

Quantitative Structure-Property Relationship Analysis on Priority PAHs Using Certain Closed Neighbourhood Topological Indices

Melaiyur Sankarraman Srinivasan^{1,*} , Ranjan Helen¹ 

¹ PG & Research Department of Mathematics, Poompuhar College (Autonomous) (Affiliated to Bharathidasan University, Tiruchirappalli), Melaiyur - 609107, Tamil Nadu, India; mssn84@gmail.com (M.S.S.); helenranjan@gmail.com (R.H);

* Correspondence: mssn84@gmail.com (M.S.S.);

Scopus Author ID 57372448200

Received: 11.06.2022; Accepted: 17.07.2022; Published: 10.09.2022

Abstract: Molecular structures and their physicochemical properties are related harmoniously. Having the potential uses in our everyday life, graph theory has become one of the most iconic and discussed areas of applied mathematics. Chemical graph theory (CGT) is a branch of graph theory that incorporates chemical aspects. The topological index (TI) is a well-defined graph-theoretical tool that offers a mathematical relationship with molecular structures and also characterizes their topology. TI has real-time applications such as isomer discrimination, drug design, QSPR, and QSAR studies in various domains of chemistry, including nanotechnology and biochemistry. Herein, seven closed neighborhood TIs are investigated. Using QSPR regression analysis, the chemical signature of the indices in predicting the physicochemical properties of priority polycyclic aromatic hydrocarbons (PAHs) is explored. These Tis exhibited a strong correlation with certain properties of priority PAHs. Certain statistical aspects of these indices are discussed, and the significant results have also been represented graphically.

Keywords: topological index; closed neighbourhood degree sum; Zagreb index; Forgotten index; PAHs.

© 2022 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Consider $G = (V, E)$ be a graph with τ edges and η vertices. The edge sw connects the vertices s and w . The number of vertices neighboring to s is the degree of a vertex $s \in V(G)$, represented by $d_G(s)$. Let $N[w]$ be the closed neighborhood set of vertex w that includes w and its neighbors. The graphs considered throughout this article are finite, connected, undirected, and loopless with no parallel edges [1].

1.1. Background.

Mathematical chemistry has made rapid progress in recent years, with the advent of novel graph theoretical approaches and interdisciplinary techniques. CGT is a mathematical field that incorporates graph theory with chemistry [2]. In CGT, chemical compounds are depicted as molecular graphs. Hence, every molecule involved in a chemical transformation can be represented by a molecular graph. In general, molecular graphs are simplified hydrogen-suppressed graphs that reflect the carbon skeleton of an organic compound, where the vertex

indicates a non-hydrogen atom and the edge denotes a covalent bond among the atoms [2]. TIs are numerical values associated with chemical substances' inter-molecular relationships and bio-activity [3]. One such important classification of TIs is a degree-based index, which is the most studied TI in structure-property relationships and supports the drug design process's theoretical framework [3]. In 1975, Milan Randić established the first degree-based TI known as the product connectivity index, widely used in drug development and theoretically associated with the normalized Laplacian matrix [4,5]. The first ($M_1(G)$) and second Zagreb ($M_2(G)$) indices were established by Gutman *et al.* [6] to calculate the π electron energy of molecules. Later, the first ($HM_1(G)$) and second hyper Zagreb ($HM_2(G)$) indices were proposed by Shirdel *et al.* [7]. Another important degree-based TI is the Forgotten index ($F(G)$), which was introduced by Furtula *et al.* [8] for predicting the Physicochemical properties of chemical compounds. The definitions of these well-known indices are given in Eq.(1) - Eq.(6). For more about TIs, we recommend the texts [9-26] and the references therein.

