and S. Karmarkar, A novel PI index and its applications to QSPR/QSAR studies, *J. Chem. Info. Comp. Sci.***41** (2001) 934–949.
12. V. R. Kulli, B. Chaluvaraju and H. S. Boregowda, The product connectivity Banhatti index of a graph, *Dis. Math. Graph Theory* **39** (2019) 505–517.
13. E. Mohammadinasab, Determination of critical properties of Alkanes derivatives using multiple linear regressions, *Iranian J. Math. Chem.* **8** (2017) 199–220.
14. F. Shafiei, Relationship between Topological indices and Thermodynamic properties and of the Monocarboxylic acids Applications in QSPR, *Iranian J. Math. Chem.* **6** (2015) 15–28.
15. N. Sivaraman, T. G. Srinivasan and P. R. Vasudeva Rao, QSPR modeling for solubility of fullerene (C_{60}) in organic solvents, *J. Chem. Inf. Comput. Sci.* **41** (2001) 1067–1074.

16. R. Todeschini, *Useful and unuseful summaries of regression Models*, 2010. http://www.moleculardescriptors.eu/tutorials/T5_moleculardescriptors_models.pdf. (Access date: 23.06.2018).
17. H. Wiener, Structural determination of paraffin boiling points, *J. Am. Chem. Soc*. **69** (1947) 17–20.