RESEARCH PAPER

Prediction of Physicochemical Properties of Antimicrobial Compounds Using Labeled Topological Indices and Regression Models

Vinaya K U[*1], Shrikanth A S[2], Mitha V U[3]

[1]*Department of Mathematics, Navkis College of Engineering (Affiliated to Visvesvaraya Technological University, Belagavi), Hassan - 573217, India*

[2&3]*Department of Mathematics, Adichunchanagiri Institute of Technology (Affiliated to Visvesvaraya Technological University, Belagavi), Chikkamagaluru - 577102, India*

ABSTRACT

Topological indices are a set of graph invariants that carry information about the structure of a molecule. In this study, we used a set of + and – signs for vertices and edges in chemical structures of 16 antimicrobial compounds in such a way that the number of + signs and – signs are balanced. Labeled topological indices were calculated for all compounds, and multiple regression models in different forms – linear, quadratic, cubic, exponential, and logarithmic – for six physicochemical properties – molecular weight (MW), polarizability, heavy atom count (HAC), boiling point (BP), molar refractivity (MR), and molar volume (MV) – are constructed. The results indicate that the labeled topological indices offer significant predictive ability. Comparison of results is performed by statistical parameters such as coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). The signed approach shows higher sensitivity for electronic effects in molecules for quantitative structure-property relationship predictions.

Keywords: Cordial labelling, Labeled Topological Indices, Physicochemical Properties, Regression Models.

How to cite this article: Vinaya K U, Shrikanth A S, Mitha V U.; Prediction of Physicochemical Properties of Antimicrobial Compounds Using Labeled Topological Indices and Regression Models. Int J Drug Deliv Technol. 2026;16(11s): 85-94; DOI: 10.25258/ijddt.16.11s.9

Source of support: Nil.

Conflict of interest: Nil

INTRODUCTION

The increasing rate of antimicrobial resistance has resulted in an immediate need for the discovery and optimization of novel antimicrobial compounds. The experimental determination of physicochemical properties and antimicrobial activity is time-consuming and requires significant resources, especially when dealing with a large number of compounds under investigation. Thus, computational methods, such as Quantitative Structure-Property Relationships (QSPR) and Quantitative Structure-Activity Relationships (QSAR), have become essential in contemporary drug discovery and development research (Cherkasov et al., 2014; Klein, 2002).

QSPR and QSAR are based on the fundamental hypothesis that physicochemical and biological properties of molecules are inherently linked to their chemical structure. By converting molecular structures into numerical values, QSPR and QSAR approaches facilitate the development of mathematical models for prediction without requiring large amounts of experimental data. Out of all the different kinds of molecular descriptors, topological indices from chemical graph theory have attracted special interest for their simplicity, mathematical soundness, and predictive ability (Balaban, 1985).

In chemical graph theory, a molecule is modeled by a graph in which vertices are associated with atoms and edges are associated with chemical bonds between these vertices. Topological indices are numerical invariants of a graph that are preserved under isomorphic transformations and carry structural information such as size, branching, connectivity, and cyclicity of the molecule (Trinajstic, 2018). Since the introduction of the Wiener index in 1947 for predicting the boiling points of paraffins, various topological indices have been developed and successfully used for predicting physicochemical properties such as molecular weight, density, molar volume, refractivity, and polarizability (Randic, 1975; Wiener, 1947).

Zagreb indices, which are based on vertex degree, were used to quantify branching in molecules, and this was successfully employed in QSPR models (Gutman and Trinajstić, 1972). The Randic index, a measure of molecular connectivity, was used to refine the relationship between molecular branching and physicochemical properties. Several authors reported excellent correlations between topological indices and molecular properties such as density, enthalpy of formation, refractivity, and surface area (Kier and Hall, 1976; Randić and Zupan, 2001). In addition, Devillers and Balaban presented a comprehensive review of the theoretical aspects and applications of topological descriptors in QSAR and QSPR (Dunn, 2000), while Todeschini and Consonni (Todeschini and Consonni, 2009) presented a comprehensive review of the classification of molecular descriptors used in cheminformatics.

Though these classical topological indices have achieved great success, they are generally based on unsigned molecular graphs and, in some way, assume that there is uniform interaction between atoms and bonds. This assumption may not be appropriate when dealing with chemically complex molecules, especially antimicrobial

**Author for Correspondence: Vinaya K U*

agents that contain many functional groups with significant electronic effects and regions with contrasting polarity (Karelson, 2000). To overcome these limitations, the theory of signed graphs proposes a promising extension of the existing graph theory. In the theory of signed graphs, signs are labeled to represent the edges and vertices of the graph to show the opposing or complementary interactions (Harary, 1953). In the chemical domain, these signs can be used to represent the donor-acceptor properties and the polarity of the molecules (Estrada, 2011).

Another significant type of signed graph is the set of balanced signed graphs, which were first defined by Harary. In this case, a signed graph is considered to be balanced if all cycles in the signed graph have an even number of negative edges, thus yielding a consistent global sign pattern (Harary, 2007). Balanced signed graphs are mathematically stable and chemically correspond to the concept of stable molecular configurations. The addition of balance constraints in the signed molecular representation is aimed at ensuring its physical correctness. Recently, researchers have made efforts to extend classical topological indices to labeled topological indices, which involve sign information in degree-based topological indices (Alameri and Al-Sharafi, 2021; Rilwan and Bharathi, 2025; Zaslavsky, 1983). These labeled topological indices have demonstrated greater sensitivity to structural details and electronic effects, making these topological indices promising tools for QSPR modeling. However, their use for antimicrobial agents and their evaluation for different physicochemical properties are limited.