$$R(G) = \sum_{sw \in E(G)} \frac{1}{\sqrt{d_G(s) \cdot d_G(w)}} \quad (1)$$

$$M_1(G) = \sum_{s \in V(G)} d_G(s)^2 = \sum_{sw \in E(G)} (d_G(s) + d_G(w)) \quad (2)$$

$$M_2(G) = \sum_{sw \in E(G)} (d_G(s) d_G(w)) \quad (3)$$

$$F(G) = \sum_{s \in V(G)} d_G(s)^3 = \sum_{sw \in E(G)} (d_G(s)^2 + d_G(w)^2) \quad (4)$$

$$HM_1(G) = \sum_{sw \in E(G)} (d_G(s) + d_G(w))^2 \quad (5)$$

$$HM_2(G) = \sum_{sw \in E(G)} (d_G(s) d_G(w))^2 \quad (6)$$

1.2. Associated TIs.

Herein, we deal with some closed neighborhood TIs [27-29] based on the sum of the degrees of all vertices in a closed neighborhood of vertex s in $V(G)$. The closed neighborhood degree-sum of the vertex s in $V(G)$ and other indices are defined as follows,

$$\Omega_G(s) = \sum_{w \in N[s]} d_G(w) \quad (7)$$

Closed neighborhood first Zagreb index

$$CM_1(G) = \sum_{sw \in E(G)} (\Omega_G(s) + \Omega_G(w)) \quad (8)$$

Modified closed neighborhood first Zagreb index

$$CM_1^*(G) = \sum_{s \in V(G)} \Omega_G(s)^2 \quad (9)$$

Closed neighborhood second Zagreb index

$$CM_2(G) = \sum_{sw \in E(G)} (\Omega_G(s) \Omega_G(w)) \quad (10)$$

Closed neighborhood Forgotten index

$$CF(G) = \sum_{sw \in E(G)} (\Omega_G(s)^2 + \Omega_G(w)^2) \quad (11)$$

Modified closed neighborhood Forgotten index

$$CF^*(G) = \sum_{s \in V(G)} \Omega_G(s)^3 \quad (12)$$

Closed neighborhood first hyper Zagreb index

$$CHM_1(G) = \sum_{sw \in E(G)} (\Omega_G(s) + \Omega_G(w))^2 \quad (13)$$

Closed neighborhood second hyper Zagreb index

$$CHM_2(G) = \sum_{sw \in E(G)} (\Omega_G(s) \Omega_G(w))^2 \quad (14)$$

1.3. Preliminary testing.

In order to determine the usefulness of a TI in predicting the physicochemical behavior of molecules, we conduct a regression analysis, as the International Academy of Mathematical Chemistry recommended.

$$\text{Linear Regression Model: } \xi = v + \mu(TI). \quad (15)$$

where ξ = Physical Property, TI = Topological Index, μ , and v are constants.

For preliminary testing of indices, octane isomers are mainly helpful as they constitute a fairly large and structurally diverse set of alkanes [30] (See Figure 1).

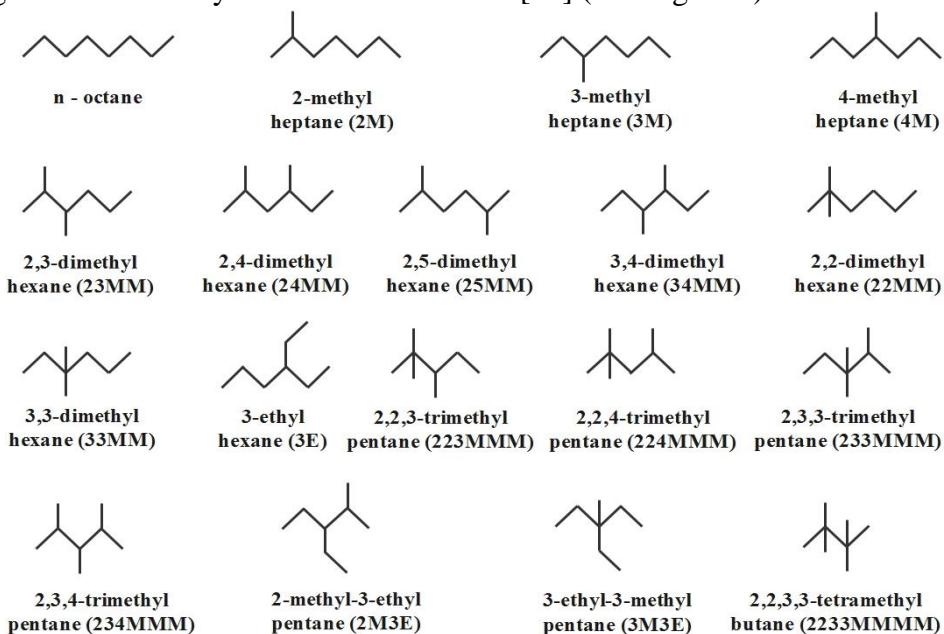


Figure 1. Different isomers of octane.

