DOI: https://doi.org/10.71317/kjard.2.11.2026.954

Kashmir Journal of Academic Research and Development

Journal homepage: https://rjsaonline.org/index.php/KJARD

Degree-Based Molecular Descriptors as Predictors of Physicochemical Properties of Antibiotic Compounds

Muhammad Hassnain
Department of Mathematics Division of Science and Technology, University of Education, Township Lahore
Email: mh9015610@gmail.com
Maria Iqbal
University of Narowal
Email: mariaiqbal0348@gmail.com
Mehr Un Nisa
Farooq Gul Khan College, Layyah, affiliated with Bahauddin Zakariya University, Multan, Pakistan
Email: mehru2971@gmail.com
Ayesha Bibi
Riphah International University, Islamabad
Email: Ayesha.bibi@riphah.edu.pk
Iqra Shaikh
Federal Urdu University of Arts, Science and Technology
Email: Iqrashaikh00@icloud.com
Sana Khurshid
University of Gujrat, Subcampus Rawalpindi
Email: sanakhurshid19977@gmail.com
Syed Hashir Ibraheem
Federal Urdu University of Arts Science and technology.
Email: syyed.hashir@gmail.com

ARTICLE INFO

Received:
September 21, 2026
Revised:
October 06, 2026
Accepted:
October 22, 2026
Available Online:
November 02, 2026

Keywords:
degree-based molecular descriptors; antibiotic drugs; QSPR analysis; physicochemical properties.

Corresponding Author:
Mehr Un Nisa
Email: mehru2971@gmail.com

ABSTRACT

Predicting the physicochemical properties of pharmaceutical compounds through mathematical frameworks o ers a powerful and cost-e ective alternative to laboratory experimentation. In this study, ten degree-based molecular descriptors are applied to eight clinically important antibiotic drugs within a quantitative structure property relationship (QSPR) framework. Descriptor values are computed via the edge partitioning technique, and simple linear regression is employed to model seven key physicochemical properties: boiling point, vapour pressure, ash point, enthalpy of vaporisation, surface tension, molar volume, and molar refractivity. The resulting correlation coe cients reveal strong positive associations between the majority of descriptors and the target properties, with molar refractivity and boiling point exhibiting the highest predictive accuracy. The statistically estimated values show excellent agreement with experimentally reported data, con rming the validity and reliability of the proposed approach. These ndings establish degree-based molecular descriptors as e cient, economical, and computationally accessible tools for characterising the physicochemical behaviour of antibiotic compounds, with promising implications for rational drug design and pharmaceutical property prediction.

Introduction

Antibiotics have saved more lives over the past century than perhaps any other class of drugs. Since Fleming's accidental discovery of penicillin in 1928 [17], these compounds have become central to how we treat bacterial disease from a simple skin infection to post-surgical complications that would otherwise be fatal [16]. The diseases they target are wideranging: respiratory tract infections, urinary tract infections, skin conditions like cellulitis and impetigo, and gut infections caused by harmful bacteria are all routinely managed with antibiotic therapy [15].

That said, the landscape is changing. Resistance is spreading faster than new drugs are being developed, and the pipeline for novel antibiotics has slowed considerably [18]. One consequence of this is growing pressure to do more with existing compounds to understand them better, to predict their behaviour more e ciently, and to reduce the cost and time involved in characterising them. Laboratory measurement of physicochemical properties such as boiling point, molar refractivity, or surface tension remains essential, but it is slow and expensive, particularly in early screening stages.

This is where mathematical modelling becomes useful. A molecule can be represented as a graph, with atoms as vertices and bonds as edges [19]. From this graph, numerical invariants called topological indices can be computed. These indices encode structural information about the molecule in a single number, and that number turns out to be surprisingly informative about physical and chemical properties. This idea underpins quantitative structure property relationship (QSPR) and quantitative structure activity relationship (QSAR) modelling, which have been used for decades to estimate molecular behaviour without the need for direct measurement [2,6 8].

Among the many topological indices proposed in the literature, degree-based descriptors are the most widely used. They are simple to compute, require nothing more than the two-dimensional molecular graph, and tend to perform well across a range of compound types. The idea goes back to Wiener, who in 1947 showed that a path-based graph invariant could predict boiling points of para ns with remarkable accuracy [10]. Building on that, Randi¢ introduced what he called the branching index [11], and from there the eld expanded rapidly.

In the present study, we take this approach and apply it to a set of eight clinically important antibiotic drugs. We compute ten degree-based topological indices for each compound and use simple linear regression to model seven physicochemical properties. The speci c contributions of this paper are as follows.

1. We compute ten degree-based topological indices for eight antibiotic drug structures using the edge partitioning method. This class of descriptors was chosen because prior work has shown it to be well-suited for pharmaceutical compounds [2,7,26].
2. We build simple linear regression models to relate each index to each target property, producing seventy regression equations in total [3,29].
3. We compare our statistically estimated property values against experimental data from the literature, using Pearson correlation coe cients and *p*values to assess the quality of each model [4,5,9].
4. We generate correlation coe cient plots in Python using Jupyter Notebook and Matplotlib to visualise how descriptor performance varies across the seven target properties [6,27].

