Prediction of Physicochemical Properties of Alkanes Using Multilinear QSPR Models Based on the Eccentricity Zagreb Index and VL-Index

Abstract:

Topological indices are mathematical descriptors that enable quantitative correlations between molecular structure and physicochemical properties or chemical reactivity. In this study, the relationship between the molecular structures of alkanes and the eccentricity-based first Zagreb index as well as the eccentricity-based VL-index is investigated. The results demonstrate that these indices serve as effective descriptors for predicting the physical properties of lower alkanes.

Keywords:Topological indices; Zagreb index; VL -index; Eccentricity; Alkanes; Physical properties; Structure–property relationship

1 Introduction

In theoretical and mathematical chemistry, topological indices *are numerical invariants derived from the* molecular graph $G = (V(G), E(G))$ *[17], where the vertex set* $V(G)$ *represents carbon atoms and the edge set* $E(G)$ *corresponds to covalent C–C bonds. These*

*

graph-theoretic descriptors provide a rigorous mathematical framework for encoding molecular connectivity and topological complexity into quantitative parameters that correlate with experimentally measurable physicochemical properties, chemical reactivity, and biological activity.

Formally, the molecular graph G of an alkane is a simple, connected, undirected graph in which each vertex $v \in V(G)$ has degree $d(v)$, representing the number of carbon atoms directly bonded to v. The distance $d(u, v)$ between two vertices $u, v \in V(G)$ is defined as the length of the shortest path connecting them. The eccentricity $\varepsilon(v)$ *of a vertex v is given by*

$$\varepsilon(v) = \max_{u \in V(G)} d(u, v),$$

which measures the topological remoteness of v from the farthest vertex in the graph. By aggregating such local and global structural characteristics, topological indices serve as molecular fingerprints in quantitative structure–property relationship (QSPR) *and* quantitative structure–activity relationship (QSAR) *studies [12].*

Among the various classes of topological indices, degree-based indices *have attracted considerable attention due to their computational simplicity and strong predictive capability. The classical* Zagreb indices*, introduced by Gutman and Trinajstić in 1972 [8], are defined by*

$$M_1(G) = \sum_{v \in V(G)} d(v)^2,$$

where the first Zagreb index $M_1(G)$ captures the second moment of the vertex degree distribution. This index has proven particularly effective in modeling physicochemical properties of organic molecules owing to its sensitivity to molecular size and branching patterns.

Recent developments have extended traditional degree-based indices by incorporating additional graph-theoretic concepts such as eccentricity [5]. One such extension is the eccentricity-based first Zagreb index*, defined as*

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2 \, \varepsilon(v),$$

which simultaneously encodes local connectivity *through the squared vertex degree and* global structural information *through vertex eccentricity. This combined representation enhances the discrimination of structural isomers and improves correlations with size- and shape-dependent molecular properties.*

The VL-index, introduced in 2020 and inspired by the Zagreb indices, is attributed to the work of Veerabhadraiah Lokesha and was later formulated by Deepika T. [6]. The VL-index of a graph G is defined as

$$VL(G) = \frac{1}{2} \sum_{uv \in E(G)} (d_e + d_f + 4),$$

where $d_e = d_u + d_v - 2$ and $d_f = d_u d_v - 2$, with d_u and d_v denoting the degrees of the vertices u and v, respectively. An equivalent and simplified expression of the VL-index is given by

$$VL(G) = \frac{1}{2} \sum_{uv \in E(G)} (d_u + d_v + d_u \cdot d_v).$$

Lower alkanes *(C_nH_{2n+2})[2, 3] constitute an ideal model system for QSPR investigations due to their structural simplicity, systematic variation through branching, and the*

availability of reliable experimental data. The molecular graphs of alkanes are tree graphs *(connected acyclic graphs), in which topological features directly reflect molecular geometry and branching patterns that influence intermolecular interactions and phase behavior.*

In this study, the effectiveness of the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and the VL-index $\xi VL(G)$ as molecular descriptors is systematically investigated using a dataset of 72 lower alkanes, each represented by its corresponding molecular graph. Comprehensive QSPR analyses are performed to explore the correlation between $M_1^{\varepsilon}(G)$, VL-index $\xi VL(G)$ and seven important physicochemical properties, namely boiling points, molar volumes, molar refractions, heats of vaporization, critical temperatures, critical pressures, and surface tensions, with experimental data obtained from reliable literature sources.