Herein, we focused on the newly developed indices for modeling physicochemical properties such as Boiling Points (BP), Critical Pressure (PC), Entropy (S), Density (D), Mean Radius (R_m^2), Heats of Formation($-\Delta H_f$), Heats of Vaporization ($-\Delta H_v$) at $25^\circ C$, Acentric Factor (AF), Molar Volumes (MV) at $20^\circ C$, Molar Refractions (MR) at $20^\circ C$, Enthalpy of Vaporization (HVAP) and Standard Enthalpy of Vaporization (DHVAP) of octane isomers.

The experimental data for the physicochemical characteristics of octane isomers (Table 1) were obtained from www.moleculardescriptors.eu. Before dealing with the linear regression model, we calculate the TIs of octane isomers, as shown in Table 2.

Table 1. Experimental values of physicochemical properties for octane isomers.

Octane	BP	PC	S	D	R_m^2	$-\Delta H_f$	$-\Delta H_v$	AF	MV	MR	DHVAP	HVAP
<i>n</i> -octane	125.7	24.64	111.67	0.7025	2.0449	208.6	41.49	0.397898	162.592	39.1922	9.915	73.19
2M	117.6	24.8	109.84	0.698	1.8913	215.4	39.67	0.377916	163.663	39.2316	9.484	70.3
3M	118.9	25.6	111.26	0.7058	1.7984	212.5	39.83	0.371002	161.832	39.1001	9.521	71.3
4M	117.7	25.6	109.32	0.7046	1.7673	210.7	39.64	0.371504	162.105	39.1174	9.483	70.91
3E	118.5	25.74	109.43	0.7136	1.7673	210.7	39.64	0.362472	160.07	38.94	9.476	71.7
22MM	106.8	25.6	103.42	0.6953	1.6744	224.6	37.28	0.339426	164.28	39.25	8.915	67.7
23MM	115.6	26.6	108.02	0.7121	1.6464	213.8	38.78	0.348247	160.39	38.98	9.272	70.2
24MM	109.4	25.8	106.98	0.7004	1.6142	219.2	37.76	0.344223	163.09	39.13	9.029	68.5
25MM	109.1	25	105.72	0.6935	1.6449	222.5	37.85	0.35683	164.69	39.25	9.051	68.6
33MM	112	27.2	104.74	0.71	1.7377	220	37.53	0.322596	160.87	39	8.973	68.5
34MM	117.7	27.4	106.59	0.72	1.523	212.8	38.97	0.340345	158.81	38.84	9.316	70.2
2M3E	115.6	27.4	106.06	0.7193	1.5525	211	38.52	0.332433	158.79	38.83	9.209	69.7
3M3E	118.3	28.9	101.48	0.7274	1.5212	214.8	37.99	0.306899	157.02	38.71	9.081	69.3
223M MM	109.8	28.2	101.31	0.7161	1.4306	220	36.91	0.300816	159.52	38.92	8.826	67.3
224M MM	99.24	25.5	104.09	0.6919	1.401	224	35.14	0.30537	165.08	39.26	8.402	64.87
233M MM	114.8	29	102.06	0.7262	1.4931	216.3	37.27	0.293177	157.29	38.76	8.897	68.1
234M MM	113.5	27.6	102.39	0.7191	1.3698	217.3	37.75	0.317422	158.85	38.86	9.014	68.37
2233M MMM	106.5	24.5	93.06	0.8242	1.4612	225.6	42.9	0.255294			8.41	66.2

Table 2. TIs of octane isomers.