Literature Review and Index De nitions

Before presenting the results, we brie y introduce the graph-theoretic framework and de ne the ten indices used in this study. Let $G = (V,E)$ be a simple connected graph, where $V(G)$ is the set of vertices and $E(G)$ is the set of edges. We work exclusively with simple two-dimensional molecular graphs, meaning no multiple edges or self-loops are present. The degree of a vertex x, written η_x, is the number of edges incident to it. In all molecular graphs considered here, vertex degrees lie between one and four.

1. Randi¢ Index [2,9,11]:

A topological measurement called the branching index was developed by Milan Randi¢ to assess the degree of branching in the carbon-atom skeleton of organic compounds.

$$R(G) = \sum_{xy \in E(G)} \frac{1}{\sqrt{\eta_x \cdot \eta_y}}$$

For enthalpy, molar refraction, and surface tension this index has strong correlation coe cients, showing it is an e ective predictor of these properties [29].

2. Reciprocal Randi¢ Index [7]:

Gutman proposed the reciprocal Randi¢ index:

$$RR(G) = \sum_{xy \in E(G)} \sqrt{\eta_x \cdot \eta_y}$$

3. Harmonic Index [8,20]: Fajtlowicz rst suggested the harmonic index in 1987. The H-index is de ned as:

$$H(G) = \sum_{xy \in E(G)} \frac{2}{\eta_x + \eta_y}$$

The H-index can be used to obtain extremal values of chemical trees in terms of the order of the molecular graph [14].

4. Sum-Connectivity Index [2,21]: Zhou and Trinajsti¢ introduced the sumconnectivity index in 2009:

$$\text{SCI}(G) = \sum_{xy \in E(G)} \frac{1}{\sqrt{\eta_x + \eta_y}}$$

It is extremely important in calculating the correlation coe cients for boiling points of antibiotic medicines.

5. First and Second Zagreb Index [12]: Gutman introduced the rst and second Zagreb indices, nding a relationship between total πelectron energy and these indices:

$$M_1(G) = \sum_{xy \in E(G)} (\eta_x + \eta_y)$$

$$M_2(G) = \sum_{xy \in E(G)} (\eta_x \cdot \eta_y)$$

Zagreb indices are important for understanding the properties of corona graphs and connected quasi-unicyclic graphs [25].

6. Hyper Zagreb Index [6,22]:

Shirdel proposed the hyper Zagreb index:

$$HM(G) = \sum_{xy \in E(G)} (\eta_x + \eta_y)^2$$

The hyper Zagreb index can be used to nd the Hamiltonicity of simple graphs and is useful in network theory.

7. Reduced Second Zagreb Index [8,23]: In 2013, Ranjini introduced the reduced Zagreb indices:

$$RM_2(G) = \sum_{xy \in E(G)} (\eta_x - 1)(\eta_y - 1)$$

8. Forgotten Index [13]:

Furtula and Gutman introduced the forgotten index in 2015 to calculate the properties of carbon compounds. They observed that the *F*index correlates with many characteristics of octane isomers:

$$F(G) = \sum_{xy \in E(G)} (\eta x2 + \eta y2)$$

This index has a correlation with physical properties of anti-tuberculosis drugs [14].

9. Hyper Forgotten Index [7,24]:

The hyper forgotten index, a vertex degree-based index, was introduced inspired by the forgotten index:

$$HF(G) = \mathrm{X}\ (\eta x2 \cdot \eta y2)2$$

$$xy \in E(G)$$

The characteristics of nanostar dendrimers can be explained using the hyper forgotten index.

Methodology

There are three common ways to compute degreebased topological indices from a molecular graph: edge partitioning, vertex partitioning, and direct degree enumeration [2,7]. We used the edge partitioning method throughout this work. The idea is straightforward edges are grouped into classes based on the degree pair (η_x,η_y) of their endpoints, and the index formula is then evaluated as a sum over these classes rather than over individual edges. For the molecules in this study, this signi cantly reduces the computational work involved.

The molecular structures of all eight antibiotic compounds were rst drawn using ChemSketch. From each structure, vertex degrees were read o and the edge partitions were identi ed. All index values were then computed by hand and double-checked with a scienti c calculator. Regression analysis was carried out in MS Excel [3,28], and the correlation coe cient plots shown later were produced in Python using Jupyter Notebook and the Matplotlib library.

Computation of Topological Indices

The eight antibiotic compounds studied here are Amoxicillin, Penicillin, Cipro oxacin, Azithromycin, Doxycycline, Clindamycin, Tetracycline, and Sulfamethazole. Their molecular structures are shown in Figures 1 8. These are three-dimensional molecules, but for the purposes of topological index computation.