By integrating chemical graph theory with experimental thermophysical data, this work contributes to the development of reliable structure–property relationships and provides a useful framework for molecular design, property prediction, and understanding alkane behavior in chemical and chemical engineering applications[1].

2 Materials and Methods

Although a wide variety of topological indices have been investigated in the literature, this section focuses exclusively on those indices pertinent to the present study.

Let $G = (V(G), E(G))$ be a simple, connected graph representing a molecular structure, where $V(G)$ denotes the vertex set and $E(G)$ denotes the edge set. For a vertex $v \in V(G)$, the degree of v is denoted by $d(v)$.

Gutman and Trinajstić introduced the classical Zagreb indices in [8, 9], which are defined as

$$M_1(G) = \sum_{uv \in E(G)} (d_u + d_v),$$

$$M_2(G) = \sum_{uv \in E(G)} d_u \cdot d_v,$$

where d_u and d_v denote the degrees of the end vertices u and v of the edge uv, respectively.

The redefined Zagreb index, denoted by $ReZG(G)$, is defined as

$$ReZG(G) = \sum_{uv \in E(G)} \frac{d(u) + d(v)}{d(u)\, d(v)},$$

where u and v are the end vertices of the edge uv, and $d(u)$ and $d(v)$ represent their respective degrees.

The VL-index, introduced in 2020 and inspired by the Zagreb indices, is attributed to the work of Veerabhadraiah Lokesha and was later formulated by Deepika T. [6]. The VL-index of a graph G is defined as

$$VL(G) = \frac{1}{2} \sum_{uv \in E(G)} (d_u + d_v + d_u \cdot d_v),$$

where d_u and d_v denote the degrees of the vertices u and v in G, respectively.

Eccentricity of a Graph

For a vertex $v \in V(G)$, the eccentricity of v, denoted by $\varepsilon(v)$, is defined as the maximum distance from v to any other vertex in G, that is,

$$\varepsilon(v) = \max\{d(v,u) \mid u \in V(G)\},$$

where $d(v,u)$ denotes the length of the shortest path between the vertices v and u.

In chemical graph theory, vertex eccentricity plays a crucial role in capturing the global structural characteristics of a molecular graph. It reflects the maximum separation of a vertex from the remainder of the graph and thus provides insight into the overall topology and spatial extent of a molecule. By incorporating eccentricity with degree-based measures, hybrid indices effectively encode both local connectivity and global structural information.

Definition 2.1. *The eccentricity-based first Zagreb index of a graph G, denoted by $M_1^{\varepsilon}(G)$, is defined as*

$$M_1^{\varepsilon}(G) = \sum_{v \in V(G)} d(v)^2\, \varepsilon(v),$$

where $V(G)$ is the vertex set of G, $d(v)$ denotes the degree of the vertex v, and $\varepsilon(v)$ represents its eccentricity.

Definition 2.2. *The eccentricity-based VL-index of a graph G is defined as [13]*

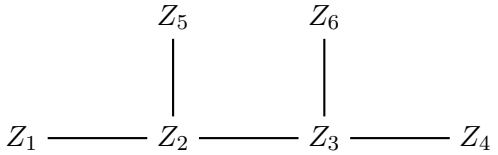
$$\xi(VL(G)) = \frac{1}{2} \sum_{uv \in E(G)} \left(\xi_u + \xi_v + \xi_u \cdot \xi_v\right),$$

where ξ_u and ξ_v denote the eccentricities of the vertices u and v in the graph G, respectively.