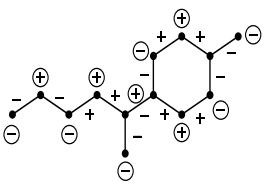
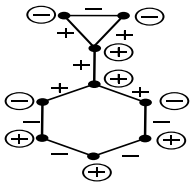
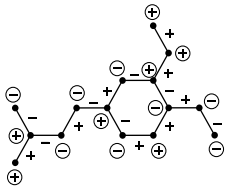
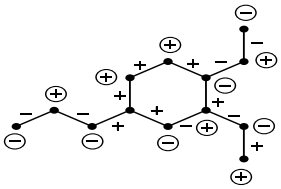
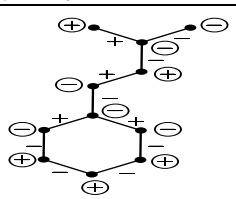
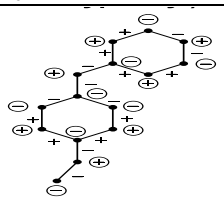
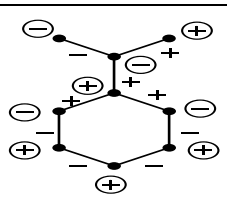
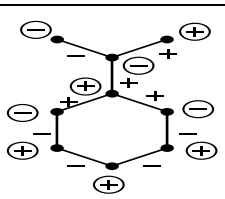
Labeled topological indices result from the extension of topological indices to their labeled versions. In signed versions of the Wiener index, sign-weighted distances are taken into account, whereas in signed versions of the Zagreb and Randić indices, sign-weighted degrees are used. These changes enable these topological indices to account for electronic effects and functional group effects, which are absent in their unsigned versions. Recently, it was found that these labeled topological indices exhibit greater discriminatory ability to differentiate between structural analogs and isomers of compounds. Such sensitivity is particularly useful in antimicrobial compounds, where structural changes may considerably influence their physicochemical properties and activity. Despite these advantages, the use of signed topological indices in QSPR modeling is relatively under investigated. Most of the existing literature on signed topological indices is devoted to their theoretical development, which reflects a research gap.

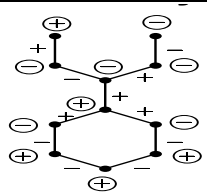
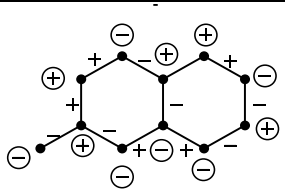
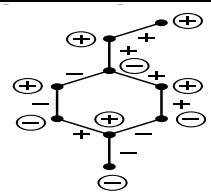
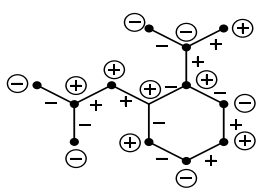
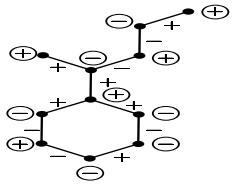
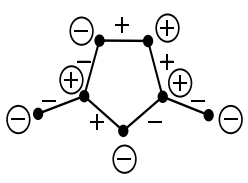
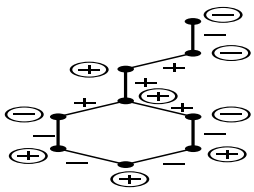
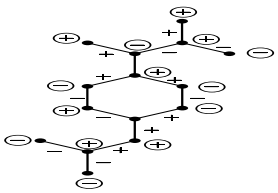
Integration of signed topological indices with regression analysis is one promising direction for the development of QSPR modeling. Linear regression is used as the baseline approach, while polynomial and nonlinear regression can be used to describe complex dependencies (Seber and Wild, 1989). Validation of the results with statistical parameters is critical, especially for small sets of data, to avoid overfitting (Montgomery et al., 2021). The present study contributes to this emerging area by systematically applying labeled topological indices and multiple regression models to predict six physicochemical properties of antimicrobial compounds, thereby extending the applicability of signed graph theory in chemical informatics.

METHODOLOGY

In the present study, the graph theory-based approach of QSPR is utilized to develop the model for the prediction of physicochemical properties of the selected antimicrobial compounds. In the first step of the study, the dataset of 16 different chemical compounds used for the purpose of antimicrobial activity was created. For the selected chemical compounds, six different physicochemical properties have been taken into consideration. These properties include molecular weight, polarizability, heavy atom count, boiling point, molar refractivity, and molar volume. These properties have been collected from the chemical databases PubChem and ChemSpider. Before the development of the model, the dataset was checked for the absence of missing values.

Table 1: Labelled Structures of Antimicrobial Compounds

			
Methyl 3-methoxy-4-methylbenzoate $C_{10}H_{12}O_3$	Styrene oxide C_8H_80	Dimethyl 2-methylterephthalate $C_{11}H_{12}O_4$	4-(4-Methoxyphenyl)-2-butanone $C_{11}H_{14}O_2$
			

Benzylacetone	(4-Phenoxyphenyl) methanol	Benzoic Acid	Acetophenone
$C_{10}H_{12}O$	$C_{13}H_{12}O_2$	$C_7H_6O_2$	C_8H_8O
			
Benzaldehyde dimethyl acetal	2-Naphthol	4-hydroxybenzaldehyde	Aspirin
$C_9H_{12}O_2$	$C_{10}H_8O$	$C_7H_6O_2$	$C_9H_8O_4$
			
Ethyl benzoate	Maleic Anhydride	Phenetole	Ibuprofen
$C_9H_{10}O_2$	$C_4H_2O_3$	$C_8H_{10}O$	$C_{13}H_{18}O_2$

Each antimicrobial compound was represented as a molecular graph in which atoms correspond to vertices and chemical bonds correspond to edges. To incorporate chemical heterogeneity and electronic effects, positive and negative signs were assigned to both vertices and edges of the molecular graphs. The sign assignment was performed as per the following definition by considering the molecular structure as a graph.