Octane	$CM_1(G)$	$CM_1^*(G)$	$CM_2(G)$	$CF(G)$	$CF^*(G)$	$CHM_1(G)$	$CHM_2(G)$
<i>n</i> -octane	74	212	198	1168	406	802	6138
2M	80	236	228	1398	480	936	8154
3M	82	244	241	1530	518	1000	9677
4M	82	246	244	1566	522	1010	10196
3E	84	254	257	1710	562	1076	11899
22MM	92	290	300	1984	658	1258	14688
23MM	90	276	285	1928	640	1210	13947
24MM	88	270	274	1802	596	1144	12452
25MM	86	260	259	1628	554	1072	10339
33MM	96	306	326	2284	744	1396	18566
34MM	92	284	297	2072	680	1274	15633
2M3E	96	286	316	2126	758	1390	17448
3M3E	100	320	349	2566	830	1528	22135
223MMM	104	336	368	2682	872	1608	23024
224MMM	95	318	323	2304	707	1353	17645
233MM	106	342	379	2808	916	1674	24813
234MMM	98	308	328	2344	764	1420	18496
2233MMMM	118	392	451	3412	1118	2020	32791

The significant outcomes of the above linear model are listed in Tables 3-10. The correlation coefficient indicates the strength of the linear relationship. The accuracy in measuring the regression coefficient is determined by the standard error of the regression. Significance F can be used to determine if the results are consistent. The result is statistically

significant if this value is less than 0.05. Unless the significance F is more than 0.05, it is highly recommended to abandon using such a set of independent variables.

Table 3. Statistical parameters for the linear QSPR model for $CM_1(G)$.

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	141.1938	-0.2975	-0.51941	0.22414	5.393	5.91126	0.02717
PC	18	20.547	0.06328	0.47057	0.17278	1.307	4.55075	0.04874
S	18	142.5055	-0.40148	-0.94992	0.89625	1.455	147.8624	1.70E-09
D	18	0.53069	0.002	0.73157	0.50614	0.021	18.42293	5.60E-04
R_m^2	18	2.89115	-0.01365	-0.80594	0.62763	0.110	29.65406	5.39E-05
$-\Delta H_f$	18	191.0258	0.27741	0.56172	0.27275	4.502	7.37563	0.01527
AF	18	0.63729	-0.00326	-0.98405	0.96639	0.006	489.7433	2.00E-13
MV	17	176.4981	-0.16927	-0.58616	0.29982	2.134	7.85121	0.01341
MR	17	40.25123	-0.01353	-0.64906	0.3827	0.145	10.91941	0.00481
DHVAP	18	11.84625	-0.02944	-0.82095	0.65358	0.226	33.07275	2.98E-05
HVAP	18	82.13176	-0.14037	-0.74046	0.52005	1.404	19.42041	4.41E-04

Table 4. Statistical parameters for the linear QSPR model for $CM_1^*(G)$.

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	137.2413	-0.08178	-0.59112	0.30877	5.0901	8.59376	0.00978
S	18	133.4398	-0.09739	-0.95405	0.9046	1.395	162.206	8.64E-10
D	18	0.5836	4.59E-04	0.69412	0.44941	0.022	14.87594	0.00139
R_m^2	18	2.58727	-0.00333	-0.81315	0.64004	0.109	31.22743	4.08E-05
$-\Delta H_f$	18	195.1683	0.07467	0.62597	0.35383	4.243	10.30867	0.00545
AF	18	0.5653	-7.98E-04	-0.99569	0.99086	0.003	1844.013	0
MV	17	170.601	-0.03368	-0.48747	0.1868	2.300	4.67529	0.04717
MR	17	39.80094	-0.00277	-0.55471	0.26155	0.158	6.66701	0.02083
DHVAP	18	11.30003	-0.00755	-0.87208	0.74555	0.193	50.81169	2.40E-06
HVAP	18	79.69864	-0.03661	-0.79958	0.61678	1.254	28.36073	6.83E-05

Table 5. Statistical parameters for the linear QSPR model for $CM_2(G)$.

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	129.3655	-0.0520	-0.52461	0.22991	5.373	6.07548	0.02541
PC	18	23.1668	0.0107	0.46047	0.16278	1.315	4.30534	0.05448
S	18	126.3572	-0.0695	-0.95099	0.8984	1.440	151.3264	1.43E-09
D	18	0.6110	0.0003	0.73384	0.50968	0.020	18.6709	5.27E-04
R_m^2	18	2.3305	-0.0023	-0.79372	0.60687	0.114	27.24237	8.43E-05
$-\Delta H_f$	18	202.0949	0.0483	0.5658	0.27764	4.487	7.53398	0.01439
AF	18	0.5065	-0.0006	-0.98811	0.97488	0.006	660.7504	1.94E-14
MV	17	169.5935	-0.0290	-0.57683	0.28824	2.152	7.4796	0.01535
Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
MR	17	39.6999	-0.0023	-0.63918	0.36913	0.146	10.3617	0.00574
DHVAP	18	10.6716	-0.0051	-0.82695	0.66409	0.222	34.60955	2.31E-05
HVAP	18	76.5177	-0.0244	-0.74452	0.52645	1.394	19.89882	3.94E-04

Table 6. Statistical parameters for the linear QSPR model for CF(G).

