Journal of Mathematical Control Science and Applications (JMCSA)
Vol. 3 No. 1 (January-June, 2017), ISSN : 0974-0570

EXPLORING THE PREDICTIVE POTENTIAL OF THE SECOND MULTIPLICATIVE HYPER-ZAGREB INDEX IN QSPR ANALYSIS OF ALDEHYDES

MANJULA K. M.

Abstract. Quantitative Structure–Property Relationship (QSPR) analysis is an effective computational approach for predicting the physicochemical properties of chemical compounds using graph-theoretical descriptors. In this study, the second multiplicative hyper-Zagreb index is employed to investigate the thermodynamic properties of twenty-four straight-chain aldehydes. The computed descriptor values are correlated with thermal energy (E_{th}), heat capacity (C_V), and entropy (S) using a logarithmic regression model. The predictive performance of the model is evaluated using the correlation coefficient (r) and the coefficient of determination (R^2). The results demonstrate that the second multiplicative hyper-Zagreb index is an effective molecular descriptor for modeling the thermodynamic properties of aldehydes.

1. Introduction

Aldehydes constitute one of the most significant classes of organic compounds and are characterized by the presence of a terminal carbonyl functional group ($(-CHO)$). They serve as fundamental intermediates in numerous chemical, pharmaceutical, agricultural, and industrial processes. Owing to their high chemical reactivity, aldehydes participate in a wide variety of organic transformations, including oxidation, reduction, condensation, and polymerization reactions. Lower aliphatic aldehydes such as ethanal and propanal are extensively employed in the manufacture of solvents, perfumes, plastics, dyes, and synthetic resins, whereas higher aldehydes find applications in the production of surfactants, lubricants, flavoring agents, and specialty chemicals. Their widespread utilization has motivated extensive investigations into their structural, thermodynamic, and physicochemical characteristics.

The physicochemical properties of organic molecules are closely associated with their molecular structures. Thermodynamic quantities such as thermal energy, heat capacity, and entropy provide valuable information regarding the energetic stability, molecular motion, and energy distribution within chemical compounds. Thermal energy represents the total internal energy arising from molecular vibrations, rotations, and translations under specified conditions. Heat capacity measures the amount of heat required to increase the temperature of a substance

2000 *Mathematics Subject Classification.* 05C90, 05C92, 05C07, 05C38.
Key words and phrases. Degree of a vertex, Topological indices, QSPR analysis.

by one degree, thereby reflecting the molecular flexibility and the number of accessible energy states. Entropy describes the degree of molecular disorder and the distribution of energy among available microscopic configurations. These parameters are indispensable in understanding molecular stability, reaction mechanisms, phase transitions, and chemical kinetics.

The homologous series of aliphatic aldehydes provides an ideal framework for investigating systematic variations in thermodynamic properties with increasing molecular size. As the number of carbon atoms increases from ethanal to pentacosanal, the molecular mass, chain length, and conformational flexibility also increase, resulting in predictable changes in thermal energy, heat capacity, and entropy. Such systematic trends facilitate the development of quantitative relationships between molecular structure and experimentally measured physicochemical properties. Consequently, homologous aldehydes are frequently employed as benchmark datasets for validating computational chemistry methods, molecular descriptors, and quantitative structureproperty relationship (QSPR) models [6].

In computational chemistry, reliable thermodynamic data are essential for validating theoretical predictions and establishing accurate structureproperty correlations. Thermal energy, heat capacity, and entropy values obtained through computational methods or experimental measurements serve as reference parameters for evaluating newly proposed molecular descriptors. These descriptors are particularly useful in graph-theoretical studies, where molecules are represented as mathematical graphs consisting of vertices and edges corresponding to atoms and chemical bonds, respectively. Graph-based descriptors capture structural information efficiently and have demonstrated considerable success in predicting diverse physicochemical, biological, and pharmacological properties without requiring expensive laboratory experiments.

The present study considers a dataset of twenty-four straight-chain aldehydes ranging from ethanal to pentacosanal. For each compound, the molecular formula, thermal energy ((E_{th})), heat capacity ((C_V)), and entropy ((S)) are collected and summarized in Table 1. These thermodynamic properties are employed as target variables for establishing quantitative structureproperty relationship models using graph-theoretical descriptors. The selected dataset offers a systematic increase in molecular complexity, enabling an effective assessment of the predictive capability of the proposed descriptors and regression models. Furthermore, the availability of experimentally or computationally determined thermodynamic data provides a reliable basis for evaluating the correlation between molecular topology and physicochemical behavior [8].