Let G be a graph with vertex set V and edge set E. Let g be a labeling from E to $\{+,-\}$. The edge labeling g induces a vertex labeling $h: V \to \{+,-\}$ defined by $h(v) = \prod g(uv)$ for u in $N(v)$, where $N(v)$ is the set of vertices adjacent to v. Let $e(+) = card\{e \in E: g(e) = +\}, e(-) = card\{e \in E: g(e) = -\}$ and $v(+) = card\{v \in V: h(v) = +\}$, $v(-) = card\{v \in V: h(v) = -\}$. A labeling g is said to be sign balanced if $|e(+) - e(-)| \leq 1$ and $|v(+) - v(-)| \leq 1$. This balanced sign configuration provides a mathematically stable representation and reflects chemically reasonable interaction patterns within the molecule.

Based on the sign balanced molecular graphs, a set of labeled topological indices were defined in table 2 by considering labeled incident of a vertex u, as $L_I(u) = \sum_{\forall v \in N(u)} f(uv)$ where N(u)is the neighborhood of u. The labeled indices of 16 compounds were derived by extending classical degree-based topological indices to the labeled setting, incorporating sign information associated with vertices and edges by considering 1 for + sign and 0 for – sign in Table-3.

Table 2: Labeled Topological Indices

Sl. No.	Topological Indices	Notation	Formulae
1	Labelled Square Index	$SQI(G)$	$\sum L_I(u)^2$
2	Labelled Product Index	$PI(G)$	$\sum L_I(u)L_I(v)$
3	Labelled Sum Index	$SI(G)$	$\sum L_I(u)+L_I(v)$
4	Labelled Nirmala Index	$NLI(G)$	$\sum \sqrt{L_I(u)+L_I(v)}$
5	Labelled Sombar Index	$SOLI(G)$	$\sum \sqrt{L_I(u)^2+L_I(v)^2}$
6	Labelled Forgotten Index	$FI(G)$	$\sum L_I(u)^2 + L_I(v)^2$

Table 3: Computed Values of the Labelled Topological Indices for Molecular Graphs

Sl. No.	Name	SQI	PI	SI	NLI	SOLI	FI
1	Methyl 3-methoxy-4-methylbenzoate	18	15	27	17.64	20.66	39
2	Styrene oxide	16	14	20	11.38	16.17	44
3	Dimethyl 2-methylterephthalate	22	21	35	22.56	25.91	49

4	4-(4-Methoxyphenyl)-2-butanone	26	25	33	18.63	25.07	65
5	Benzylacetone	18	17	26	15.90	19.25	40
6	(4-Phenoxyphenyl)methanol	28	24	36	21.76	28.26	64
7	Benzoic Acid	16	14	20	11.38	16.17	44
8	Acetophenone	16	14	20	11.38	16.17	44
9	Benzaldehyde dimethyl acetal	18	17	24	14.11	18.82	48
10	2-Naphthol	18	16	28	17.95	21.48	42
11	4-hydroxybenzaldehyde	16	14	22	13.71	16.96	36
12	Aspirin	18	14	28	18.63	22.25	42
13	Ethyl benzoate	20	17	27	16.63	21.41	51
14	Maleic Anhydride	8	7	14	9.71	10.72	18
15	Phenetole	16	14	19	10.97	15.17	41
16	Ibuprofen	28	26	39	23.41	30.53	73

In Table-4 physicochemical properties of antimicrobial compound are mentioned. The labeling scheme allowed differentiation between atomic contributions and bond characteristics, enabling the indices to capture subtle structural and electronic variations among the antimicrobial compounds. All index computations were performed using custom algorithms implemented in a computational environment.

Table 4: Physicochemical Properties of Antimicrobial Compounds.

Sl. No.	**Name**	MW	**α**	**HAC**	**BP**	**MR**	**V_m**
1	Methyl 3-methoxy-4methylbenzoate	180.2	19.6	13	257.4	49.5	167.6
2	Styrene oxide	120.2	14.0	09	194.0	35.3	108.5
3	Dimethyl 2-methylterephthalate	208.2	21.7	15	295.2	54.6	181.5
4	4-(4-Methoxyphenyl)-2-butanone	178.2	20.5	13	280.3	51.8	176.4
5	Benzylacetone	148.2	17.9	11	233.5	45.1	152.4
6	(4-Phenoxyphenyl) methanol	200.2	23.4	15	333.2	59.1	173.8
7	Benzoic Acid	122.1	13.2	09	249.3	33.2	102.0
8	Acetophenone	120.2	14.4	09	202.0	36.3	121.0
9	Benzaldehyde dimethyl acetal	152.2	17.4	11	199.7	43.9	152.2
10	2-Naphthol	144.2	18.2	11	285.5	46.0	122.0
11	4-hydroxybenzaldehyde	122.1	13.8	09	246.6	34.9	099.5
12	Aspirin	180.2	17.7	13	321.4	44.5	139.6
13	Ethyl benzoate	150.2	16.9	11	211.7	42.7	143.8
14	Maleic Anhydride	098.1	7.89	07	202.0	19.9	066.0
15	Phenetole	122.2	14.9	09	169.8	37.6	129.9
16	Ibuprofen	206.3	24.1	15	319.6	60.9	200.3

The correlation coefficients, which measure the linear correlations between different topological descriptors and the compounds physicochemical characteristics, are shown in Table 5. Bold highlights statistically significant links between these traits, emphasizing the largest positive correlations. Strong correlation values indicate the predictive capacity of these degree-based indices and show how useful they can be in physicochemical property prediction.