Table 1: Physicochemical properties of antibiotic medicines [1,3]

Drug Name	BP	VP	EV	FP	MR	ST	MV
Amoxicillin	743.5±60	0.0±2.6	113.7	403.3±32.9	91.5	85.3	236.2
Penicillin	663.3±55	0.0±2.1	102.5	355.0±31.5	86.3	67.9	235.2
Cipro oxacin	581.8±50	0.0±1.7	91.5	305.6±30.1	83.3	67.4	226
Azithromycin	822.1±65	0.0±0.6	136	451.0±34.3	197.6	50.6	632.7
Doxycycline	762.6±6	0.0±32.7	166.5	415±32.9	109	99.2	271.1
Clindamycin	628.1±55	0.0±4.2	106.5	333.6±31.5	107.9	56.2	327.5
Tetracycline	738.2±70	0.0±2.6	113	400.2±35.7	106.9	73	266.3
Sulfamethazole	482.1±55	0.0±1.2	74.7	245.4±31.5	62.5	70.9	173.1

Table 2: Topological indices of antibiotic medicines

Drug Name	$R(G)$	$RR(G)$	$H(G)$	SCI(G)	$M_1(G)$	$M_2(G)$	$HM(G)$	$RM_2(G)$	$F(G)$	$HF(G)$
Amoxicillin	11.6303	65.83	10.88	12.077	138	168	726	57	90	6106
Penicillin	10.809	60.36	10.21	11.26	126	160	654	53	350	5386
Cipro oxacin	11.55	65.42	11.1667	12.23	134	164	682	57	354	5002
Azithromycin	23.514	109.37	34.8090	20.65	268	320	1406	105	681	12298
Doxycycline	13.866	82.194	13.89	15.50	184	237	1006	88	538	8914
Clindamycin	12.70	65.447	12	12.89	136	161	684	53	362	5096
Tetracycline	14.77	8.70	13.72	15.41	186	239	1028	88	550	9584
Sulfamethazole	7.97	41.8020	7.51	8.20	88	100	440	29	240	3516

We work with their two-dimensional graph representations, which is standard practice in QSPR modelling [2,7,26].

The experimentally reported physicochemical property values for these compounds are listed in Table 1 [1,5]. For each compound, we identi ed all edge types from the molecular graph using the notation

$[e = \{xy \in E(G) : \eta_x = a, \eta_y = b\}]$, where a and b are the endpoint degrees. The computed values of all ten indices are presented in Table 2 [2, 7, 8, 14]. These serve as the independent variables in the regression models of the next section. In general, the estimated property values agreed well with the experimental benchmarks [9,27], which gives us con dence in the approach.

Regression Models

To explore how well each topological index predicts the physicochemical properties of the eight antibiotic compounds, we tted a simple linear regression model for each descriptor property combination [3, 4, 29]. The model takes the standard form:

$$Y = mX + c \qquad (1)$$

where X is the topological index value, Y is the physicochemical property being estimated, m is the slope, and c is the intercept [3]. In total, seventy regression equations were derived ten descriptors across seven target properties. All models were tted and visualised in Python using Jupyter Notebook with the NumPy and Matplotlib libraries. The equations for each descriptor are listed below.

