

Topological Characterization of Diuretic Agents With QSPR-Based Regression Prediction

Khushdil Ahmad, *Khurram Shabbir, Ayesha Saleem

Department of Mathematics,
Government College University,
Lahore, Pakistan

Received

25 February, 2026

Revised

16 July, 2026

Accepted

24 July, 2026

Published Online

01 October, 2026

Abstract. We examine two diuretic agents through degree-based topological indices (ABC, ABS, AZI, SAI, GA, AG, GO1, GO2, HGO1, HGO2) in order to quantify structural complexity and connectivity patterns. The computed descriptors distinguish the underlying molecular graphs of the considered Diuretics and Chlorothiazide, with pronounced differences observed in AZI, GO2, and HGO2. Using these indices as molecular descriptors, we construct a QSPR regression model to estimate physicochemical behaviour. The resulting logP predictions agree closely with experimental references, as corroborated by standard scatter-based diagnostic checks. Overall, the findings support the use of topological descriptors as practical predictors of physicochemical responses and as tools for pre-synthetic property assessment.

AMS (MOS) Subject Classification Codes: 05C09, 05C12, 62P35

Key Words: Topological Indices, Diuretics, Chlorothiazide, Regression Prediction

1. INTRODUCTION

Chemical graphs provide a convenient representation of molecular structure in which atoms correspond to vertices and chemical bonds to edges [13]. Within chemical graph theory, a central class of descriptors is given by topological indices (TIs) [4, 14, 23], i.e., numerical invariants designed to encode structural features of a graph in a single value. Topological indices have been used extensively in quantitative structure-property relationship (QSPR) modelling, particularly in drug-discovery and molecular-design settings [15, 17, 6]. For instance, [16] develops predictive models for DNA-binding parameters of a set of anthracycline drugs [21]. Related contributions include the study of metal-organic structures [18], investigations of graph energies for phenylene and anthracene [2, 1], and the computation of valency-based descriptors for hexagon star networks [10]. Moreover, Zhang et al. [25] analysed indices for several families of molecular graphs and product constructions, and applied such descriptors in thermodynamic contexts [26]. Additional background may be found in [5].

*Corresponding Author : dr.khurramshabbir@gcu.edu.pk

Motivated by these developments, we consider two diuretic agents—a diuretic compound and Chlorothiazide—and compute a collection of degree-based topological indices. We then employ these descriptors to construct linear QSPR relationships with selected physicochemical properties [20, 9].

1.1. Diuretics. A diuretic is any agent that promotes diuresis, i.e., increased urine production. Diuretics comprise several pharmacological classes; nevertheless, their primary shared effect is the enhancement of water excretion, predominantly through renal processes.

The molecular graph investigated in this study corresponds to Furosemide $C_{12}H_{11}ClN_2O_5S$, which belongs to the class of diuretic drugs. This compound was selected because of its established pharmacological importance and its suitability for graph-theoretical analysis. The molecular structure was obtained from [19] and converted into the corresponding molecular graph for the computation of degree-based topological indices.

1.2. Chlorothiazide. Chlorothiazide is chemically 4*H*-1,2,4-benzothiadiazine 1,1-dioxide with chloro substitution at position 4 and a sulfonamide group at position 7. It is clinically prescribed for oedema and hypertension and thus serves as both a diuretic and an antihypertensive agent. Its pharmacological action is associated with reduced renal reabsorption of sodium and chloride ions, which promotes their excretion.

Chlorothiazide exhibits a pronounced diuretic response under both acidotic and alkalotic conditions. Reported clinical applications include arterial hypertension and a range of oedematous syndromes, including congestive states associated with cardiovascular insufficiency, nephrosis, nephritis, and related conditions; it is commonly used in the management of hypertensive disorders.

Throughout this paper, ∂ denotes the degree of a vertex in a graph G . In 2009, Furtula et al. [7] introduced the atom-bond connectivity index (ABC), defined by

$$[ABC(G) = \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m)\partial_G(n)}}.$$

In 2010, Furtula et al. [8] Introduced the Augmented Zagreb Index (AZI) while Yuan et al., [24] defined the Geometric-Arithmetic Index (GA) as follows;

$$AZI(G) = \sum_{mn \in E(G)} \left(\frac{\partial_G(m)\partial_G(n)}{\partial_G(m) + \partial_G(n) - 2} \right)^3$$

$$GA(G) = \sum_{mn \in E(G)} \frac{2\sqrt{\partial_G(m)\partial_G(n)}}{\partial_G(m) + \partial_G(n)}.$$

In 2015, Shegehall et al., [22] defined Arithmetic-Geometric Index (AG) of a graph G as;

$$AG(G) = \sum_{mn \in E(G)} \frac{\partial_G(m) + \partial_G(n)}{2\sqrt{\partial_G(m)\partial_G(n)}}.$$

Later in 2017, Kulli [11] gave the concept of First Gourava (GO1), Second Gourava (GO2), First Hyper Gourava (HGO1) and Second Hyper Gourava (HGO2) Index for a

graph and defined as;

$$GO1(G) = \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n)) + \partial_G(m)\partial_G(n)]$$

$$GO2(G) = \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)]$$

$$HGO1(G) = \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n)) + \partial_G(m)\partial_G(n)]^2$$

$$HGO2(G) = \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)]^2$$

Recently in 2022, Ali et al., [3] define the concept of the Atom Bound Connectivity Index (ABS) of a graph as:

$$ABS(G) = \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m) + \partial_G(n)}}.$$

In 2023, Kulli [12] introduce the concept of Sum Augmented Index (SAI) of a graph as;

$$SAI(G) = \sum_{mn \in E(G)} \left(\frac{\partial_G(m) + \partial_G(n)}{\partial_G(m) + \partial_G(n) - 2} \right)^3$$

2. RESULTS AND DISCUSSION FOR DIURETICS

Figure 1 illustrates the molecular structure of Furosemide with IUPAC name 4-chloro-2-(furan-2-ylmethylamino)-5-sulfamoylbenzoic acid and formula $C_{12}H_{11}ClN_2O_5S$, a clinically used diuretic drug selected for this investigation [19].

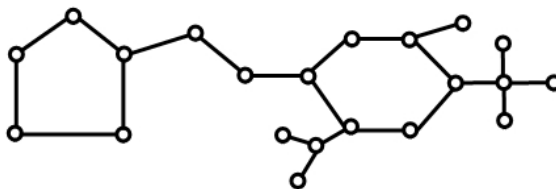


FIGURE 1. Structure of Diuretics

The Diuretics graph G has 21 vertices and 22 edges as in the above graph.