Table 5: The Correlation Between Physicochemical Properties and Labelled Topological Indices

Index	**MW**	**α**	**HAC**	**BP**	**MR**	**V_m**
SQI(G)	0.8513	0.9324	0.8734	0.6932	0.9333	0.8840
PI(G)	0.8190	0.9097	0.8379	0.6317	0.9107	0.8852
SI(G)	0.9410	**0.9629**	0.9628	0.8219	**0.9629**	**0.8999**
NLI(G)	**0.9590**	0.9277	**0.9665**	**0.8594**	0.9273	0.8533

SOLI(G)	0.9418	0.9612	0.9565	0.8215	0.9613	0.8927
FI(G)	0.7211	0.8316	0.7424	0.5443	0.8327	0.8102

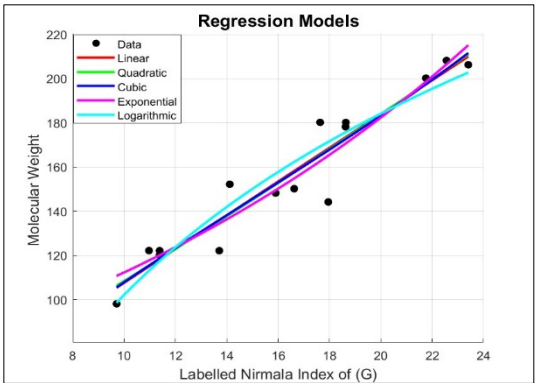
The revelation provides an avenue to understand the molecular structure and how it defines the fundamental physical properties. From the table above, it is clear that NLI(G) and SI(G) are the best, with NLI(G) showing good correlation with molecular weight, heavy atoms, and boiling point. However, SI(G) revealed its capacity in the prediction of molar volume, molar refractivity, and polarizability. To establish quantitative relationships between the labeled topological indices and the physicochemical properties, several regression models were developed based on the Pearson correlation coefficients.

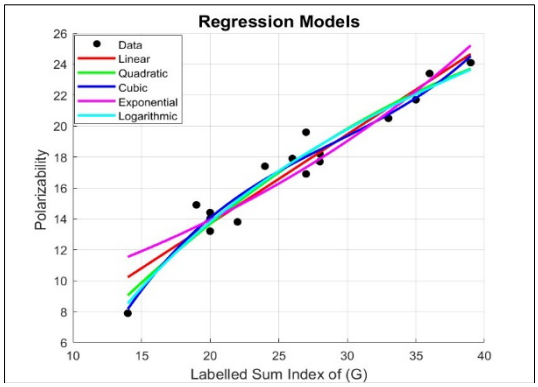
REGRESSION MODELLING AND STATISTICAL ANALYSIS

The QSPR analysis was conducted using a variety of regression models, including linear ($y = ax + b$), quadratic ($y = ax^2 + bx + c$), cubic ($y = ax^3 + bx^2 + cx + d$), exponential ($y = ae^{bx}$), and logarithmic ($y = a\,log\,x + b$) to further investigate the relationships suggested by the correlation coefficients between topological indices and physicochemical properties. Table 6 shows the coefficients, constants and significance of the data used. The physicochemical property is indicated as "y" while the topological index with the highest correlation coefficient is indicated as "x". Figure 1 shows a graph of each physicochemical property acting with its optimal topological index. All statistical computations and regression modeling were performed using MATLAB software packages. The modeling process provided a systematic evaluation of the ability of labeled topological indices to predict physicochemical properties of antimicrobial agents by ensuring reproducibility and consistency in the modeling results.

Table 6: Regression Models of the Physicochemical Properties with its Best Topological Index

Properties	**MW**	**α**	**HAC**	**BP**	**MR**	**V_m**
TI	**NLI(G)**	**SI(G)**	**NLI(G)**	**NLI(G)**	**SI(G)**	**SI(G)**
Linear Regression Model						
a	07.641	0.577	0.5537	10.052	1.459	04.622
b	31.167	2.139	2.3995	89.412	5.342	19.045
Quadratic Regression Model						
a	00.035	-0.010	0.001	00.130	-00.024	-00.067
b	06.490	1.099	0.527	05.782	02.763	08.224
c	39.975	-4.440	2.602	122.08	-11.124	-26.426
Cubic Regression Model						
a	0.007	0.001	0.001	-00.162	0.003	0.0111
b	-0.313	-0.107	-0.037	08.278	-0.272	-0.909
c	12.079	3.543	1.126	-125.05	8.975	29.378
d	11.396	-23.85	-0.461	791.02	-60.441	-194.39
Exponential Regression Model						
a	69.116	7.448	5.132	131.147	18.766	60.907
b	00.049	0.031	0.048	00.040	00.031	00.031
Logarithmic Regression Model						
a	118.392	14.771	8.593	154.657	37.314	117.943
b	-170.523	-30.459	-12.254	-172.939	-77.004	-240.975