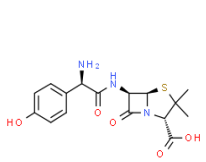


Figure 1: Amoxicillin

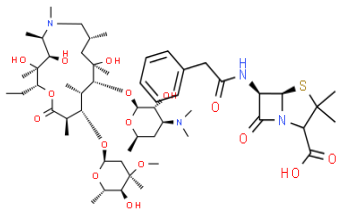


Figure 2: Penicillin

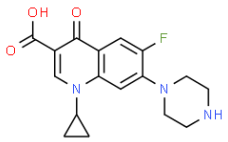


Figure 3: Cipro oxacin

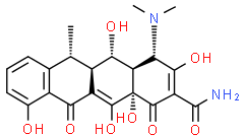


Figure 4: Azithromycin

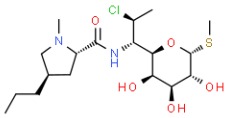


Figure 5: Doxycycline

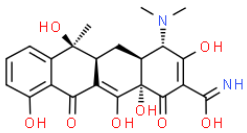


Figure 6: Clindamycin

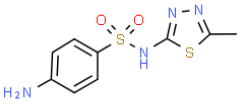


Figure 7: Tetracycline

Figure 8: Sulfamethazole

1. Randi¢ Index $R(G)$

$$BP = 426.4443 + 18.8199R(G)$$

$$VP = 5.5505 + 0.0309R(G)$$

$$EV = 64.5846 + 3.63R(G)$$

$$FP = 211.6642 + 11.3809R(G)$$

$$MR = -10.3754 + 8.6884R(G)$$

$$ST = 88.5008 - 1.2874R(G)$$

$$MV = -101.1703 + 29.7489R(G)$$

2. Reciprocal Randi¢ Index $RR(G)$

$$BP = 328.8706 + 4.8864RR(G)$$

$$VP = -1.6329 + 0.1064RR(G)$$

$$EV = 38.5694 + 1.0433RR(G)$$

$$FP = 152.6915 + 2.9545RR(G)$$

$$MR = -30.3738 + 1.905RR(G)\ ST = 81.3104 - 0.14RR(G)$$

$$MV = -149.0613 + 6.2344RR(G)$$

3. Harmonic Index $H(G)$

$$BP = 552.8484 + 8.7481H(G)$$

$$V\,P = 7.1347 - 0.0821H(G)$$

$$EV = 89.7797 + 1.6303H(G)$$

$$FP = 288.0758 + 5.2922H(G)$$

$$MR = 39.4203 + 4.6384H(G)$$

$$ST = 83.085 - 0.8248H(G)$$

$$MV = 62.183 + 16.3824H(G)$$

4. Sum Connectivity Index SCI(G)

$$BP = 332.7473 + 25.5017SCI(G)$$

$$V\,P = -1.6385 + 0.5619SCI(G)$$

$$EV = 38.528 + 5.5091SCI(G)$$

$$FP = 155.0358 + 15.4191SCI(G)$$

$$MR = -33.4412 + 10.2805SCI(G)\ ST = 83.3019 - 0.8863SCI(G)$$

$$MV = -161.5426 + 33.825SCI(G)$$

5. First Zagreb Index $M_1(G)$

$$BP = 406.8515 + 1.7198M_1(G)$$

$$V\,P = 0.8751 + 0.0323M_1(G)$$

$$EV = 56.0185 + 0.3621M_1(G)$$

$$FP = 199.8338 + 1.0399M_1(G)$$

$$MR = -4.525 + 0.6994M_1(G)$$

$$ST = 80.498 - 0.0583M_1(G)$$

$$MV = -67.1629 + 2.3059M_1(G)$$

6. Second Zagreb Index $M_2(G)$

$$BP = 401.389 + 1.4271M_2(G)$$

$$V\,P = -1.0592 + 0.0363M_2(G)$$

$$EV = 53.3171 + 0.3085M_2(G)$$

$$FP = 196.5418 + 0.8629M_2(G)$$

$$MR = 1.2772 + 0.5389M_2(G)$$

$$ST = 76.8541 - 0.0286M_2(G)$$

$$MV = -40.0938 + 1.7359M_2(G)$$

7. Hyper Zagreb Index *HM*(*G*)

$$BP = 413.953 + 0.3185HM(G)$$

$$VP = -0.1631 + 0.0074HM(G)$$

$$EV = 56.6604 + 0.0681HM(G)$$

$$FP = 204.1392 + 0.1925HM(G)$$

$$MR = 3.9024 + 0.1228HM(G)$$

$$ST = 77.269 - 0.0072HM(G)\ MV = -33.565 + 0.3979HM(G)$$

8. Reduced Second Zagreb Index

$$RM_2(G)$$

$$BP = 416.6582 + 3.9404RM_2(G)$$

$$VP = -3.5105 + 0.143RM_2(G)$$

$$EV = 53.7678 + 0.8948RM_2(G)$$

$$FP = 205.7972 + 2.3821RM_2(G)$$

$$MR = 16.9169 + 1.339RM_2(G)$$

$$ST = 71.1867 + 0.0019RM_2(G)$$

$$MV = 21.9439 + 4.1369RM_2(G)$$

9. Forgotten Index *F*(*G*)

$$BP = 378.8184 + 0.6901F(G)$$

$$VP = -2.8829 + 0.0204F(G)$$

$$EV = 47.5025 + 0.1513F(G)$$

$$FP = 182.9141 + 0.4172F(G)$$

$$MR = -1.3832 + 0.2471F(G)$$

$$ST = 73.9314 - 0.006F(G)$$

$$MV = -42.1976 + 0.7809F(G)$$

10. Hyper Forgotten Index *HF*(*G*)

$$BP = 448.9259 + 0.0327HF(G)$$

$$VP = 0.0855 + 0.0008HF(G)$$

$$EV = 64.3202 + 0.007HF(G)$$

$$FP = 225.2955 + 0.0198HF(G)$$

$$MR = 22.6159 + 0.0119HF(G)$$

$$ST = 73.6921 - 0.0003HF(G)\ MV = 32.8025 + 0.0377HF(G)$$

Statistical Analysis

For each regression model we report six parameters [3, 4]: the sample size $N = 8$, the intercept c, the slope m, the Pearson correlation coe cient r, the coe cient of determination r^2, the Fisher statistic F, and the two-tailed signi cance value P.

The Pearson coe cient r is the most important of these for our purposes [9]. A value near +1 means the index and the property rise together; a value near −1 means they move in opposite directions; a value near zero means there is little or no linear relationship between them. We treat $|r| \geq 0.90$ as the threshold for a strong predictor. The P-value tells us whether the observed correlation could plausibly have arisen by chance values below 0.05 are taken as statistically signi cant [2,6].

To give a concrete example, Table 3 shows the results for the Randi¢ index. The strongest correlation is with molar refractivity ($r = 0.988$, $P < 0.001$), and the weakest is with vapour pressure ($r = 0.013$). This already tells us something useful: the Randi¢ index is an excellent predictor of molar refractivity for these antibiotic compounds, but it carries essentially no information about vapour pressure [2, 26]. The full results for all remaining descriptors are given in Tables 4 12.

Table 3: Statistical data for Randi¢ Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	426.4443	18.8199	0.7789	0.6067	9.2560	0.0227
VP	8	5.5505	0.0309	0.0131	0.0002	0.0010	0.9755
EV	8	64.5846	3.6300	0.5961	0.3554	3.3078	0.1188
FP	8	211.6642	11.3809	0.7792	0.6071	9.2724	0.0227
MR	8	−10.3754	8.6884	0.9883	0.9767	251.4860	0.0000
ST	8	88.5008	−1.2874	−0.3838	0.1473	1.0366	0.3479
MV	8	−101.1703	29.7489	0.9553	0.9126	62.6887	0.0002