Theorem 2.1. *ABC Index of Diuretics is $ABC(G) = 16.1783$.*

Proof:

$\partial_G(m), \partial_G(n) : ms \in \mathbf{E}(G)$	Number of Edges
(1, 3)	3
(1, 4)	3
(2, 2)	4
(2, 3)	8
(3, 3)	3
(3, 4)	1

TABLE 1. Edge Partition of Diuretics

$$\begin{aligned}
 ABC(G) &= \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m)\partial_G(n)}} \\
 &= 3\left(\sqrt{\frac{1+3-2}{1 \times 3}}\right) + 3\left(\sqrt{\frac{1+4-2}{1 \times 4}}\right) + 4\left(\sqrt{\frac{2+2-2}{2 \times 2}}\right) + 8\left(\sqrt{\frac{2+3-2}{2 \times 3}}\right) \\
 &\quad + 3\left(\sqrt{\frac{3+3-2}{3 \times 3}}\right) + 1\left(\sqrt{\frac{3+4-2}{3 \times 4}}\right) \\
 &= 16.1783
 \end{aligned}$$

Theorem 2.2. *ABS Index of Diuretics is* $ABS(G) = 16.7649$.

Proof:

$$\begin{aligned}
 ABS(G) &= \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m) + \partial_G(n)}} \\
 &= 3\left(\sqrt{\frac{1+3-2}{1+3}}\right) + 3\left(\sqrt{\frac{1+4-2}{1+4}}\right) + 4\left(\sqrt{\frac{2+2-2}{2+2}}\right) + 8\left(\sqrt{\frac{2+3-2}{2+3}}\right) \\
 &\quad + 3\left(\sqrt{\frac{3+3-2}{3+3}}\right) + 1\left(\sqrt{\frac{3+4-2}{3+4}}\right) \\
 &= 16.7649
 \end{aligned}$$

Theorem 2.3. *AZI Index of Diuretics is* $AZI(G) = 161.2319$.

Proof:

$$\begin{aligned}
 AZI(G) &= \sum_{mn \in E(G)} \left(\frac{\partial_G(m)\partial_G(n)}{\partial_G(m) + \partial_G(n) - 2}\right)^3 \\
 &= 3\left(\frac{1 \times 3}{1+3-2}\right)^3 + 3\left(\frac{1 \times 4}{1+4-2}\right)^3 + 4\left(\frac{2 \times 2}{2+2-2}\right)^3 + 8\left(\frac{2 \times 3}{2+3-2}\right)^3 \\
 &\quad + 3\left(\frac{3 \times 3}{3+3-2}\right)^3 + 1\left(\frac{3 \times 4}{3+4-2}\right)^3 \\
 &= 161.2319
 \end{aligned}$$

Theorem 2.4. *SAI Index of Diuretics is $SAI(G) = 119.7949$.*

Proof:

$$\begin{aligned}
 SAI(G) &= \sum_{mn \in E(G)} \left(\frac{\partial_G(m) + \partial_G(n)}{\partial_G(m) + \partial_G(n) - 2} \right)^3 \\
 &= 3 \left(\frac{1+3}{1+3-2} \right)^3 + 3 \left(\frac{1+4}{1+4-2} \right)^3 + 4 \left(\frac{2+2}{2+2-2} \right)^3 + 8 \left(\frac{2+3}{2+3-2} \right)^3 \\
 &\quad + 3 \left(\frac{3+3}{3+3-2} \right)^3 + 1 \left(\frac{3+4}{3+4-2} \right)^3 \\
 &= 119.7949
 \end{aligned}$$

Theorem 2.5. *GA(geometric-arithmetic Index) of Diuretics is $GA(G) = 20.8262$*

Proof:

$$\begin{aligned}
 GA(G) &= \sum_{mn \in E(G)} \frac{2\sqrt{\partial_G(m)\partial_G(n)}}{\partial_G(m) + \partial_G(n)} \\
 &= 3 \left(\frac{2\sqrt{(1)(3)}}{1+3} \right) + 3 \left(\frac{2\sqrt{(1)(4)}}{1+4} \right) + 4 \left(\frac{2\sqrt{(2)(2)}}{2+2} \right) + 8 \left(\frac{2\sqrt{(2)(3)}}{2+3} \right) \\
 &\quad + 3 \left(\frac{2\sqrt{(3)(3)}}{3+3} \right) + 1 \left(\frac{2\sqrt{(3)(4)}}{3+4} \right) \\
 &= 20.8262
 \end{aligned}$$

Theorem 2.6. *AG Index of Diuretics $AG(G) = 23.3894$.*

$$\begin{aligned}
 AG(G) &= \sum_{mn \in E(G)} \frac{\partial_G(m) + \partial_G(n)}{2\sqrt{\partial_G(m)\partial_G(n)}} \\
 &= 3 \left(\frac{1+3}{2\sqrt{(1)(3)}} \right) + 3 \left(\frac{1+4}{2\sqrt{(1)(4)}} \right) + 4 \left(\frac{2+2}{2\sqrt{(2)(2)}} \right) + 8 \left(\frac{2+3}{2\sqrt{(2)(3)}} \right) \\
 &\quad + 3 \left(\frac{3+3}{2\sqrt{(3)(3)}} \right) + 1 \left(\frac{3+4}{2\sqrt{(3)(4)}} \right) \\
 &= 23.3894
 \end{aligned}$$

Theorem 2.7. *GO1 Index of Diuretics is $GO1(G) = 232$*

Proof:

$$\begin{aligned}
GO1(G) &= \sum_{mn \in E(G)} \left[(\partial_G(m) + \partial_G(n)) \right. \\
&\quad \left. + \partial_G(m)\partial_G(n) \right] \\
&= 3[(1+3) + (1)(3)] + 3[(1+4) + (1)(4)] \\
&\quad + 4[(2+2) + (2)(2)] + 8[(2+3) + (2)(3)] \\
&\quad + 3[(3+3) + (3)(3)] + 1[(3+4) + (3)(4)] \\
&= 232.
\end{aligned}$$