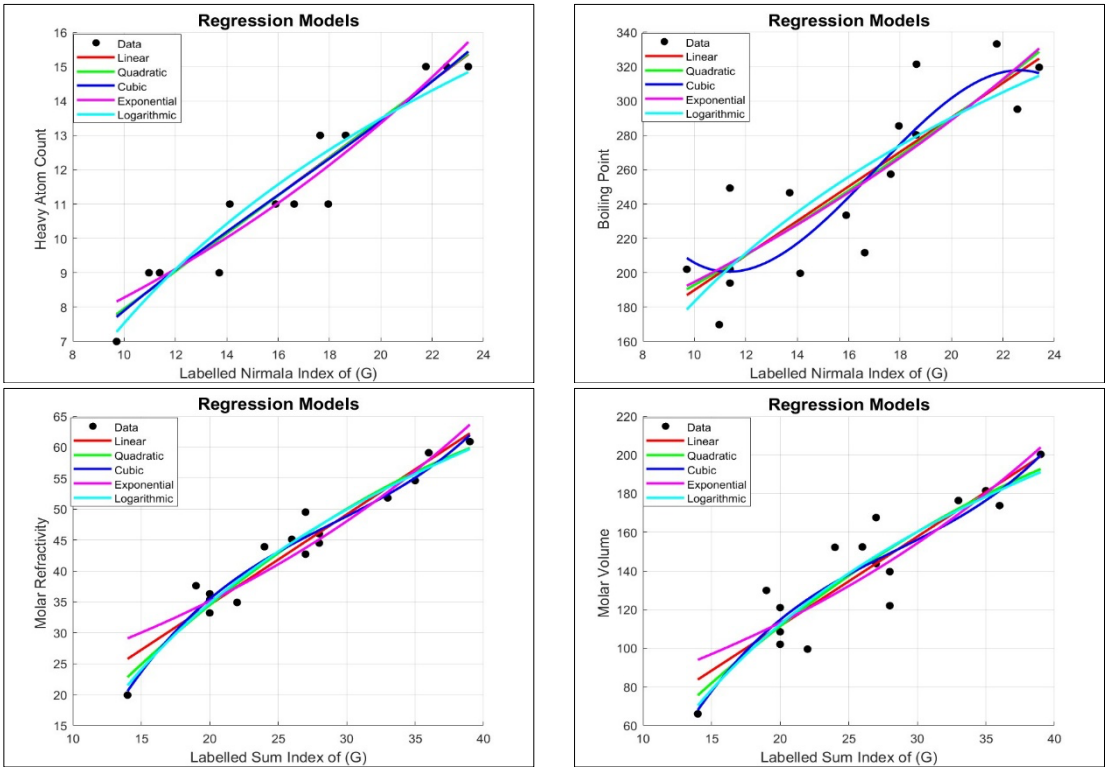


Figure 1: Graphically Representation of the Regression Models

The performance of all regression models was assessed using statistical performance measures such as the coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) presented in Table 7. Residual analysis was also performed to check for model adequacy, outliers and regression assumptions. Comparative analysis of regression forms was performed to establish which regression model was best for each physicochemical property of interest.

Table 7: Statistical Indicators

MAE	MAPE	RMSE	R^2
$\frac{1}{n}\sum_{i=1}^{n}\lvert p_i - a_i\rvert$	$\frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{p_i - a_i}{a_i}\right\rvert$	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\lvert p_i - a_i\rvert^2}$	$1 - \frac{\sum_{i=1}^{n}(p_i - a_i)^2}{\sum_{i=1}^{n}(p_i - \bar{a})^2}$

Here n is the number of compounds, a_i and p_i are the actual as well as predicted values of the physicochemical property, and $\bar{a}$ is the mean of the actual values. All the results obtained from the model's prediction were compared with the actual physicochemical properties.

RESULTS AND DISCUSSION

The calculated labeled topological indices showed considerable variation in the set of 16 antimicrobial compounds. This is necessary for effective QSPR modeling because the descriptors having less dispersion are found to be less effective in prediction. The balanced signs helped in the better differentiation of the structurally similar set of compounds by incorporating the information regarding the electron-donor and electron-withdrawer effects within the molecule. The statistical values of the minimum absolute percentage error (Min. Error %), maximum absolute percentage error (Max. Error %), RMSE, MAE, MAPE and R^2 values for every model corresponding to the physicochemical properties of the compounds are shown in Table 8.

Table 8: The Statistical Values of the Regression Models

Regression Model	Max Error %	Min Error %	MAE	RMSE	MAPE	R^2
Molecular Weight						
Linear	16.7679	1.4088	7.6853	9.6009	5.2942	0.9053
Quadratic	16.3905	1.2195	7.6573	9.5838	5.2692	0.9065
Cubic	**16.0432**	**1.3638**	**7.7278**	**9.5802**	**5.2644**	**0.9071**

Exponential	14.5105	0.0920	7.5450	9.8978	5.3375	0.9060
Logarithmic	18.8622	0.5107	8.2367	10.4916	5.5574	0.8865
Polarizability						
Linear	29.5678	0.5852	0.9135	1.0929	6.3392	0.9119
Quadratic	14.5830	1.8440	0.8759	0.9840	5.6467	0.9280
Cubic	**12.2946**	**0.5768**	**0.9871**	**1.1990**	**5.5304**	**0.9376**
Exponential	46.3147	0.5059	0.9889	1.3312	7.2508	0.8788
Logarithmic	12.5394	1.3156	0.8247	0.9435	5.5359	0.9340
Heavy Atom Count						
Linear	12.1815	0.7280	0.5343	0.6251	5.1352	0.9240
Quadratic	**12.0541**	**0.6368**	**0.5340**	**0.6250**	**5.1009**	**0.9244**
Cubic	10.2354	0.3066	0.6010	0.6802	5.5295	0.9242
Exponential	16.5871	0.2336	0.5384	0.6574	5.1097	0.9236
Logarithmic	14.1925	0.9106	0.5625	0.6836	5.2912	0.9069
Boiling Point						
Linear	21.1712	0.9019	21.1250	25.4369	8.8348	0.7004
Quadratic	20.0482	1.3559	21.2796	25.3484	8.8382	0.7048
Cubic	**19.8741**	**0.6708**	**19.0229**	**23.9621**	**7.9424**	**0.7337**
Exponential	19.4630	1.7810	21.4338	25.4191	8.8763	0.7050
Logarithmic	23.6599	0.3366	21.7050	26.5134	9.1779	0.6743
Molar Refractivity						
Linear	29.4149	0.4154	2.3083	2.7606	6.3393	0.9116
Quadratic	14.6203	1.6282	2.2201	2.4901	5.6677	0.9276
Cubic	**11.3676**	**0.08717**	**1.8040**	**2.3409**	**4.3272**	**0.9406**
Exponential	46.0851	0.5825	2.5032	3.3600	7.2611	0.8784
Logarithmic	12.6875	1.2566	2.0928	2.3903	5.1416	0.9337
Molar Volume						
Linear	26.8872	0.0175	12.2936	15.1189	10.1045	0.7616
Quadratic	24.1554	1.0593	12.2688	14.7562	9.6919	0.7778
Cubic	**25.4581**	**0.6852**	**11.7261**	**14.3336**	**8.9762**	**0.7868**
Exponential	42.4309	0.6890	13.4784	16.1685	11.3563	0.7301
Logarithmic	24.6190	1.7336	12.1087	14.6193	9.3461	0.7842