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	124.9322	-0.00541	-0.49755	0.20053	5.474	5.26398	0.03564
PC	18	23.84501	0.00123	0.48114	0.18347	1.299	4.81976	0.04323
S	18	121.0963	-0.00757	-0.94213	0.88059	1.561	126.361	5.27E-09
D	18	0.6364	3.82E-05	0.73478	0.51115	0.0204	18.7754	5.14E-04
R_m^2	18	2.16084	-2.56E-04	-0.79578	0.61035	0.113	27.62886	7.83E-05
$-\Delta H_f$	18	206.2099	0.00504	0.53701	0.24391	4.590	6.48406	0.02156
AF	18	0.46483	-6.23E-05	-0.988	0.97465	0.005	654.5058	2.09E-14
MV	17	167.727	-0.00332	-0.60654	0.32576	2.094	8.73026	0.00984
MR	17	39.54524	-2.62E-04	-0.66521	0.40534	0.142	11.90632	0.00357
DHVAP	18	10.27397	-5.54E-04	-0.81253	0.63897	0.230	31.08764	4.18E-05
HVAP	18	74.59751	-0.00262	-0.72782	0.50033	1.432	18.02264	6.17E-04

Table 7. Statistical parameters for the linear QSPR model for $CF^*(G)$.

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	124.5837	-0.01588	-0.45795	0.16033	5.610	4.24606	0.05599
PC	18	23.70354	0.00393	0.48242	0.18477	1.298	4.85309	0.0426
S	18	121.8705	-0.02403	-0.93912	0.87457	1.6	119.5392	7.83E-09
D	18	0.62915	1.26E-04	0.7619	0.55427	0.02	22.13961	2.38E-04
R_m^2	18	2.17461	-7.95E-04	-0.77554	0.57655	0.118	24.14635	1.56E-04
$-\Delta H_f$	18	206.3303	0.01508	0.50424	0.20765	4.699	5.45507	0.03286
AF	18	0.46983	-1.96E-04	-0.9749	0.94733	0.008	306.7905	7.32E-12
MV	17	168.7326	-0.01156	-0.6517	0.38636	1.998	11.07381	0.00459
MR	17	39.61864	-9.05E-04	-0.70739	0.4671	0.134	15.02431	0.00149
DHVAP	18	10.28637	-0.00169	-0.78017	0.58421	0.247	24.88647	1.34E-04
HVAP	18	74.5977	-0.00794	-0.69139	0.4454	1.509	14.65283	0.00148

Table 8. Statistical parameters for the linear QSPR model for CHM₁(G).

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	126.4782	-0.00992	-0.48585	0.18831	5.516	4.94391	0.04094
PC	18	23.46945	0.00227	0.47382	0.17603	1.305	4.63189	0.04699
S	18	123.7392	-0.01424	-0.94488	0.8861	1.525	133.2598	3.60E-09
D	18	0.62129	7.32E-05	0.75104	0.53681	0.199	20.70187	3.28E-04
R_m^2	18	2.23914	-4.73E-04	-0.78375	0.59016	0.116	25.47922	1.19E-04
HF	18	204.6422	0.00933	0.53008	0.23605	5.614	6.25274	0.02365
AF	18	0.48511	-1.16E-04	-0.98125	0.96052	0.007	414.6436	7.25E-13
MV	17	169.1668	-0.00647	-0.62163	0.34551	2.063	9.4466	0.00772
Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
MR	17	39.65771	-5.11E-04	-0.68016	0.42679	0.139	12.91321	0.00266
DHVAP	18	10.44314	-0.00102	-0.80018	0.61781	0.237	28.48008	6.68E-05
HVAP	18	75.37395	-0.00482	-0.71393	0.47905	1.462	16.63294	8.75E-04

Table 9. Statistical parameters for the linear QSPR model for CHM₂(G).