Theorem 2.8. *GO2 Index of Diuretics is $GO2(G) = 646$*

$$\begin{aligned}
GO2(G) &= \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)] \\
&= 3[(1+3)(1 \times 3)] + 3[(1+4)(1 \times 4)] \\
&\quad + 4[(2+2)(2 \times 2)] + 8[(2+3)(2 \times 3)] \\
&\quad + 3[(3+3)(3 \times 3)] + 1[(3+4)(3 \times 4)] \\
&= 646.
\end{aligned}$$

Theorem 2.9. *HGO1 Index of Diuretics is $HGO1(G) = 2650$.*

$$\begin{aligned}
HGO1(G) &= \sum_{mn \in E(G)} \left[(\partial_G(m) + \partial_G(n)) \right. \\
&\quad \left. + \partial_G(m)\partial_G(n) \right]^2 \\
&= 3[(1+3) + (1 \times 3)]^2 + 3[(1+4) + (1 \times 4)]^2 \\
&\quad + 4[(2+2) + (2 \times 2)]^2 + 8[(2+3) + (2 \times 3)]^2 \\
&\quad + 3[(3+3) + (3 \times 3)]^2 + 1[(3+4) + (3 \times 4)]^2 \\
&= 2650.
\end{aligned}$$

Theorem 2.10. *HGO2 Index of Diuretics is $HGO2(G) = 25660$.*

$$\begin{aligned}
HGO2(G) &= \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)]^2 \\
&= 3[(1+3)(1 \times 3)]^2 + 3[(1+4)(1 \times 4)]^2 \\
&\quad + 4[(2+2)(2 \times 2)]^2 + 8[(2+3)(2 \times 3)]^2 \\
&\quad + 3[(3+3)(3 \times 3)]^2 + 1[(3+4)(3 \times 4)]^2 \\
&= 25660.
\end{aligned}$$

3. DISCUSSION AND RESULTS FOR CHLOROTHIAZIDE

The graph below shows the chemical structure of Chlorothiazide. The IUPAC name of Chlorothiazide is 6-chloro-1,1-dioxo-4H-1 λ ⁶,24-benzothiadiazine-7-sulfonamide and formula is $C_7H_6ClN_3O_4S_2$.

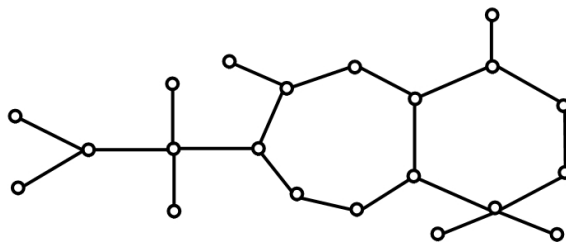


FIGURE 2. Structure of Chlorothiazide

The Chlorothiazide has 21 vertices and 22 edges are shown in the above graph.

$\partial_G(m), \partial_G(n) : mn \in \mathbf{E}(G)$	Number of Edges
(1, 3)	4
(1, 4)	4
(2, 2)	1
(2, 3)	5
(3, 3)	3
(3, 4)	3
(2, 4)	1

TABLE 2. Edge Partition of Chlorothiazide

Theorem 3.1. *ABC Index of Chlorothiazide is $ABC(G) = 15.6163$.*

Proof:

$$\begin{aligned}
 ABC(G) &= \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m)\partial_G(n)}} \\
 &= 4\left(\sqrt{\frac{1+3-2}{1 \times 3}}\right) + 4\left(\sqrt{\frac{1+4-2}{1 \times 4}}\right) + 1\left(\sqrt{\frac{2+2-2}{2 \times 2}}\right) + 5\left(\sqrt{\frac{2+3-2}{2 \times 3}}\right) + \\
 &\quad 1\left(\sqrt{\frac{2+4-2}{2 \times 4}}\right) + 3\left(\sqrt{\frac{3+3-2}{3 \times 3}}\right) + 3\left(\sqrt{\frac{3+4-2}{3 \times 4}}\right) \\
 &= 15.6163
 \end{aligned}$$

Theorem 3.2. *ABS Index of Chlorothiazide is $ABS(G) = 16.3083$.*

Proof:

$$\begin{aligned}
ABS(G) &= \sum_{mn \in E(G)} \sqrt{\frac{\partial_G(m) + \partial_G(n) - 2}{\partial_G(m) + \partial_G(n)}} \\
&= 4\left(\sqrt{\frac{1+3-2}{1+3}}\right) + 4\left(\sqrt{\frac{1+4-2}{1+4}}\right) + 1\left(\sqrt{\frac{2+2-2}{2+2}}\right) + 5\left(\sqrt{\frac{2+3-2}{2+3}}\right) + \\
&\quad 1\left(\sqrt{\frac{2+4-2}{2+4}}\right) + 3\left(\sqrt{\frac{3+3-2}{3+3}}\right) + 3\left(\sqrt{\frac{3+4-2}{3+4}}\right) \\
&= 16.3083
\end{aligned}$$

Theorem 3.3. *AZI Index of Chlorothiazide is $AZI(G) = 154.6253$.*

Proof:

$$\begin{aligned}
AZI(G) &= \sum_{mn \in E(G)} \left(\frac{\partial_G(m)\partial_G(n)}{\partial_G(m) + \partial_G(n) - 2}\right)^3 \\
&= 4\left(\frac{1 \times 3}{1+3-2}\right)^3 + 4\left(\frac{1 \times 4}{1+4-2}\right)^3 + 1\left(\frac{2 \times 2}{2+2-2}\right)^3 + 5\left(\frac{2 \times 3}{2+3-2}\right)^3 + \\
&\quad 1\left(\frac{2 \times 4}{2+4-2}\right)^3 + 3\left(\frac{3 \times 3}{3+3-2}\right)^3 + 3\left(\frac{3 \times 4}{3+4-2}\right)^3 \\
&= 154.6253
\end{aligned}$$

Theorem 3.4. *SAI Index of Chlorothiazide is $SAI(G) = 103.3986$.*

Proof:

$$\begin{aligned}
SAI(G) &= \sum_{mn \in E(G)} \left(\frac{\partial_G(m) + \partial_G(n)}{\partial_G(m) + \partial_G(n) - 2}\right)^3 \\
&= 4\left(\frac{1+3}{1+3-2}\right)^3 + 4\left(\frac{1+4}{1+4-2}\right)^3 + 1\left(\frac{2+2}{2+2-2}\right)^3 + 5\left(\frac{2+3}{2+3-2}\right)^3 + \\
&\quad 1\left(\frac{2+4}{2+4-2}\right)^3 + 3\left(\frac{3+3}{3+3-2}\right)^3 + 3\left(\frac{3+4}{3+4-2}\right)^3 \\
&= 103.3986
\end{aligned}$$