Table 4: Statistical data for Reciprocal Randi¢ Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	328.8706	4.8864	0.8721	0.7606	19.0575	0.0047
VP	8	−1.6329	0.1064	0.1940	0.0377	0.2348	0.6452
EV	8	38.5694	1.0433	0.7388	0.5459	7.2115	0.0363
FP	8	152.6915	2.9545	0.8723	0.7608	19.0882	0.0047
MR	8	−30.3738	1.9050	0.9344	0.8731	41.2892	0.0007
ST	8	81.3104	−0.1400	−0.1800	0.0324	0.2010	0.6696
MV	8	−149.0613	6.2344	0.8633	0.7453	17.5602	0.0057

Table 5: Statistical data for Harmonic Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	552.8484	8.7481	0.6735	0.4536	4.9810	0.0671
VP	8	7.1347	−0.0821	−0.0646	0.0042	0.0252	0.8792
EV	8	89.7797	1.6303	0.4980	0.2480	1.9791	0.2091
FP	8	288.0758	5.2922	0.6740	0.4543	4.9941	0.0668
MR	8	39.4203	4.6384	0.9814	0.9632	156.9670	0.0000
ST	8	83.0850	−0.8248	−0.4574	0.2092	1.5874	0.2545
MV	8	62.1830	16.3824	0.9786	0.9577	135.7078	0.0000

Table 6: Statistical data for Sum Connectivity Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	332.7473	25.5017	0.8507	0.7236	15.7087	0.0074
VP	8	−1.6385	0.5619	0.1915	0.0367	0.2285	0.6495
EV	8	38.5280	5.5091	0.7292	0.5317	6.8115	0.0401
FP	8	155.0358	15.4191	0.8508	0.7239	15.7305	0.0074
MR	8	−33.4412	10.2805	0.9425	0.8882	47.6892	0.0005
ST	8	83.3019	−0.8863	−0.2130	0.0454	0.2851	0.6126
MV	8	−161.5426	33.8250	0.8754	0.7664	19.6854	0.0044

Table 7: Statistical data for First Zagreb Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	406.8515	1.7198	0.8478	0.7188	15.3335	0.0078
VP	8	0.8751	0.0323	0.1627	0.0265	0.1632	0.7002
EV	8	56.0185	0.3621	0.7083	0.5017	6.0406	0.0493
FP	8	199.8338	1.0399	0.8480	0.7191	15.3595	0.0078
MR	8	−4.5250	0.6994	0.9475	0.8978	52.7188	0.0003
ST	8	80.4980	−0.0583	−0.2071	0.0429	0.2689	0.6226
MV	8	−67.1629	2.3059	0.8820	0.7779	21.0171	0.0038

Table 8: Statistical data for Second Zagreb Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	401.3890	1.4271	0.8752	0.7659	19.6288	0.0044
VP	8	−1.0592	0.0363	0.2273	0.0516	0.3268	0.5883
EV	8	53.3171	0.3085	0.7506	0.5635	7.7449	0.0319
FP	8	196.5418	0.8629	0.8753	0.7662	19.6578	0.0044
MR	8	1.2772	0.5389	0.9083	0.8250	28.2768	0.0018
ST	8	76.8541	−0.0286	−0.1264	0.0160	0.0975	0.7655
MV	8	−40.0938	1.7359	0.8259	0.6822	12.8786	0.0115

Table 9: Statistical data for Hyper Zagreb Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	413.9530	0.3185	0.8672	0.7521	18.1995	0.0053
VP	8	−0.1631	0.0074	0.2058	0.0424	0.2654	0.6248
EV	8	56.6604	0.0681	0.7357	0.5412	7.0774	0.0375
FP	8	204.1392	0.1925	0.8674	0.7523	18.2239	0.0053
MR	8	3.9024	0.1228	0.9192	0.8449	32.6845	0.0012
ST	8	77.2690	−0.0072	−0.1411	0.0199	0.1218	0.7390
MV	8	−33.5650	0.3979	0.8408	0.7069	14.4718	0.0089

Table 10: Statistical data for Forgotten Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	378.8184	0.6901	0.8905	0.7930	22.9794	0.0030
VP	8	−2.8829	0.0204	0.2693	0.0725	0.4692	0.5189
EV	8	47.5025	0.1513	0.7749	0.6004	9.0152	0.0239
FP	8	182.9141	0.4172	0.8905	0.7931	22.9945	0.0030
MR	8	−1.3832	0.2471	0.8762	0.7677	19.8268	0.0043
ST	8	73.9314	−0.0060	−0.0562	0.0032	0.0190	0.8948
MV	8	−42.1976	0.7809	0.7818	0.6112	9.4332	0.0219

Table 11: Statistical data for Hyper Forgotten Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	448.9259	0.0327	0.8752	0.7659	19.6295	0.0044
VP	8	0.0855	0.0008	0.2297	0.0528	0.3343	0.5842
EV	8	64.3202	0.0070	0.7396	0.5470	7.2460	0.0360
FP	8	225.2955	0.0198	0.8752	0.7660	19.6439	0.0044
MR	8	22.6159	0.0119	0.8727	0.7615	19.1614	0.0047
ST	8	73.6921	−0.0003	−0.0656	0.0043	0.0259	0.8774
MV	8	32.8025	0.0377	0.7812	0.6103	9.3956	0.0221