Overall, the structural and thermodynamic information presented in Table 1 forms the foundation for the subsequent QSPR analysis. By examining the relationship between graph-based descriptors and thermodynamic properties, the study aims to demonstrate the effectiveness of mathematical graph theory in predicting important physicochemical characteristics of aldehydes. Such investigations contribute to the development of efficient computational tools for molecular property prediction, molecular design, and chemical informatics, thereby reducing the dependence on extensive experimental measurements while enhancing the understanding of structureproperty relationships in organic compounds.

2. QSPR Analysis of Aldehydes

Quantitative Structure–Property Relationship (QSPR) analysis has emerged as an effective computational approach for establishing mathematical relationships between the structural characteristics of chemical compounds and their experimentally determined physicochemical properties. By representing molecules as mathematical graphs and extracting appropriate topological descriptors, QSPR models enable the prediction of molecular properties without the need for extensive laboratory experiments. Such predictive models significantly reduce experimental time and cost while providing valuable insights into the influence of molecular structure on physicochemical behavior. Consequently, graph-theoretical descriptors have been widely employed in chemistry, pharmaceutical sciences, material science, and chemical engineering for predicting various thermodynamic, electronic, and physicochemical properties of molecular compounds [2, 3].

In the present study, a homologous series of twenty-four straight-chain aldehydes, ranging from ethanal to pentacosanal, was selected for QSPR analysis. Aldehydes are an important class of organic compounds containing the terminal carbonyl functional group ($-$CHO) and are extensively used in the manufacture of pharmaceuticals, polymers, fragrances, solvents, and fine chemicals. Owing to the systematic increase in their carbon chain length, these compounds provide an ideal dataset for investigating the relationship between molecular topology and thermodynamic properties. The structural details together with the corresponding thermal energy (E_{th}), heat capacity (C_V), and entropy (S) of the selected aldehydes are presented in Table 2. These experimentally reported physicochemical properties are considered as the dependent variables in the present QSPR investigation [1].

Thermal energy (E_{th}) represents the internal energy associated with molecular vibrations, rotations, and translational motion, whereas heat capacity (C_V) measures the amount of heat required to raise the temperature of a substance by one unit. Entropy (S) describes the degree of molecular disorder and the distribution of energy among the accessible microscopic states. Since these thermodynamic properties are closely influenced by molecular size, connectivity, and structural complexity, they are well suited for correlation with graph-theoretical molecular descriptors [5].

In this study, the proposed graph-theoretical descriptor is employed as the independent variable to establish quantitative relationships with the selected physicochemical properties of aldehydes. The descriptor values are computed from the molecular graphs of the considered compounds, where vertices represent atoms and edges correspond to chemical bonds. These descriptor values are subsequently correlated with the thermal energy (E_{th}), heat capacity (C_V), and entropy (S) using linear, quadratic, cubic, logarithmic, power, and exponential regression models. The predictive capability of each regression model is assessed using the correlation coefficient (r) and the coefficient of determination (R^2). The computed values of the proposed descriptor for all considered aldehydes are presented in Table 1, and the corresponding regression analyses are discussed in the subsequent sections.

To investigate and predict the physicochemical properties of Aldehydes, the second multiplicative hyper-Zagreb index $HII_2(G)$ was employed as a molecular descriptor. The second multiplicative hyper-Zagreb index of a graph was introduced by Kulli in [4] and is defined as follows:

$$HII_2(G) = \prod_{uv \in E(G)} \left(d_G(u) \cdot d_G(v)\right)^2,$$

here $d_u(G)$ and $d_v(G)$ denote the degrees of the vertices u and v in the graph G, respectively, while $E(G)$ represents the edge set of G.

TABLE 1. Second Multiplicative Hyper-Zagreb Index of the considered aldehydes

Compound	$HII_2(G)$
Ethanal	16
Propanal	256
Butanal	4096
Pentanal	65536
Hexanal	1048576
Heptanal	16777216
Octanal	268435456
Nonanal	4294967296
Decanal	68719476736
Undecanal	1.09951×10^{12}
Dodecanal	1.75922×10^{13}
Tridecanal	2.81475×10^{14}
Tetradecanal	4.50360×10^{15}
Pentadecanal	7.20576×10^{16}
Hexadecanal	1.15292×10^{18}
Heptadecanal	1.84467×10^{19}
Octadecanal	2.95148×10^{20}
Nonadecanal	4.72237×10^{21}
Eicosanal	7.55579×10^{22}
Heneicosanal	1.20893×10^{24}
Docosanal	1.93428×10^{25}
Tricosanal	3.09485×10^{26}
Tetracosanal	4.95176×10^{27}
Pentacosanal	7.92282×10^{28}

TABLE 2. Structural details and their thermal energy (E_{th}), heat capacity (C_V) and entropy (S) for the aldehydes used in the present study