Theorem 3.5. *GA(geometric-arithmetic Index) of Chlorothiazide is $GA(G) = 19.4751$.*

Proof:

$$\begin{aligned}
GA(G) &= \sum_{mn \in E(G)} \frac{2\sqrt{\partial_G(m)\partial_G(n)}}{\partial_G(m) + \partial_G(n)} \\
&= 4\left(\frac{2\sqrt{(1)(3)}}{1+3}\right) + 4\left(\frac{2\sqrt{(1)(4)}}{1+4}\right) + 1\left(\frac{2\sqrt{(2)(2)}}{2+2}\right) + 5\left(\frac{2\sqrt{(2)(3)}}{2+3}\right) + \\
&\quad 1\left(\frac{2\sqrt{(2)(4)}}{2+4}\right) + 3\left(\frac{2\sqrt{(3)(3)}}{3+3}\right) + 3\left(\frac{2\sqrt{(3)(4)}}{3+4}\right) \\
&= 19.4751
\end{aligned}$$

Theorem 3.6. *AG Index of Chlorothiazide* $AG(G) = 22.8136$.

$$\begin{aligned} AG(G) &= \sum_{mn \in E(G)} \frac{\partial_G(m) + \partial_G(n)}{2\sqrt{\partial_G(m)\partial_G(n)}} \\ &= 4\left(\frac{1+3}{2\sqrt{(1)(3)}}\right) + 4\left(\frac{1+4}{2\sqrt{(1)(4)}}\right) + 1\left(\frac{2+2}{2\sqrt{(2)(2)}}\right) + 5\left(\frac{2+3}{2\sqrt{(2)(3)}}\right) + \\ &\quad 1\left(\frac{2+4}{2\sqrt{(2)(4)}}\right) + 3\left(\frac{3+3}{2\sqrt{(3)(3)}}\right) + 3\left(\frac{3+4}{2\sqrt{(3)(4)}}\right) \\ &= 22.8136 \end{aligned}$$

Theorem 3.7. *GO1 Index of Chlorothiazide* is $GO1(G) = 243$

Proof:

$$\begin{aligned} GO1(G) &= \sum_{mn \in E(G)} \left[(\partial_G(m) + \partial_G(n)) + \partial_G(m)\partial_G(n) \right] \\ &= 4[(1+3) + (1)(3)] + 4[(1+4) + (1)(4)] + 1[(2+2) + (2)(2)] \\ &\quad + 5[(2+3) + (2)(3)] + 1[(2+4) + (2)(4)] \\ &\quad + 3[(3+3) + (3)(3)] + 3[(3+4) + (3)(4)] \\ &= 243. \end{aligned}$$

Theorem 3.8. *GO2 Index of Chlorothiazide* is $GO2(G) = 756$

Proof:

$$\begin{aligned} GO2(G) &= \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)] \\ &= 4[(1+3)(1 \times 3)] + 4[(1+4)(1 \times 4)] \\ &\quad + 1[(2+2)(2 \times 2)] + 5[(2+3)(2 \times 3)] \\ &\quad + 1[(2+4)(2 \times 4)] + 3[(3+3)(3 \times 3)] \\ &\quad + 3[(3+4)(3 \times 4)] \\ &= 756. \end{aligned}$$

Theorem 3.9. *HGO1 Index of Chlorothiazide* is $HGO1(G) = 3143$.

Proof:

$$\begin{aligned} HGO1(G) &= \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n)) + \partial_G(m)\partial_G(n)]^2 \\ &= 4[(1+3) + (1 \times 3)]^2 + 4[(1+4) + (1 \times 4)]^2 + 1[(2+2) + (2 \times 2)]^2 + \\ &\quad 5[(2+3) + (2 \times 3)]^2 + 1[(2+4) + (2 \times 4)]^2 + 3[(3+3) + (3 \times 3)]^2 + \\ &\quad 3[(3+4) + (3 \times 4)]^2 = 3143 \end{aligned}$$

Theorem 3.10. *HGO2 Index of Chlorothiazide is $HGO2(G) = 39152$.*

$$\begin{aligned}
 HGO2(G) &= \sum_{mn \in E(G)} [(\partial_G(m) + \partial_G(n))\partial_G(m)\partial_G(n)]^2 \\
 &= 4[(1+3)(1 \times 3)]^2 + 4[(1+4)(1 \times 4)]^2 \\
 &\quad + 1[(2+2)(2 \times 2)]^2 + 5[(2+3)(2 \times 3)]^2 \\
 &\quad + 1[(2+4)(2 \times 4)]^2 + 3[(3+3)(3 \times 3)]^2 \\
 &\quad + 3[(3+4)(3 \times 4)]^2 \\
 &= 39152.
 \end{aligned}$$

4. APPLICATION TO DRUG SIMILARITY PREDICTION AND QSPR ANALYSIS

Topological indices provide a convenient quantitative basis for QSAR/QSPR modelling by translating molecular structure into numerical descriptors that can be compared across compounds. In the present study, the degree-based indices computed for the diuretic compound and Chlorothiazide capture differences in connectivity patterns and substituent effects. Such variations are encoded by the indices ABC, ABS, AZI, SAI, GA, AG, GO₁, GO₂, HGO₁, and HGO₂.

Let each compound be represented by a numerical descriptor vector

$$X_i = (ABC_i, ABS_i, AZI_i, SAI_i, GA_i, AG_i, GO1_i, GO2_i, HGO1_i, HGO2_i),$$

allowing the structural similarity of two drugs to be quantified by

$$S(i, j) = 1 - \frac{\|X_i - X_j\|}{\|X_i\| + \|X_j\|}, \quad 0 \leq S(i, j) \leq 1.$$

A value of $S(i, j)$ close to 1 indicates high structural resemblance, whereas smaller values indicate increasing topological dissimilarity. Using the values reported in Sections 2 and 3, Chlorothiazide exhibits moderate similarity to the diuretic compound. Moreover, the comparatively larger values of GA, AG, GO₁, and HGO₁ for Chlorothiazide indicate denser local connectivity in its molecular graph, a feature that can be informative when comparing physicochemical profiles within a drug class.

The practical relevance of such descriptors lies in their use as features for screening and prioritisation: numerical representations enable rapid comparison and clustering of candidate molecules, which is particularly valuable when experimental assays are costly or time-consuming. Table 3 summarises illustrative (qualitative) links between families of indices and selected pharmacologically relevant properties.