The higher value of R^2 implies a better fit and helps in the selection of the model. The cubic regression model has the highest value of R^2, which results in the minimum value of MAE and RMSE and MAPE, which implies less prediction error, which indicates a good fit. The result indicates that the quadratic and cubic models accurately describe the relationship between the input and output variables in each case at a reasonable level of agreement. The graphical representation of the predicted values and the actual values obtained from the best-fit model is shown in Figure 2. The ability of the model in generalizing the relationship between the physicochemical properties and the labeled topological indices is shown in the best-fit curve.

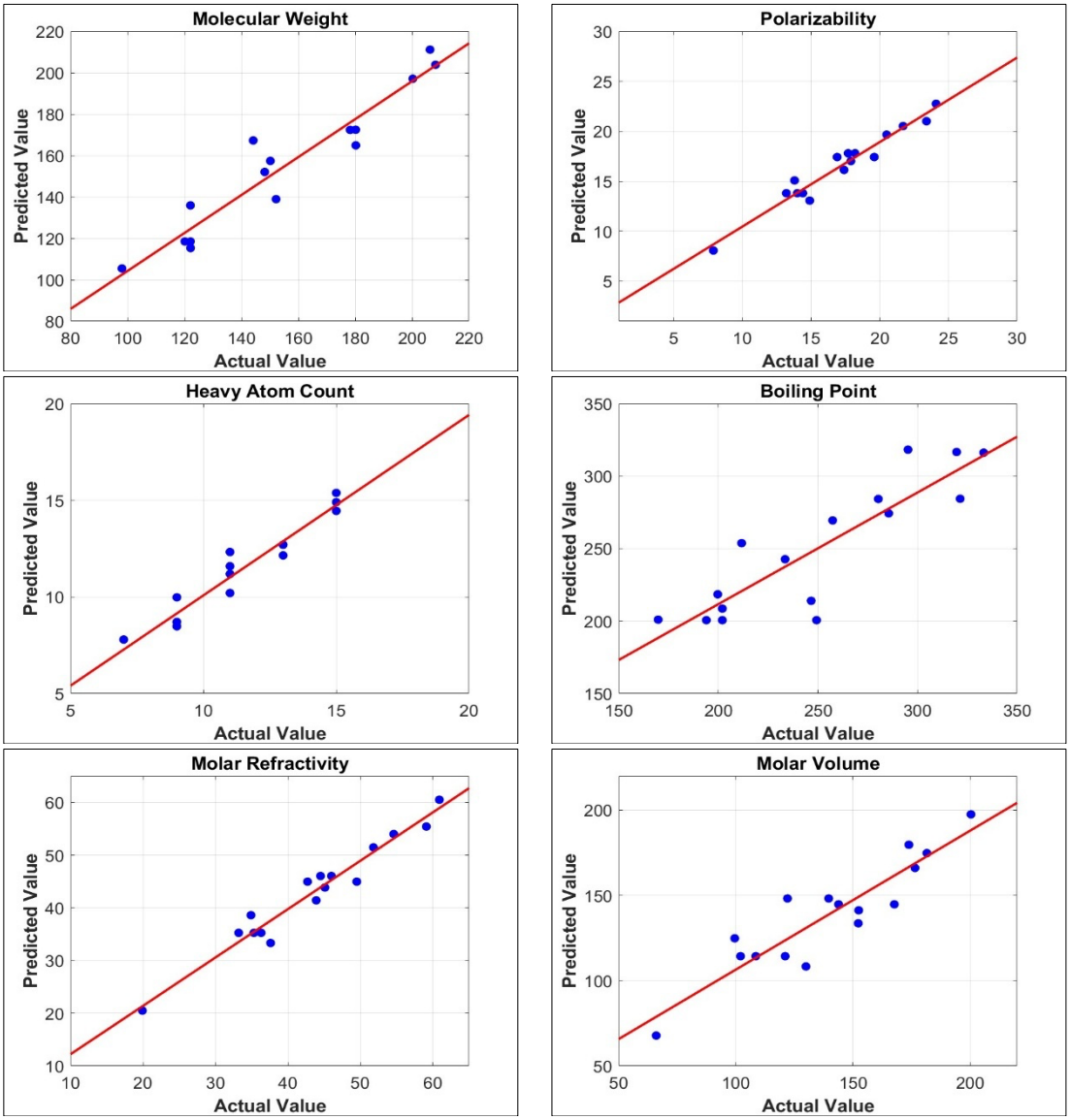
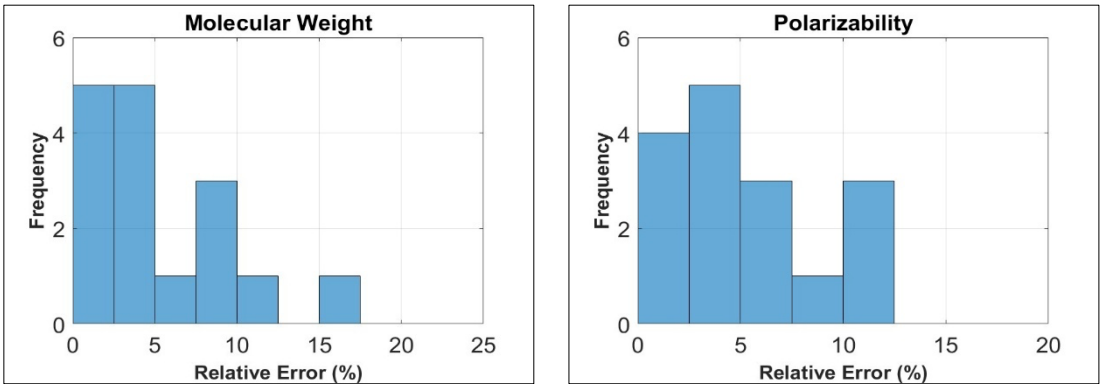