Example 2.1 (Using Definition 2.1).

Let G be the molecular graph of 2,3-dimethylbutane, which consists of six vertices. The corresponding molecular graph is shown below.

Graph 1:



The degree, eccentricity, and corresponding contributions to $M_1^{\varepsilon}(G)$ for each vertex are summarized in Table 1.

Using Definition 2.1, we compute

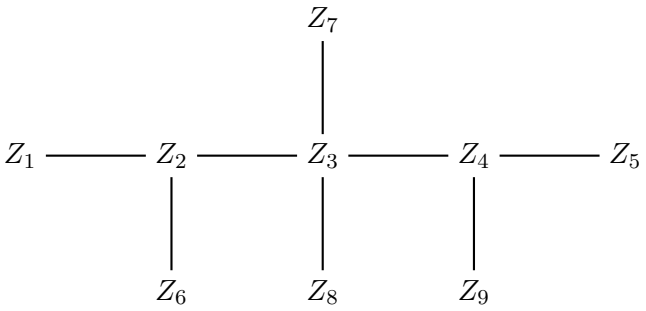
$$\begin{aligned} M_1^{\varepsilon}(G) &= 1^2(3) + 3^2(2) + 3^2(2) + 1^2(3) + 1^2(3) + 1^2(3) \\ &= 3 + 18 + 18 + 3 + 3 + 3 \\ &= 48. \end{aligned}$$

$$\boxed{M_1^{\varepsilon}(G) = 48}$$

Vertex	$d(v)$	$\varepsilon(v)$	$d(v)^2\varepsilon(v)$
Z_1	1	3	3
Z_2	3	2	18
Z_3	3	2	18
Z_4	1	3	3
Z_5	1	3	3
Z_6	1	3	3

Example 2.2 (Using Definition 2.2).
Let $G = (V(G), E(G))$ be the molecular graph of 2, 3, 3, 4-tetramethylpentane, consisting of $|V(G)| = 9$ vertices corresponding to its molecular formula C_9H_{20}. The associated molecular graph is depicted below.

Graph 2:



Here, $\xi_{Z_1} = 4$, $\xi_{Z_4} = 3$, $\xi_{Z_3} = 2$, $\xi_{Z_2} = 3$, $\xi_{Z_5} = 4$, $\xi_{Z_6} = 4$, $\xi_{Z_7} = 3$, $\xi_{Z_8} = 3$, $\xi_{Z_9} = 4$

Therefore, by Definition 2.2, the eccentricity-based VL-index of G is given by

$$\xi_{VL}(G) = \frac{1}{2}(19 + 19 + 11 + 11 + 11 + 11 + 19 + 19) = 60.$$

$$\boxed{\xi_{VL}(G) = 60}$$

3 Preliminary Results & Discussion

Let G be the molecular graph of an alkane, where the vertices correspond to carbon atoms and the edges represent carbon–carbon bonds. The *eccentricity-based first Zagreb index* of G is defined as in [5], while the *eccentricity-based VL-index* of a graph G is defined in [13].

Using these definitions, the values of $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ are computed for various alkanes and their structural isomers. The computed values, together with the corresponding molecular graph representations, are summarized in Table 1.