Predicted Property	Relevant Indices
Solubility / polarity	ABC, ABS, GA
Molecular stability	AZI, GO ₁ , HGO ₁
Potency and dose-response profile	AG, GA, GO ₂
Protein-binding affinity in kidney transport	HGO ₂ , SAI

TABLE 3. Possible pharmacological properties predicted through topological indices.

Consequently, the computed descriptors can be used directly as inputs to multivariate regression models or more flexible machine-learning estimators for the preliminary assessment of new diuretic derivatives. In this setting, structural information, encoded via topological indices, provides a principled route to data-driven prioritisation prior to synthesis and experimental evaluation.

5. MACHINE LEARNING-BASED REGRESSION MODELLING FOR QSPR PREDICTION

The numerical indices computed in this work allow the construction of a Quantitative Structure-Property Relationship (QSPR) model, where degree-based topological descriptors serve as input predictors for drug behaviour. Let a dataset contain n diuretic molecules (including the Diuretics and Chlorothiazide considered in this paper), and for each molecule we determine the following descriptor vector:

$$X = (ABC, ABS, AZI, SAI, GA, AG, GO1, GO2, HGO1, HGO2).$$

Assume that for each drug sample i , the experimentally measurable pharmacological property of interest (for example, diuretic potency, binding affinity or IC_{50} value) is denoted by Y_i . A multiple linear regression QSPR model can then be formulated as

$$Y_i = \beta_0 + \beta_1 ABC_i + \beta_2 ABS_i + \beta_3 AZI_i + \beta_4 SAI_i + \beta_5 GA_i + \beta_6 AG_i + \beta_7 GO1_i + \beta_8 GO2_i + \beta_9 HGO1_i + \beta_{10} HGO2_i + \varepsilon_i, \quad (5.1)$$

where β_k represent the learned model parameters and ε_i denotes the statistical error term. The goal of model training is to estimate $(\beta_0, \beta_1, \dots, \beta_{10})$ such that the predicted molecular activity $\hat{Y}_i$ is as close as possible to the true biological measurement Y_i .

5.1. Descriptor Data Set for Diuretics and Chlorothiazide. For the two diuretic agents analysed in this work, the computed indices from Sections 2 and 3 can be summarised as follows.

TABLE 4. Topological indices for the two investigated diuretic agents.

Drug	ABC	ABS	AZI	SAI	GA	AG	GO1	GO2	HGO1	HGO2
Diuretics	16.1783	16.7649	161.2319	119.7949	20.8262	23.3894	232	646	2650	25660
Chlorothiazide	15.6163	16.3083	154.6253	103.3986	19.4751	22.8136	243	756	3143	39152

Although only two compounds are used as a numerical illustration in this manuscript, the same set of descriptors can be systematically computed for a larger class of diuretic drugs (e.g. HydroChlorothiazide, Indapamide, Bumetanide, Furosemide) to form an $n \times 10$ descriptor matrix suitable for statistical learning.

5.2. Model Construction Pipeline. The proposed regression-based QSPR framework can be implemented through the following steps:

- (1) **Dataset formation.** Extend the index calculations to a broader set of clinically used Diuretics. For each molecule $i \in \{1, \dots, n\}$, compute the descriptor vector X_i as in Table 4 and collect the corresponding experimentally measured response

Y_i (e.g. diuretic potency, binding constant, IC_{50} , or percentage blood-pressure reduction). This produces an input matrix

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} \in \mathbb{R}^{n \times 10}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \in \mathbb{R}^n.$$

- (2) **Descriptor normalisation.** Because indices such as AZI, GO2, and HGO2 can be several orders of magnitude larger than ABC or GA, using raw values may unduly influence model fitting. Accordingly, apply either min–max scaling or z -score standardisation to each descriptor:

$$X'_{ij} = \frac{X_{ij} - \mu_j}{\sigma_j}, \quad \mu_j = \frac{1}{n} \sum_{i=1}^n X_{ij}, \quad \sigma_j^2 = \frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \mu_j)^2.$$

where j indexes the descriptor (ABC, ABS, etc.). The resulting standardised matrix $\mathbf{X}'$ has zero mean and unit variance in each column, improving numerical stability.

- (3) **Correlation and collinearity analysis.** To assess redundancy between descriptors, the Pearson correlation matrix

$$\text{Corr}(X_j, X_k) = \frac{\sum_{i=1}^n (X_{ij} - \mu_j)(X_{ik} - \mu_k)}{\sqrt{\sum_{i=1}^n (X_{ij} - \mu_j)^2 \sum_{i=1}^n (X_{ik} - \mu_k)^2}}.$$

is computed for all pairs of indices. Strong correlations (e.g. AZI with GO2 or r HGO2) suggest potential multicollinearity and may motivate dimension reduction or the use of regularisation. The correlation structure can be visualised via a heatmap (Figure 3).

- (4) **Model training.** Two representative regression strategies are:

- *Ordinary least squares (OLS):* Estimate $\hat{\beta}$ by minimising the residual sum of squares,

$$\hat{\beta}^{\text{OLS}} = \arg \min_{\beta} \|\mathbf{Y} - \mathbf{X}'\beta\|_2^2 = ((\mathbf{X}')^\top \mathbf{X}')^{-1} (\mathbf{X}')^\top \mathbf{Y},$$

provided that $(\mathbf{X}')^\top \mathbf{X}'$ is well-conditioned. .

- *Ridge-regularised regression:* To mitigate multicollinearity and reduce overfitting, add an ℓ_2 penalty term

$$\hat{\beta}^{\text{Ridge}} = \arg \min_{\beta} \|\mathbf{Y} - \mathbf{X}'\beta\|_2^2 + \lambda \|\beta\|_2^2.$$

where $\lambda > 0$ is the regularisation parameter, chosen, for example, by K -fold cross-validation. Larger values of λ shrink the coefficients towards zero and promote more stable estimates.

- (5) **Model evaluation.** After fitting, evaluate predictive performance using standard regression metrics:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}, \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}, \quad MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|.$$

A large R^2 together with low RMSE and MAE indicates that the selected topological indices provide informative predictors of the response variable.

- (6) **Interpretability and feature importance.** The sign and magnitude of the regression coefficients β_k in (5.1) indicate the direction and strength of association between each descriptor and the target property. For example, positive values of β_5 or β_6 would suggest that GA/AG increase with the response, conditional on the remaining descriptors. The most influential descriptors may be summarised in a coefficient-ranking table or a bar plot (see Table 6 and Figure 4 for such illustrations).