Physical properties	n	ν	μ	R	Adj. R square	SE	F	Sig. F
BP	18	120.325	-4.14E-04	-0.45285	0.15539	5.626	4.12773	0.05914
PC	18	24.78825	1.00E-04	0.46787	0.17009	1.309	4.48409	0.05022
S	18	115.5368	-6.33E-04	-0.93891	0.87415	1.603	119.0786	8.05E-09
D	18	0.66182	3.36E-06	0.77007	0.56757	0.019	23.31275	1.85E-04
R_m^2	18	1.957	-2.04E-05	-0.75687	0.54615	0.122	21.45712	2.77E-04
HF	18	210.3204	3.96E-04	0.50283	0.20614	4.703	5.41434	0.03343
AF	18	0.41844	-5.17E-06	-0.97711	0.95191	0.007	337.5268	3.53E-12
MV	17	165.7173	-3.07E-04	-0.6486	0.38206	2.004	10.89247	0.00486
MR	17	39.38108	-2.39E-05	-0.70139	0.45808	0.135	14.5248	0.0017
DHVAP	18	9.84008	-4.46E-05	-0.78017	0.58421	0.247	24.88573	1.34E-04
HVAP	18	72.49425	-2.08E-04	-0.68876	0.44154	1.514	14.44074	0.00157

Table 10. Linear correlation among the physicochemical properties of octane isomers with TIs.

	$CM_1(G)$	$CM_1^*(G)$	$CM_2(G)$	$CF(G)$	$CF^*(G)$	$CHM_1(G)$	$CHM_2(G)$
BP	-0.519	-0.591	-0.525	-0.498	-0.458	-0.486	-0.453
PC	0.471	0.433	0.460	0.481	0.482	0.474	0.468
S	-0.950	-0.954	-0.951	-0.942	-0.939	-0.945	-0.939
D	0.732	0.694	0.734	0.735	0.762	0.751	0.770
R_m^2	-0.806	-0.813	-0.794	-0.796	-0.776	-0.784	-0.757
$-\Delta H_f$	0.562	0.626	0.566	0.537	0.504	0.530	0.503
$-\Delta H_v$	-0.125	-0.182	-0.124	-0.122	-0.073	-0.094	-0.061
AF	-0.984	-0.996	-0.988	-0.988	-0.975	-0.981	-0.977
MV	-0.586	-0.487	-0.577	-0.607	-0.652	-0.622	-0.649
MR	-0.649	-0.555	-0.639	-0.665	-0.707	-0.680	-0.701
DHVAP	-0.821	-0.872	-0.827	-0.813	-0.780	-0.800	-0.780
HVAP	-0.740	-0.801	-0.745	-0.728	-0.691	-0.714	-0.689

We reserve the symbols n, ν , μ , ψ , R, adj.R square, SE, F, Sig F for the number of pairs, intercept, slope, correlation coefficient, adjusted R square, standard error, F-test, and significance F, respectively.

It is worth mentioning that the value of the absolute correlation coefficient ($|R|$) represents the suitability of the TI as a tool for predicting the physicochemical characteristics of chemical compounds in QSPR research. It is noteworthy that these TIs are chemically significant because they disclose better correlations ($|R| > 0.8$) with AF, S, DHVAP, R_m^2 , and HVAP of octane isomers, as displayed in Table 10.

1.3.1. Degeneracy test.

The goal of a TI is to decipher the structural property of compounds as much as feasible. A good topological descriptor should be able to distinguish the different structural formulae. More structural information is captured by TIs with strong discriminating power. The fact that two or more isomers have the same TI is known as degeneracy, which seems to be a major flaw among the TIs. We employ Konstantinova's sensitivity [31] as a measure of degeneracy, which is defined below,

$$S_I = \frac{H - H_I}{H} \quad (16)$$

where H denotes the total number of isomers investigated and H_I represents the number of isomers that cannot be distinguished by the TI I . As S_I grows, so does the isomer-discrimination power of TIs.

The mean isomer degeneracy (M_d) was defined by Bonchev et al. in 1981 [32].

$$M_d = \frac{H}{X} \quad (17)$$

where X denotes the number of unique values assigned to these isomers by the index. Unambiguously, the least value of d is 1. The isomer-discrimination power of TIs decreases as d increases. Compared to other kinds of molecular descriptors, vertex degree-based TIs offer more discriminating power. The considered indices outperform well-established degree-based indices and exhibit a better response to octane isomers (Table 11).