Figure 2: Comparing the Values of the Selected Properties Actual and Predicted

The results indicate that a significant relationship exists between the observed values and the predicted values, which proves that these cubic and quadratic regression models perform well in recognizing the underlying patterns. In this regard, HAC revealed that the quadratic model performed best, while all the other properties revealed that the cubic model performed best. In addition, the relative error percentage is determined, and its histogram is displayed in Figure 3 for evaluating the precision of the prediction made by the best-fit regression model in regard to the differences in values for the predicted and actual physicochemical properties. The results indicate how well topological indices can predict physicochemical properties while revealing low relative errors in all characteristics examined.



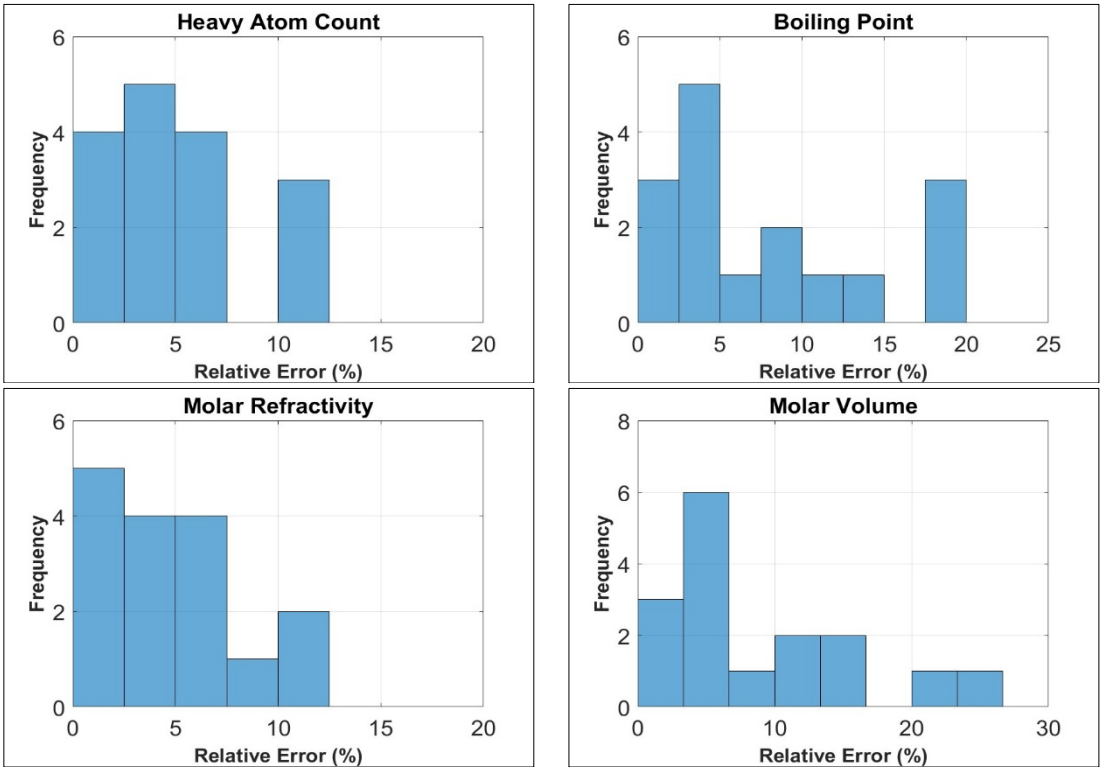


Figure 3: Histogram of Relative Error Percentage

Some minor deviations in the results for certain compounds can be explained by unique structural characteristics not fully captured by the chosen topological descriptors, which is a limitation in this approach in general. Significantly, the inclusion of balanced signed information resulted in a substantial improvement in the interpretability of the results. In addition, the labeled indices used in the study showed improved sensitivity and prediction accuracy compared to the unlabeled topological indices reported in the literature.

However, it should be noted that despite these promising results, it is also important to acknowledge that the relatively limited dataset used here does limit the overall generalization of these models. Even though these results show statistically significant correlations, it is also important to validate these results on larger datasets to confirm the overall effectiveness of this proposed approach. However, it should also be noted that the overall trends are consistent across these regression models, which also indicate that the labeled topological indices are useful for the prediction of physicochemical properties.