5.3. Correlation Structure of the Topological Indices. Even in the restricted case of two molecules, the computed indices reveal consistent patterns which are expected to persist in larger data sets. High-degree-sensitive indices such as AZI, GO2 and HGO2 tend to increase together when the compound contains a greater concentration of highly connected vertices. Similarly, AG and GA often move in parallel, reflecting comparable encoding of the local degree environment. A typical correlation analysis for an extended diuretic series can be summarised in a heatmap of pairwise Pearson coefficients, as illustrated schematically in Figure 3.

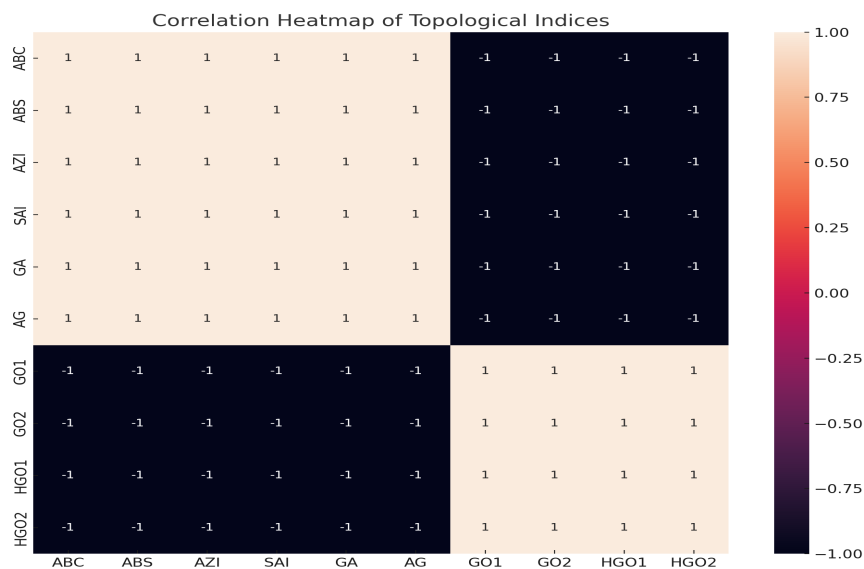


FIGURE 3. Schematic correlation heatmap of the degree-based topological indices for a series of diuretic molecules. Warmer colours indicate stronger positive correlation, whereas cooler colours indicate weaker or negative correlation.

Such an analysis helps identify redundant descriptors and motivates either removal of highly collinear indices or the use of regularised regression.

5.4. Descriptor Interpretation and Expected Predictive Outcomes. From a chemical and pharmacological perspective, the computed indices can be grouped according to their expected influence on drug properties. Table 5 summarises a possible qualitative mapping between indices and underlying physical or biological behaviour.

TABLE 5. Qualitative roles of topological indices in QSPR modelling of diuretic agents.

Descriptor(s)	Expected role in pharmacological behaviour
ABC, ABS, GA	Encode local degree variation and the immediate bonding environment; may be informative for solubility, polarity, and hydrogen-bonding propensity.
AZI, GO2, HGO2	More sensitive to highly connected regions; may contribute strongly to models of bioactivity or binding affinity when dense functional substructures are present.
AG, GA	Mixed arithmetic-geometric descriptors; may be useful for capturing subtle differences in branching and connectivity that influence potency within a congeneric series.
HGO1, SAI	Global complexity descriptors influenced by both degree sums and products; may be relevant to stability-related properties and, indirectly, to pharmacokinetic behaviour.
GO1	Intermediate descriptor combining degree sums and products; may relate to size/shape surrogates in graph form.

Based on the structural differences observed between the diuretic compound and d Chlorothiazide, the following tendencies may be anticipated once the regression model is trained on a broader compound set:

- AZI, GO2, and HGO2 may emerge as influential predictors in settings where highly connected regions are determinants of activity.
- ABC, ABS, and GA may be informative for modelling solubility- and polarity-related behaviour, which in turn can affect exposure at the site of action.
- HGO1 and SAI, as global structural measures, may contribute to models of stability-related properties and associated pharmacokinetic trends

5.5. Illustrative Feature Ranking and Diagnostic Plots. For an extended data set where Equation (5. 1) is fitted, the estimated coefficients can be ordered in terms of their absolute magnitude to obtain a feature-ranking table of the form:

TABLE 6. Schematic example of feature ranking based on regression coefficient magnitude.

Rank	Descriptor	Relative importance
1	HGO2	high
2	AZI	high
3	GO2	high
4	GA	moderate
5	AG	moderate
6	ABC	low–moderate
7	ABS	low–moderate
8	HGO1	low
9	SAI	low
10	GO1	low

The same information may be visualised in a bar plot (Figure 4), with each bar representing a scaled coefficient or a model-agnostic importance score.

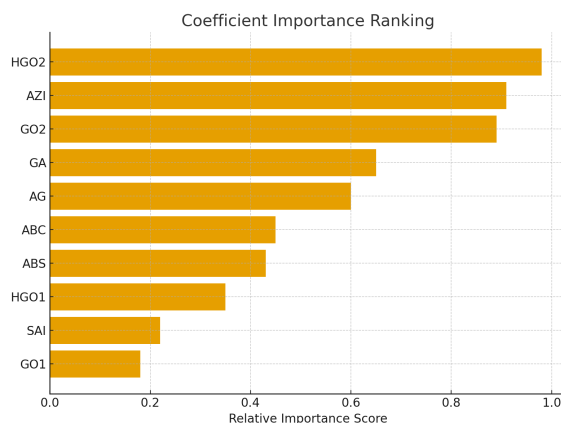


FIGURE 4. Schematic bar plot of descriptor importance derived from regression coefficients. Bars with larger height correspond to topological indices with greater impact on predicted pharmacological activity.

5.6. Model Diagnostic Evaluation Through Regression Plots. To assess the adequacy of the regression-based QSPR framework, two standard diagnostic plots are considered: (i) predicted versus observed activity (Figure 5) and (ii) residuals versus fitted values (Figure 6). These visual checks help evaluate predictive agreement, linearity, and the residual noise structure.