Table 12: Statistical data for Rede ned Zagreb Index

Prop.	N	c	m	r	*r2*	F	P
BP	8	416.6582	3.9404	0.8861	0.7852	21.9273	0.0034
VP	8	−3.5105	0.1430	0.3286	0.1080	0.7263	0.4268
EV	8	53.7678	0.8948	0.7984	0.6375	10.5503	0.0175
FP	8	205.7972	2.3821	0.8861	0.7852	21.9323	0.0034
MR	8	16.9169	1.3390	0.8275	0.6848	13.0347	0.0112
ST	8	71.1867	0.0019	0.0031	0.0000	0.0001	0.9942
MV	8	21.9439	4.1369	0.7218	0.5210	6.5258	0.0432

Graphical Representation

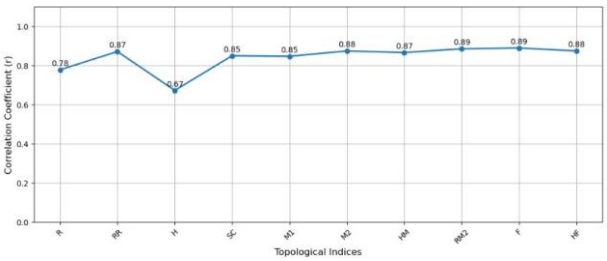


Figure 9: Correlation plot for boiling point (BP).

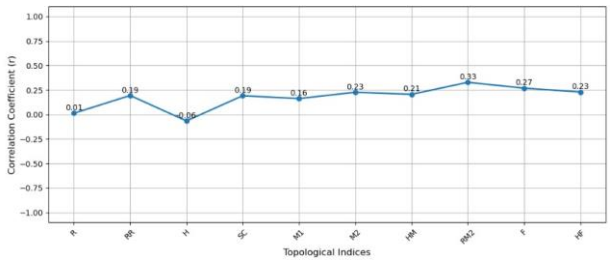


Figure 10: Correlation plot for vapour pressure (VP).

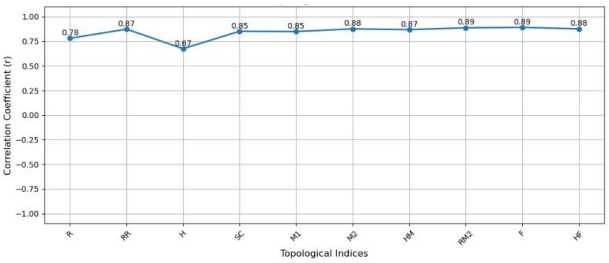


Figure 11: Correlation plot forash point (FP).

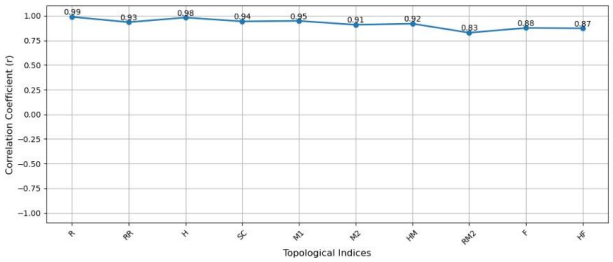


Figure 12: Correlation plot for molar refractivity (MR).

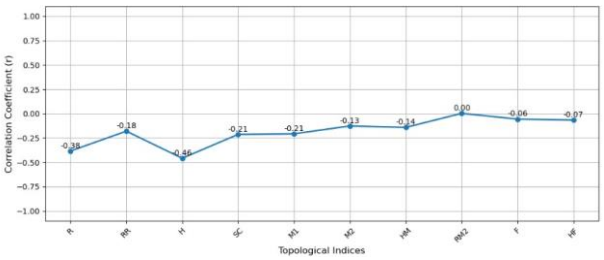


Figure 13: Correlation plot for surface tension (ST).

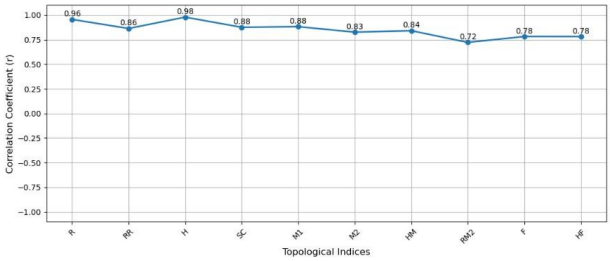


Figure 14: Correlation plot for molar volume (MV).

Graphical Analysis of Correlation Coe cients

To make the pattern of results easier to see at a glance, we plotted the correlation coe cients for each physicochemical property against the ten topological indices. The *x*-axis lists the descriptors and the *y*-axis shows the *r* value. These line graphs were produced in Python using Jupyter Notebook and Matplotlib, and are shown in Figures 9 14 [2,6,9,29].

Looking at the boiling point graph rst, the correlations are consistently high most descriptors give *r* values close to 1. This makes sense. Boiling point in organic molecules is strongly in uenced by molecular size and the degree of branching, and those are exactly the things that degree-based indices capture [1, 5]. Flash point shows almost the same pattern, which is not surprising given how closely the two properties are related in organic compounds. Molar refractivity and molar volume also come out well, with strong positive correlations across nearly all ten descriptors [7,8,26].