Table 1: Calculated $M_1^{\varepsilon}(G)$ and ξVL-indices values for alkanes

Sl. No.	Name of the Chemical Compound	$M_1^{\varepsilon}(G)$	$\xi VL(G)$
01	Ethane	2	1.5
02	Propane	8	5
03	Butane	22	15
04	2-Methylpropane(isobutane)	15	7.5
05	Pentane	40	30
06	2-Methylbutane	35	20.5
07	2,2-Dimethylpropane	24	10
08	Hexane	66	55.5
09	2-Methylpentane	39.5	89
10	3-Methylpentane	53	39.5
11	2,2-Dimethylbutane	47	35.5
12	2,3-Dimethylbutane	48	26
13	Heptane	96	26
14	2-Methylhexane	117	70
15	3-Methylhexane	123	65
16	3-Ethylhexane	97	79.5
17	2,2-Dimethylpentane	93	49
18	2,3-Dimethylpentane	82	45
19	2,4-Dimethylpentane	56	49
20	3,3-Dimethylpentane	70	41
21	2,3,3-Trimethylbutane	65	31.5
22	Octane	134	55.5
23	2-Methylheptane	187	109.5
24	3-Methylheptane	193	103.5
25	4-Methylheptane	171	98.5
26	2,2-Dimethylhexane	194	84.5
27	2,3-Dimethylhexane	213	89
28	2,4-Dimethylhexane	227	79.5
29	2,5-Dimethylhexane	154	84.5
30	3,3-Dimethylhexane	121	74.5
31	3,4-Dimethylhexane	134	74.5
32	3-Ethyl-2-methylpentane	145	54.5
33	3-Ethyl-3-methylpentane	77	50.5
34	2,2,3-Dimethylpentane	131	54.5
35	2,2,4-Dimethylpentane	161	58.5
36	2,3,3-Trimethylpentane	112	50.5
37	2,3,4-Trimethylpentane	113	54.5
38	2,2,3,3-Tetramethylbutane	82	37
39	Nonane	176	196
40	2-Methyloctane	255	164.5
41	3-Methyloctane	250	157.5
42	4-Methyloctane	240	151.5
43	3-Ethylheptane	187	124
44	2,2-Dimethylheptane	307	130

45	2,3-Dimethylheptane	199	124
46	2,4-Dimethylheptane	222	119
47	2,5-Dimethylheptane	204	124
48	2,6-Dimethylheptane	224	130
49	3,3-Dimethylheptane	310	118
50	3,4-Dimethylheptane	191	113
51	3,5-Dimethylheptane	199	118
52	4,4-Dimethylheptane	212	108
53	3-Ethyl-2-methylhexane	196	94
54	4-Ethyl-2-methylhexane	170	94
55	3-Ethyl-3-methylhexane	190	89
56	3-Ethyl-4-methylhexane	169	87
57	2,2,3-Trimethylhexane	203	94
58	2,2,4-Trimethylhexane	197	94
59	2,2,5-Trimethylhexane	201	99
60	2,3,3-Trimethylhexane	166	89
61	2,3,4-Trimethylhexane	150	89
62	2,3,5-Trimethylhexane	154	94
63	2,4,4-Trimethylhexane	188	89
64	3,3,4-Trimethylhexane	190	84
65	3,3-Diethylpentane	136	60
66	2,3-Diethyl-3-ethylpentane	464	60
67	2,2-Diethyl-3-ethylpentane	264	64
68	2,2,3,3-Tetramethylpentane	108	60
69	2,2,3,4-Tetramethylpentane	132	64
70	2,2,4,4-Tetramethylpentane	100	68
71	2,3,3,4-Tetramethylpentane	132	60
72	4-Ethylpentane	160	113

4 Correlation Analysis: Eccentricity-based First Zagreb Index, VL Index and Physicochemical Properties

Quantitative Structure–Property Relationship (QSPR) and Quantitative Structure–Activity Relationship (QSAR) methodologies provide a robust mathematical framework for correlating molecular structure with experimentally measurable physicochemical properties. In the present study, the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and the vl-index $\xi VL(G)$ index are employed as graph-theoretical descriptors to model physicochemical properties of alkanes such as boiling point, molar volume, entropy, and enthalpy.

The calculated values of $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ for seventy-two alkane isomers are listed in Table 1. These indices encode essential structural characteristics related to vertex eccentricity, molecular size, and branching, which are known to strongly influence thermodynamic and physical properties.