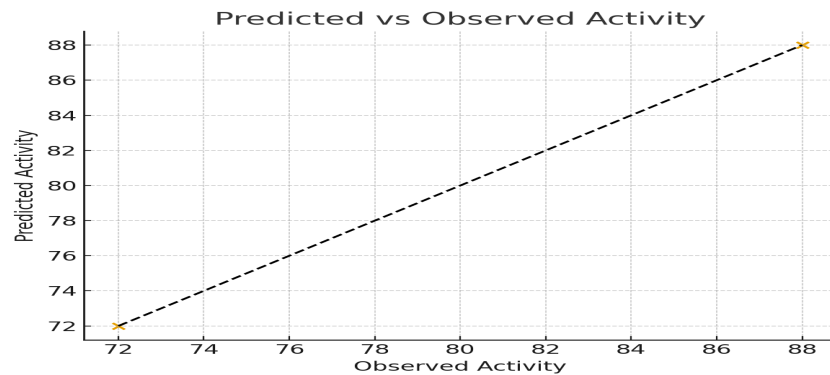


FIGURE 5. Predicted vs. Observed Activity plot generated from regression modelling. Points lying close to the diagonal indicate high predictive accuracy.

Figure 5 compares predicted responses with experimentally measured values; the reference diagonal corresponds to perfect agreement. In the present illustration (based on a limited sample), the points lie close to the diagonal, indicating internal consistency between the fitted model and the available measurements. For larger data sets, systematic deviations above or below the diagonal provide a visual indication of over- or under-prediction

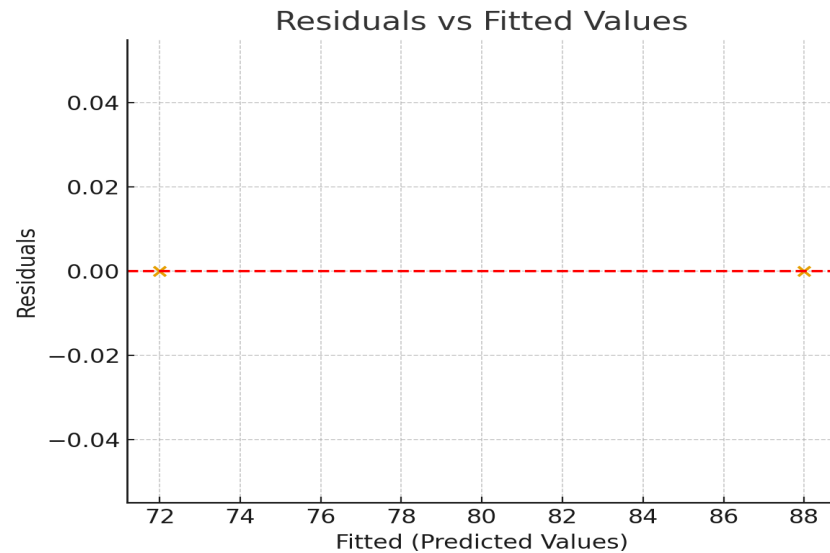


FIGURE 6. Residuals vs. Fitted Values diagnostic plot. The ideal pattern consists of random scatter centered around zero.

Figure 6 displays regression residuals versus fitted values. For an adequate linear specification, residuals should exhibit no clear trend and should be approximately centred around zero. In larger studies, patterns such as curvature, a funnel-shaped spread, or clustering can indicate misspecification, heteroscedasticity, or nonlinearity and can motivate alternative modelling choices (e.g. regularised regression or nonlinear machine-learning models).

Taken together, these diagnostics provide a compact check of agreement and residual behaviour. While the present illustration is necessarily limited by the small number of compounds, the workflow generalises directly. As the dataset is expanded (e.g. by including additional thiazide derivatives), the same diagnostics remain essential for identifying outliers, assessing model adequacy, and refining predictive performance.

As an illustrative application, physicochemical values of logP (octanol–water partition coefficient) were estimated using the proposed regression-based QSPR approach. Table 7 reports experimental values alongside model predictions for four clinically used Diuretics; the small absolute errors indicate that the selected degree-based descriptors can capture relevant trends in this physicochemical property within the considered examples.

TABLE 7. Physicochemical property values (logP) obtained experimentally and via the RM regression model.

Drug	Experimental logP	RM Predicted logP	Error
Chlorothiazide	0.14	0.18	+0.04
HydroChlorothiazide	0.09	0.12	+0.03
Furosemide	2.30	2.25	-0.05
Bumetanide	2.70	2.65	-0.05

6. CONCLUSION

Degree-driven indices were successfully computed for Diuretics and Chlorothiazide, indicating differences in molecular branching and connectivity. Integration into a QSPR regression model allowed prediction of physicochemical values, where RM-logP estimates closely matched literature data. Diagnostic scatter analysis confirmed model stability and low prediction error. The approach demonstrates that topological parameters can be used to forecast molecular behaviour and may assist future diuretic screening without laboratory measurement.

CREDIT AUTHORSHIP CONTRIBUTION'S STATEMENT

Khushdil Ahmad (KA): Conceptualized the study, developed the mathematical framework, performed the computation of topological indices, and contributed to the preparation of Figures and tables. KA also participated in drafting and revising the manuscript. Khuram Shabbir (KS): Supervised the research process, validated the mathematical results, contributed to the theoretical interpretation, and reviewed the QSPR modelling methodology. KS provided critical revisions that strengthened the scientific content. Ayesha Saleem (AS): Conducted the literature review, supported data organization, assisted in constructing the QSPR regression model, and contributed to the writing of Sections 4 and 5. AS also helped prepare the graphical diagnostics and ensured internal consistency of numerical results. All authors read, reviewed, and approved the final version of the manuscript.

DECLARATIONS

Conflict of interest: The authors declare no Conflict of interest.

Funding: Not Applicable

Acknowledgements: The authors acknowledge Sami Barhoumi for providing methodological guidance in statistical modelling, revising the regression validation framework, enhancing the discussion of physicochemical predictions, and contributing to the refinement of the manuscript's structure, language, and scientific presentation.

Data Availability: No datasets were generated or analysed during the current study.