Vapour pressure tells a di erent story [4]. The correlations are low, and for several descriptors they are slightly negative. Surface tension is similarly poorly predicted [3,28]. These results are not entirely surprising. Both vapour pressure and surface tension are driven by intermolecular forces, particularly hydrogen bonding and van der Waals interactions, that depend on three-dimensional molecular geometry in ways a two-dimensional graph cannot fully represent. Enthalpy of vaporisation sits somewhere in the middle [2]. The correlations are moderate and positive, suggesting the degree-based descriptors pick up some of the relevant structural variation but not all of it. Overall, the graphs make it clear that boiling point, ash point, molar refractivity, and molar volume are where degree-based QSPR modelling works best for this set of antibiotic compounds [2,6 8,30].

Discussion

The results point fairly clearly in one direction: for four of the seven properties studied boiling point, ash point, molar refractivity, and molar volume degree-based topological indices work well as predictors. For the remaining three, particularly vapour pressure and surface tension, they do not. Understanding why requires looking at the physics.

Boiling point and ash point are both governed largely by the strength of intermolecular attractions and by how easily molecules can escape from the liquid phase. Heavier, more branched molecules tend to boil at higher temperatures, and degree-based indices are essentially measures of molecular size and branching. So it is not surprising that descriptors like the rst Zagreb, hyper Zagreb, and sum connectivity indices all correlate strongly with boiling point [3,5,7,8,26].

Molar refractivity gave the highest individual correlation in the entire study. The Randi¢ index achieved $r = 0.988$ with this property [2], and the reciprocal Randi¢ was not far behind at $r = 0.934$. Molar refractivity re ects the polarisability of a molecule how easily its electron cloud deforms in an electric eld. Larger, more connected molecules tend to be more polarisable, and degree-based indices track exactly this kind of structural information [6,9,30].

Vapour pressure and surface tension are another matter. Both depend heavily on the speci c arrangement of polar groups and hydrogen bonding sites in three-dimensional space features that a two-dimensional graph simply cannot capture [2 4, 7, 28].

The low and sometimes negative correlations we found for these properties are consistent with what others have reported for similar compound classes, and they highlight a known limitation of twodimensional topological indices.

Enthalpy of vaporisation sits between the two groups. The correlations are real but moderate, reecting the fact that enthalpy depends partly on molecular size which indices can capture and partly on speci c intermolecular interactions that they cannot [2,4,28].

Overall, these ndings line up well with recent QSPR work on other drug families. Degree-based descriptors are a reliable and practical tool for a subset of physicochemical properties, particularly those that scale with molecular weight and branching. For properties that depend more on three-dimensional structure or electronic speci cs, more sophisticated descriptors will be needed [1,6,8,9,29].

Conclusion

This study set out to test how well ten degree-based topological indices can predict seven physicochemical properties of eight common antibiotic drugs. We computed all index values using the edge partitioning method, built simple linear regression models in MS Excel, and visualised the results in Jupyter Notebook using Python and Matplotlib.

The short answer is that these indices work well for some properties and poorly for others. Boiling point, ash point, molar refractivity, and molar volume all showed strong, statistically signi cant correlations with most of the ten descriptors. The best individual result was the Randi¢ index predicting molar refractivity at r = 0.988. The harmonic index also performed notably well for both molar refractivity and molar volume. The sum connectivity index was among the most consistent performers across multiple properties. The rst and second Zagreb indices, the hyper Zagreb index, the forgotten index, the hyper forgotten index, and the rede ned Zagreb index all gave reliable results for boiling point and ash point. Vapour pressure and surface tension were harder to predict. For most descriptors, the correlations were low and in some cases slightly negative. This is not a failure of the method so much as a re ection of what these descriptors can and cannot encode. Properties that depend on three-dimensional geometry and intermolecular forces are simply beyond what a twodimensional graph-based descriptor can fully capture. What this study shows, in practical terms, is that degree-based topological indices can be a quick, lowcost way to estimate several important physicochemical properties of antibiotic compounds before any laboratory work is done. The approach is not perfect, but for screening purposes it is genuinely useful.

Future Work

We intend to extend this work to other drug classes, including anticancer, antidiabetic, and cardiovascular compounds. We also plan to test multilinear and curvilinear regression models, which may improve accuracy for properties like vapour pressure where simple linear ts fall short. Including three-dimensional molecular descriptors in future models would be a natural next step to address the limitations identied here.

Acknowledgement

This research was supported by the Ongoing Research Funding Program (ORF-2026-996) at King Saud University in Riyadh, Saudi Arabia.

References

M. Reinhard and A. Drefahl, Handbook for Estimating Physicochemical Properties of Organic Compounds. Wiley, New York, 1999, p. 39.

M. W. Rasheed, A. Mahboob, and I. Hanif, An estimation of physicochemical properties of heart attack treatment medicines by using molecular descriptors, South African Journal of Chemical Engineering, vol. 45, pp. 20 29, 2023.

B. Lian and S. H. Yalkowsky, Uni ed physicochemical property estimation relationships (UPPER), Journal of Pharmaceutical Sciences, vol. 103, no. 9, pp. 2710 2723, 2014.

G. Isaacman-VanWertz and B. Aumont, Impact of organic molecular structure on the estimation of atmospherically relevant physicochemical parameters, Atmospheric Chemistry and Physics, vol. 21, no. 8, pp. 6541 6563, 2021.