Let P denote a physicochemical property of interest. The linear association between

P and a given topological index (TI) is first examined using the Pearson correlation coefficient r, defined as

$$r = \frac{\sum_{i=1}^{n}(TI_i - \overline{TI})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(TI_i - \overline{TI})^2 \sum_{i=1}^{n}(P_i - \overline{P})^2}},$$

where n is the number of molecules, and $\overline{TI}$ and $\overline{P}$ represent the mean values of the corresponding variables.

4.1 Linear Regression Models

To evaluate the predictive ability of each descriptor independently, simple linear regression models are constructed. The relationship between a physicochemical property P and the eccentricity-based first Zagreb index is expressed as[7]

$$P = a_1 + b_1\, M_1^{\varepsilon}(G),$$

while the corresponding model based on the $\xi VL(G)$ index is given by

$$P = a_2 + b_2\, \xi VL(G),$$

where a_1, a_2 are intercepts and b_1, b_2 are regression coefficients determined using the least squares method.

4.2 Multilinear Regression Model

To enhance predictive accuracy and account for the combined influence of both structural descriptors, a multilinear regression (MLR) model is also developed[7]:

$$P = a + b_1\, M_1^{\varepsilon}(G) + b_2\, \xi VL(G),$$

where a is the intercept and b_1, b_2 represent the relative contributions of the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and the VL-index $\xi VL(G)$, respectively.

4.3 Statistical Summary of Multilinear QSPR Models Using $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ for the Prediction of Physicochemical Properties of Alkanes

Table 2 summarizes the statistical performance of Multilinear QSPR models developed using the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and VL-index $\xi VL(G)$ as a single molecular descriptor[10]. The adoption of correlation-based validation measures is well established in QSPR and QSAR modeling studies [12, 17].

For each physicochemical property, two widely accepted statistical indicators are reported:

- r **(Pearson correlation coefficient)**, which quantifies the strength and direction of the linear relationship between $M_1^{\varepsilon}(G)$, $\xi VL(G)$ and the corresponding physicochemical property [14];

Table 2: Statistical performance of Multilinear QSPR models based on the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and VL-index $\xi VL(G)$

Physicochemical Property	r	**Correlation Strength**
Boiling Point (bp)	0.982	Very Strong
Molar Volume (mv)	0.961	Very Strong
Molar Refraction (mr)	0.968	Very Strong
Heat of Vaporization (hv)	0.961	Very Strong
Critical Temperature (ct)	0.984	Very Strong
Critical Pressure (cp)	-0.902	Very Strong
Surface Tension (st)	0.951	Very Strong

- **Correlation strength**, a qualitative classification based on the magnitude of $|r|$ following conventional statistical criteria [4].

The results clearly demonstrate the strong predictive capability of $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ across all the considered physicochemical properties of alkanes. All seven properties exhibit *very strong* correlations, with correlation coefficients ranging from $|r| = 0.951$ to 0.984. Among these, the critical temperature shows the highest positive correlation ($r = 0.984$), indicating an exceptionally strong linear dependence on the proposed topological descriptors.

Critical pressure exhibits a *very strong negative correlation* ($r = -0.902$), which is chemically meaningful. Higher values of $M_1^{\varepsilon}(G)$ and VL-index $\xi VL(G)$, associated with increased molecular branching and structural complexity, correspond to reduced intermolecular van der Waals interactions, leading to lower critical pressures.

Similarly, boiling point, molar volume, molar refraction, heat of vaporization, and surface tension all display very strong positive correlations ($r > 0.95$), highlighting the ability of $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ to effectively capture structural features governing thermodynamic and physicochemical behavior in alkane molecular graphs.

Overall, the consistently high correlation strengths observed in this study confirm that the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and VL-index $\xi VL(G)$ serves as a reliable and efficient single-parameter descriptor for QSPR modeling of alkane physicochemical properties. These findings are in good agreement with earlier QSPR studies employing Zagreb and VL-index-type topological indices [16, 11].