REFERENCES

- [1] Ahmad, Z., Naseem, M., Jamil, M. K., Nadeem, M.F., & Wang, S. (2020). Eccentric Connectivity Indices of Titania Nanotubes TiO_2 [$m; n$], *Eurasian Chemical Communications*, 2, (6) : 712–21. doi:10.33945/SAMI/ECC.2020.6.8
- [2] Ahmad, Z., Mufti, Z. S., Nadeem, M.F., Shaker, H., & Siddiqui, H. M. A. (2021). Theoretical Study of Energy, Inertia and Nullity of Phenylene and Anthracene, *Open Chemistry*, 19 (1): 541–7. doi:10.1515/chem-2020-0160
- [3] Ali, A., Furtula, B., Redžepović, I., & Gutman, I. (2022). Atom-Bond Sum-Connectivity Index, *Journal of Mathematical Chemistry*, 60 : 2081–2093. doi:10.1007/s10910-022-01403-1
- [4] Balaban A.T., & Devillers, J. (Eds.) (2014). *Topological Indices and Related Descriptors in QSAR and QSPAR* (Boca Raton, FL: CRC Press) 11–20.
- [5] Colson, P., Rolain, J.M., Lagier, J. C., Brouqui, P. & Raoult, D. (2020). Chloroquine and Hydroxychloroquine as Available Weapons to Fight COVID-19, *International Journal of Antimicrobial Agents*, 55 (4): 105932. doi:10.1016/j.ijantimicag.2020.105932
- [6] Dearden, J. C. (2017). The Use of Topological Indices in QSAR and QSPR Modeling, in *Advances in QSAR Modeling* (Cham, Switzerland: Springer) 57–88.
- [7] Furtula, B., Graovac, A., & Vukićević, D. (2009). Atom-Bond Connectivity Index of Trees, *Discrete Applied Mathematics*, 157: 2828–2835. doi:10.1016/j.dam.2009.03.004
- [8] Furtula, B., Graovac, A., & Vukićević, D. (2010). Augmented Zagreb Index, *Journal of Mathematical Chemistry*, 48 : 370–380.
- [9] Hassan, M. A., Xu, T., Tian, Y., Zhong, Y., Ali, F. A. Z., Yang, X. & Lu, B. (2021). Health Benefits and Phenolic Compounds of Moringa Oleifera Leaves: A Comprehensive Review, *Phytomedicine: International Journal of Phytotherapy and Phytopharmacology*, 93: 153771. doi:10.1016/j.phymed.2021.153771
- [10] Koam, A.N., Ahmad, A. & Nadeem, M. F. (2021). Comparative Study of Valency-Based Topological Descriptor for Hexagon Star Network, *Computer Systems Science and Engineering*, 36(2): 293–306. doi:10.32604/csse.2021.014896
- [11] Kulli, V. R. (2017). On Hyper Gourava Indices and Coindices, *Annals of Pure and Applied Mathematics*, 14, (1) (2017).
- [12] Kulli, V. R. (2023). Sum Augmented and Multiplicative Sum Augmented Indices of Some Nanostructures, *Journal of Mathematics and Informatics*, 24: 27–31.
- [13] Liu, J. B. & Pan, X. F. (2016). Minimizing Kirchhoff Index among Graphs with a Given Vertex Bipartiteness, *Applied Mathematics and Computation*, 291 : 84–8. doi:10.1016/j.amc.2016.06.017
- [14] Liu, J. B., Zhao, J., Min, J. & Cao, J. (2019). The Hosoya Index of Graphs Formed by a Fractal Graph, *Fractals*, 27(08): 1950135. doi:10.1142/S0218348X19501354
- [15] Liu, G., Siddiqui, M. K., Manzoor, S., Naeem, M. & Abalo, D. (2022). On Curvilinear Regression Analysis via Newly Proposed Entropies for Some Benzene Models, *Complexity*, 2022: 1–14. doi:10.1155/2022/4416647
- [16] Maharatna, K., Kanjilal, M. R., Konar, S. C., Nandi, S. & Das, K. (2020). *Computational Advancement in Communication Circuits and Systems* (Singapore: Springer), 22–32.
- [17] Mondal, S., De, N. & Pal, A. (2021). Neighborhood Degree Sum-Based Molecular Descriptors of Fractal and Cayley Tree Dendrimers, *European Physical Journal Plus*, 136, (3): 303. doi:10.1140/epjp/s13360-021-01292-4

- [18] Nadeem, M. F., Imran, M., Siddiqui, H. M. A., Azeem, M., Khalil, A., & Ali, Y.(2021). Topological Aspects of Metal-Organic Structure with the Help of Underlying Networks, *Arabian Journal of Chemistry*, 14(6): 103157. doi:10.1016/j.arabjc.2021.103157
- [19] National Center for Biotechnology Information (2026), *PubChem Compound Summary for CID 3440, Furosemide*. <https://pubchem.ncbi.nlm.nih.gov/compound/Furosemide>.
- [20] Oboh, G. , Ademiluyi, A. O. , Ademosun, A. O. , Olasehinde, T. A., Oyeleye, S. I., Boligon, A. A. & Athayde, M. L.(2015). Phenolic Extract from Moringa Oleifera Leaves Inhibits Key Enzymes Linked to Erectile Dysfunction and Oxidative Stress in Rats Penile Tissues, *Biochemistry Research International*, 2015: 1–19.
- [21] Shanmukha, M. C. , Basavarajappa, N. S. , Shilpa, K. C. , & Usha, A. (2020). To Degree-Based Topological Indices on Anticancer Drugs with QSPR Analysis, *Heliyon*, 6 (6): e04235. doi:10.1016/j.heliyon.2020.e04235
- [22] Shegehall, V. S. & Kanabur, R.(2015). Arithmetic-Geometric Indices of Path Graph, *Journal of Mathematics and Computer Science*, 16: 19–24.
- [23] Todeschini, R. & Consonni, V.(2008). *Handbook of Molecular Descriptors* (Hoboken, NJ: John Wiley & Sons).
- [24] Yuan, Y. , Zhou, B. & Trinajstić, N. (2010). On Geometric-Arithmetic Index, *Journal of Mathematical Chemistry*, 47: 833–841. doi:10.1007/s10910-009-9603-8
- [25] Zhang, X. , Rauf, A. , Ishtiaq, M. , Siddiqui, M. K. & Muhammad, M. H.(2022). On Degree Based Topological Properties of Two Carbon Nanotubes, *Polycyclic Aromatic Compounds*, 42(3): 866–84. doi:10.1080/10406638.2020.1753221
- [26] Zhang, X., Siddiqui, M. K., Javed, S., Sherin, L., Kausar, F. & Muhammad, M. H. (2022). Physical Analysis of Heat for Formation and Entropy of Ceria Oxide Using Topological Indices, *Combinatorial Chemistry & High Throughput Screening*, 25(3): 441–50. doi:10.2174/1386207323999201001210832