W. J. Lyman, Estimation of physical properties, in Environmental Exposure from Chemicals, CRC Press, 2018, pp. 13 48.

I. Haider, M. Li, and M. K. Jamil, Graph-based stacking ensemble approach for physicochemical properties prediction of oncology-relevant compounds, The Journal of Supercomputing, vol. 82, no. 8, p. 428, 2026.

S. Sorgun and K. Birgin, Vertex-edge-weighted molecular graphs: a study on topological indices and their relevance to physicochemical properties of drugs used in cancer treatment, Journal of Chemical Information and Modeling, vol. 65, no. 4, pp. 2093 2106, 2025.

H. Hayat, M. K. Siddiqui, and S. Ahmad, A graph theoretical and multicriteria decision framework for physicochemical evaluation of drugs used in cancer treatment via graph energy, Chemical Papers, pp. 1 25, 2026.

W. J. Boecklen and G. J. Niemi, Multivariate association of graph-theoretic variables and physicochemical properties, SAR and QSAR in Environmental Research, vol. 2, no. 1 2, pp. 79 87, 1994.

H. Wiener, Structural determination of para n boiling points, Journal of the American Chemical Society, vol. 69, no. 1, pp. 17 20, 1947.

M. Randi¢, Characterization of molecular branching, Journal of the American Chemical Society, vol. 97, no. 23, pp. 6609 6615, 1975.

I. Gutman and N. Trinajsti¢, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, Chemical Physics Letters, vol. 17, no. 4, pp. 535 538, 1972.

B. Furtula and I. Gutman, A forgotten topological index, Journal of Mathematical Chemistry, vol. 53, no. 4, pp. 1184 1190, 2015.

M. Adnan, S. A. U. H. Bokhary, G. Abbas, and T. Iqbal, Degree-based topological indices and QSPR analysis of antituberculosis drugs, Journal of Chemistry, 2022, Art. no. 5748626.

R. I. Aminov, A brief history of the antibiotic era: lessons learned and challenges for the future, Frontiers in Microbiology, vol. 1, p. 134, 2010.

M. I. Hutchings, A. W. Truman, and B. Wilkinson, Antibiotics: past, present and future, Current Opinion in Microbiology, vol. 51, pp. 72 80, 2019.

A. Fleming, On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. in uenzae, British Journal of Experimental Pathology, vol. 10, no. 3, pp. 226 236, 1929.

E. M. Darby, E. Trampari, P. Siasat et al., Molecular mechanisms of antibiotic resistance revisited, Nature Reviews Microbiology, vol. 21, no. 5, pp. 280 295, 2023.

N. Trinajsti¢, Chemical Graph Theory, 2nd ed. CRC Press, Boca Raton, FL, 1992.

S. Fajtlowicz, On conjectures of Gra ti, Annals of Discrete Mathematics, vol. 38, pp. 113 118, 1988.

B. Zhou and N. Trinajsti¢, On a novel connectivity index, Journal of Mathematical Chemistry, vol. 46, no. 4, pp. 1252 1270, 2009.

G. H. Shirdel, H. Rezapour, and A. M. Sayadi, The hyper-Zagreb index of graph operations,

Iranian Journal of Mathematical Chemistry, vol. 4, no. 2, pp. 213 220, 2013.

P. S. Ranjini, V. Lokesha, and A. Usha, Relation between phenylene and hexagonal squeeze using harmonic index, International Journal of Graph Theory, vol. 1, no. 4, pp. 116 121, 2013.

B. Chaluvaraju and V. Munaswamy, The hyperforgotten topological index of a graph, Journal of Computer and Mathematical Sciences, vol. 8, no. 9, pp. 473 481, 2017.

M. Javaid, M. U. Rehman, and J. Cao, Computing analysis of Zagreb indices for generalized sum graphs under strong product, Journal of Chemistry, 2021, Art. no. 6663624.

A. Mahboob, M. W. Rasheed, I. Hanif, L. Amin, and A. Alameri, QSPR analysis of some important drugs used in heart attack treatment via degree-based topological indices and regression models, Polycyclic Aromatic Compounds, vol. 44, no. 8, pp. 5237 5246, 2024.

M. Kalaimathi and B. J. Balamurugan, Topological indices of molecular graphs of monkeypox drugs for QSPR analysis to predict physicochemical and ADMET properties, International Journal of Quantum Chemistry, vol. 123, no. 19, p. e27210, 2023.

M. S. Alam, A. Mahboob, I. Siddique, and M. W. Rasheed, On quantitative structureproperty relationship (QSPR) analysis of physicochemical properties and anti-hepatitis prescription drugs using a linear regression model, Heliyon, vol. 10, no. 4, p. e26125, 2024.

A. A. Khabyah, S. Zaman, A. N. Koam, A. Ahmad, and A. Ullah, QSAR analysis of drugs using graph-based degree-based topological indices and regression models, Scienti c Reports, vol. 15, p. 9832, 2025.

T. Arockiaraj, A. B. Greeni, and A. A. Kalaam, A QSPR analysis and curvilinear regression models for various degree-based topological indices: quinolone antibiotics, Heliyon, vol. 10, no. 12, p. e32919, 2024.



2026 by the authors; Journal of *Kashmir Journal of Academic Research and Development*. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).