5 Conclusion

MultiLinear QSPR models were developed to assess the predictive ability of the eccentricity-based first Zagreb index $M_1^{\varepsilon}(G)$ and VL-index $\xi VL(G)$ for several physicochemical properties of alkanes. The results confirm that $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ is a highly effective single-parameter descriptor, exhibiting very strong correlations for all seven properties, with Pearson correlation coefficients in the range $|r| = 0.951$–0.984.

The strongest positive correlation is observed for critical temperature, while critical pressure shows a very strong negative correlation, both of which are chemically interpretable in terms of molecular branching and intermolecular interactions. These outcomes

indicate excellent linear relationships between alkane molecular structure and physicochemical behavior.

Overall, the consistently high correlation strengths demonstrate that $M_1^{\varepsilon}(G)$ and $\xi VL(G)$ successfully captures both local connectivity and global topological features, enabling reliable prediction of thermodynamic and physicochemical properties within a simple and interpretable framework. The findings are consistent with earlier QSPR studies employing Zagreb and VL-index-type indices and highlight the applicability of eccentricity-based descriptors in chemical graph theory and computational chemistry [12, 16, 11, 8, 5].

References

[1] Z. Aliannejadi and S. S. Alamoti, Graph edge hyper-Zagreb index: Application and thermodynamic property prediction for linear acenes molecules, *J. Discrete Math. Sci. Cryptogr.* **20**(1) (2026), 49–64. DOI: 10.47974/JDMSC-2135.

[2] M. S. Y. Alsharafi, A. Q. Alameri and Y. Zeren, The second hyper Zagreb index of $VC_5C_7[p,q]$ and $HC_5C_7[p,q]$ nanotubes, *J. Discrete Math. Sci. Cryptogr.* **27**(3) (2024), 915–928.

[3] S. S. Alamotie, M. Alaeiyan and A. Gilani, Studying thermodynamic properties of linear acenes molecules ($C_{4n+2}H_{2n+4}$), *J. Discrete Math. Sci. Cryptogr.* **22**(7) (2019), 1261–1268.

[4] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, 1988.

[5] P. Dankelmann, I. Gutman, and S. Mukwembi, On eccentricity-based molecular descriptors, *Discrete Applied Mathematics*, **160** (2012), 1444–1451.

[6] T. Deepika, An Emphatic study on VL Index and their bounds on tensor product of F sum graph, *TWMS Journal of Applied and Engineering Mathematics*, vol. 11, no. 2, pp. 374–385, 2021.

[7] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed., Wiley, New York, 1998.

[8] I. Gutman and N. Trinajstić, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, *Chemical Physics Letters*, **17** (1972), 535–538.

[9] I. Gutman, N. Trinajstić, C. F. Wilcox, and S. M. Wilcox, Graph theory and molecular orbitals. XII. Acyclic polyenes, *Journal of Chemical Physics*, **62** (1975), 3399–3405.

[10] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, New York, 2009.

[11] H. I. Ahmad, Correlation of Zagreb-type indices with physicochemical properties of hydrocarbons, *Heliyon*, **6** (2020), e03455.

[12] L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1986.

[13] V. H. Narendra and P. Mahalakshmi, *A theoretical study on correlation analysis between eccentricity of VL-index and the physical properties of alkanes*, *Asian Research Journal of Mathematics*, vol. 19, no. 10, pp. 103–113, 2023.

[14] J. L. Rodgers and W. A. Nicewander, Thirteen ways to look at the correlation coefficient, *The American Statistician*, **42** (1988), 59–66.

[15] P. G. Seybold, R. M. Bagal, and G. J. A. Madura, Molecular structure–property relationships, *Journal of Chemical Information and Computer Sciences*, **41** (2001), 156–160.

[16] K. Sunil, V. Lokesha, and P. S. Ranjini, Degree-based topological indices and QSPR analysis of alkanes, *Iranian Journal of Mathematical Chemistry*, **8** (2017), 89–104.

[17] N. Trinajstić, *Chemical Graph Theory*, 2nd ed., CRC Press, Boca Raton, 1992.

COMPETING INTERESTS DISCLAIMER:
Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.