Compound	**Formula**	**No.**	$\mathbf{E}_{\text{th}}$ (kcal/mol)	$\mathbf{C_V}$ (cal/molK)	**S** (cal/molK)
Ethanal	C_2H_4O	1	40.074	10.472	62.483
Propanal	C_3H_6O	2	60.162	14.719	69.255
Butanal	C_4H_8O	3	80.131	19.195	76.521
Pentanal	$C_5H_{10}O$	4	100.175	23.569	83.110
Hexanal	$C_6H_{12}O$	5	132.672	26.877	90.594
Heptanal	$C_7H_{14}O$	6	139.940	32.586	97.626
Octanal	$C_8H_{16}O$	7	159.879	37.083	104.813
Nonanal	$C_9H_{18}O$	8	179.812	41.597	112.474
Decanal	$C_{10}H_{20}O$	9	199.764	46.095	119.946
Undecanal	$C_{11}H_{22}O$	10	219.717	50.587	127.356
Dodecanal	$C_{12}H_{24}O$	11	239.673	55.061	134.061
Tridecanal	$C_{13}H_{26}O$	12	259.626	59.553	141.197
Tetradecanal	$C_{14}H_{28}O$	13	279.545	64.080	149.267
Pentadecanal	$C_{15}H_{30}O$	14	299.493	68.575	156.600
Hexadecanal	$C_{16}H_{32}O$	15	319.434	73.070	163.611
Heptadecanal	$C_{17}H_{34}O$	16	339.444	77.480	171.456
Octadecanal	$C_{18}H_{36}O$	17	360.034	81.825	178.205
Nonadecanal	$C_{19}H_{38}O$	18	380.088	86.245	185.815
Eicosanal	$C_{20}H_{40}O$	19	400.070	90.739	193.172
Heneicosanal	$C_{21}H_{42}O$	20	420.055	95.226	200.520
Docosanal	$C_{22}H_{44}O$	21	440.039	99.715	207.846
Tricosanal	$C_{23}H_{46}O$	22	460.024	104.206	215.418
Tetracosanal	$C_{24}H_{48}O$	23	480.006	108.692	222.462
Pentacosanal	$C_{25}H_{50}O$	24	499.992	113.185	230.223

TABLE 3. Statistical parameters for the logarithmic QSPR model for the second multiplicative hyper-Zagreb index $HII_2(G)$ and the physicochemical properties of aldehydes

Physicochemical properties	R	R^2	Adjusted R^2	F	Sig.
$\mathbf{E}_{\text{th}}$ (kcal/mol)	1.000	1.000	1.000	67012.79	0.000
$\mathbf{C_V}$ (cal/molK)	1.000	1.000	1.000	293586.81	0.000
$\mathbf{S}$ (cal/molK)	1.000	1.000	1.000	383621.43	0.000

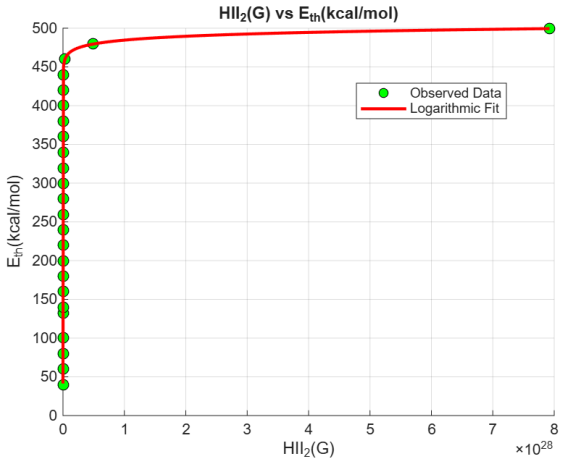
Logarithmic Regression Model:
By utilizing the data presented in Tables 1 and 2, logarithmic regression models are developed for the selected physicochemical properties of aldehydes. The logarithmic regression model is expressed as:

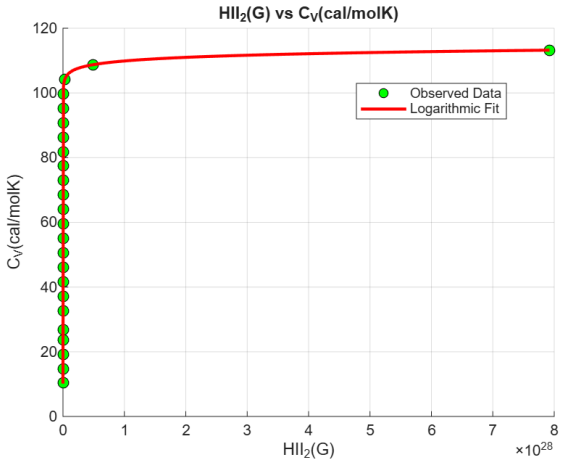
$$y = a + b\log(x),$$

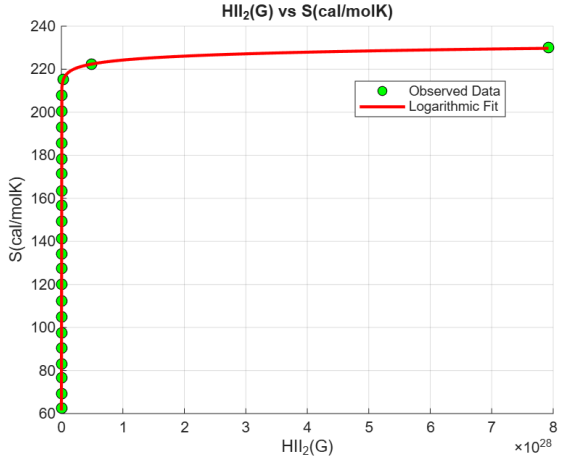
where y denotes the physicochemical property, a is the regression coefficient (intercept), b is the regression constant (slope), and x represents the molecular descriptor. The logarithmic regression models corresponding to the second multiplicative hyper-Zagreb index are given below:
The logarithmic regression models for the second multiplicative hyper-Zagreb index $HII_2(G)$ are given by:

$$\begin{aligned} \mathbf{E}_{\text{th}}\,(\text{kcal/mol}) &= 21.4813 + 7.1826\,\log(HII_2), \\ \mathbf{C_V}\,(\text{cal/molK}) &= 5.6697 + 1.6163\,\log(HII_2), \\ \mathbf{S}\,(\text{cal/molK}) &= 54.1387 + 2.6386\,\log(HII_2). \end{aligned}$$

The correlations of the second multiplicative hyper-Zagreb index, $HII_2(G)$, with the selected physicochemical properties of aldehydes obtained using the logarithmic regression model are depicted in the figures below:







3. Conclusion

In this study, the second multiplicative hyper-Zagreb index was employed as a molecular descriptor to investigate the thermodynamic properties of twenty-four straight-chain aldehydes through QSPR analysis. The descriptor values were correlated with thermal energy (E_{th}), heat capacity (C_V), and entropy (S) using a logarithmic regression model. The obtained values of the correlation coefficient (R) and the coefficient of determination (R^2) demonstrate a strong correlation between the proposed descriptor and the selected physicochemical properties. Therefore, the second multiplicative hyper-Zagreb index can be regarded as an effective molecular descriptor for predicting the thermodynamic properties of aldehydes and may serve as a useful tool in future QSPR investigations of related organic compounds.

References

[1] A. Alaghebandi and F. Shafiei, *QSPR modeling of heat capacity, thermal energy and entropy of aliphatic Aldehydes by using Topological Indices and MLR method*, Iranian Journal of Mathematical Chemistry, **7**(2) (2016), 235–251.

[2] D. Bonchev, *Over all connectivities/topological complexities: A new powerful tool for QSPR/QSAR*, J. Chem. Inf. Comput. Sci., **40** (2000), 934–941.

[3] A. A. Gakh, E. G. Gakh, B. G. Sumpter, D. W. Noid, *Neural network graph theory approach to the prediction of the physical properties of organic compounds*, J. Chem. Inf. Comput. Sci., **34**(4) (1994), 832–839.

[4] V. R. Kulli, *Multiplicative hyper-Zagreb indices and coindices of graphs*, International Journal of Pure Algebra, **6**(7) (2016), 342–347.

[5] Z. Lin, J. Xu, X. Zheng, Z. Li, *Study on quantitative structure- property relationship of chain hydrocarbons, aldehydes and alkanones by molecular distance-edge vector*, Acta Phys. Chem. Sin., **16** (2000), 153–161.

[6] V. Sharma, R. Goswami, A. K. Madan, *A novel highly discriminating topological descriptor for structureproperty and structureactivity studies*, J. Chem. Inf. Comput. Sci., **37** (1997), 273–282.

[7] G. H. Shirdel, H. Rezapour and A. M. Sayadi, *The Hyper–Zagreb Index of Graph Operations*, Iranian Journal of Mathematical Chemistry, **4**(2) (2013), 213–220.

[8] N. Trinajstić, *Chemical Graph Theory*, 2nd Edition, CRC Press, Boca Raton, (1992).

Manjula K. M.: Department of Mathematics, Government First Grade College for Women, Hassan-573 202, India.

E-mail address: manjula.km.gowda@gmail.com