CONCLUSION

In the current research, a well-rounded, graph-based QSPR model has been designed for the prediction of the essential physicochemical properties of antimicrobial agents. In the current model, the molecular structure has been successfully represented with the help of a sign labeled graph, and the labeled topological index has been calculated for the successful incorporation of the structural and electronic properties of the molecule. Regression models of different complexities, like linear, quadratic, cubic, exponential and logarithmic models, have been employed for the purpose of building quantitative relationships.

This research work aims at explaining the benefits associated with the use of labeled degree-based topological indices in the QSPR model in the prediction of physicochemical properties for antibacterial drugs. Degree-based topological indices are advantageous in the estimation of the risk to the environment, whereas the QSPR study was able to show the correlation between topological indices and the physicochemical properties of antibacterial compounds through the calculation of the correlation coefficients. The topological index with the highest correlation coefficient to molecular weight is NLI (0.959029); SI (0.962902) for polarizability; NLI (0.96654) for heavy atom count; NLI (0.859393) for boiling point; SI (0.962918) for molar refractivity; and SI (0.899971) for molar volume.

These results show that labeled topological indices have a strong predictive ability and offer better sensitivity than the conventional unlabeled indices, especially for properties that depend on the distribution of electrons, such as polarizability and molar refractivity. Comparison of regression models showed that nonlinear models are better than linear models in most cases. Overall, this work establishes labeled topological indices as promising descriptors for physicochemical property prediction and provides a foundation for future studies involving larger datasets, multivariate modeling and biological activity prediction of antimicrobial agents

REFERENCE

1. Alameri, A.Q.S., Al-Sharafi, M.S.Y., 2021. Topological Indices of Some New Graph Operations and

Their Possible Applications. Asian Journal of Probability and Statistics 16–30. https://doi.org/10.9734/ajpas/2021/v13i230302

2. Balaban, A.T., 1985. Applications of graph theory in chemistry. J. Chem. Inf. Comput. Sci. 25, 334–343. https://doi.org/10.1021/ci00047a033

3. Cherkasov, A., Muratov, E.N., Fourches, D., Varnek, A., Baskin, I.I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y.C., Todeschini, R., Consonni, V., Kuz'min, V.E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A., Tropsha, A., 2014. QSAR Modeling: Where Have You Been? Where Are You Going To? J. Med. Chem. 57, 4977–5010. https://doi.org/10.1021/jm4004285

4. Dunn, W.J., 2000. Topological Indices and Related Descriptors in QSAR and QSPR Edited by James Devillers and Alexandru T. Balaban. Gordon and Breach Science Publishers, Amsterdam, Netherlands. 1999. x + 811 pp. 15.5 × 23.5 cm. ISBN 90-5699-239-2. $198.00. J. Med. Chem. 43, 5055–5056. https://doi.org/10.1021/jm000448a

5. Estrada, E., 2011. The Structure of Complex Networks: Theory and Applications. Oxford University Press, New York.

6. Gutman, I., Trinajstić, N., 1972. Graph theory and molecular orbitals. Total φ-electron energy of alternant hydrocarbons. Chem. Phys. Lett. 17, 535–538. https://doi.org/10.1016/0009-2614(72)85099-1

7. Harary, F., 2007. On the measurement of structural balance. Behavioral Science 4, 316–323. https://doi.org/10.1002/bs.3830040405

8. Harary, F., 1953. On the notion of balance of a signed graph. Michigan Mathematical Journal 2. https://doi.org/10.1307/mmj/1028989917

9. Karelson, Mati., 2000. Molecular descriptors in QSAR/QSPR. Wiley-Interscience. .

10. Kier, L., Hall, H., 1976. Molecular Connectivity in Chemistry and Drug Research. Academic Press, London.

11. Klein, D.J., 2002. Topological Indices and Related Descriptors in QSAR and QSPR Edited by James Devillers & Alexandru T. Balaban. Gordon and Breach Science Publishers: Singapore. 1999. 811 pp. 90-5699-239-2. $198.00. J. Chem. Inf. Comput. Sci. 42, 1507–1507. https://doi.org/10.1021/ci010441h

12. Montgomery, D., Peck, E., Vining, G., 2021. Introduction to Linear Regression Analysis. John Wiley & Sons Inc., New Jersey.

13. Randic, M., 1975. Characterization of molecular branching. J. Am. Chem. Soc. 97, 6609–6615. https://doi.org/10.1021/ja00856a001

14. Randić, M., Zupan, J., 2001. On Interpretation of Well-Known Topological Indices. J. Chem. Inf. Comput. Sci. 41, 550–560. https://doi.org/10.1021/ci000095o

15. Rilwan, N.M., Bharathi, R.D., 2025. Exploring the Degree Based Topological Indices through M-Polynomial and Entropy Measures of Porous Graphitic Framework via Logarithmic Regression Analysis. https://doi.org/10.21203/rs.3.rs-6919353/v1

16. Seber, G.A.F., Wild, C.J., 1989. Nonlinear Regression. Wiley. https://doi.org/10.1002/0471725315

17. Todeschini, R., Consonni, V., 2009. Molecular Descriptors for Chemoinformatics. Wiley. https://doi.org/10.1002/9783527628766

18. Trinajstic, N., 2018. Chemical Graph Theory. Routledge. https://doi.org/10.1201/9781315139111

19. Wiener, H., 1947. Structural Determination of Paraffin Boiling Points. J. Am. Chem. Soc. 69, 17–20. https://doi.org/10.1021/ja01193a005

20. Zaslavsky, T., 1983. Signed graphs. Discrete Appl. Math. (1979). 5, 248. https://doi.org/10.1016/0166-218X(83)90047-1.