Table 11. The measure of sensitivity (S_I) of different indices for octane isomers.

TI	Sensitivity (S_I)	Mean Degeneracy
$CM_1^*(G)$	1.000	1.000
CF(G)	1.000	1.000
$CM_2(G)$	1.000	1.000
$CF^*(G)$	1.000	1.000
$CHM_1(G)$	1.000	1.000
$CHM_2(G)$	1.000	1.000
$CM_1(G)$	0.833	1.200
Merrifield - Simmons index	0.833	1.200
Hyper Zagreb index	0.833	1.200
Hosoya index Z	0.778	1.286
Second Zagreb index $M_2(G)$	0.722	1.385
Forgotten Index F(G)	0.389	2.571
First Zagreb index $M_1(G)$	0.333	3.000

1.4. A brief sketch on PAHs.

PAHs are identified as common environmental pollutants. The combustion of fossil fuels, oil refining, coal industry, tobacco smoking, wood preservation (creosote), and different kinds of cooking have all been manufactured sources of PAHs. Wildfires and volcanoes are a few examples of natural sources [33]. The United States Environmental Protection Agency (USEPA) has categorized 16 PAHs as high-priority contaminants out of the hundreds of known PAHs (Figure 2). Due to their potential toxicity in living organisms and their abundance and persistence, these 16 PAHs seem to be of environmental concern. Many PAHs are known or suspected carcinogens [34].

1.5. Objectives and construction.

In this scenario, we primarily focus on revealing the chemical applicability of the closed neighborhood TIs [12] in terms of QSPR analysis on priority PAHs, which is precisely the contribution of this study. The design of this article is as follows: Section 1 explains the key

concepts and preliminary testing of the considered indices. The methods and techniques employed in this study are given in section 2. Section 3 deals with the QSPR studies of priority PAHs. Concluding remarks are provided in Section 4.

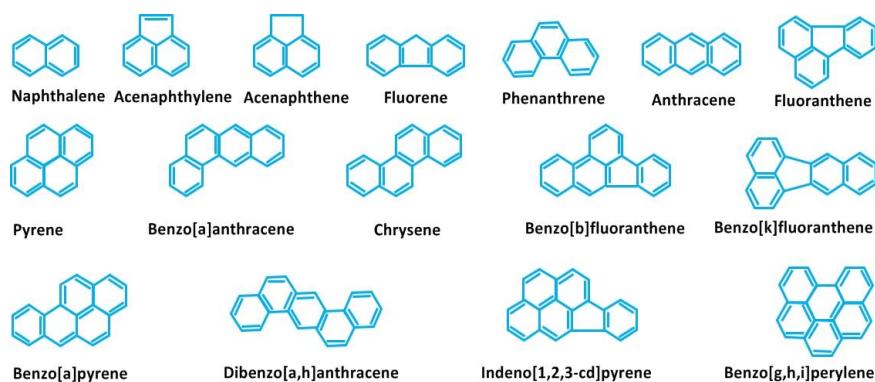


Figure 2. Molecular Structure of priority PAHs.

2. Materials and Methods

We adopt statistical linear regression analysis to establish the prediction efficacy of TIs towards the physicochemical characteristics of priority PAHs. Graphical representation of indices was done by SPSS 16.0 software. We use combinatorial computing, graph theoretical tools, degree counting, and vertex and edge partitioning methods to obtain our key results.

3. Results and Discussion

In this part, we investigated the chemical applicability of TIs in predicting the various physicochemical properties [Boiling Point (BP), Solubility in Water (C_w^{sat}), Molecular Weight (MW), Melting Point (MP), n-octanol air partition coefficient ($\text{Log } K_{oa}$) and n-octanol water partition coefficient ($\text{Log } K_{ow}$)] of priority PAHs. The experimental data for the physical and chemical properties of priority PAHs are shown in Table 12.

Table 12. Experimental data of priority PAHs [35].

PAHs	BP (°C)	MP (°C)	MW g/mol	C_w^{sat} m g/l	$\text{Log } K_{ow}$	$\text{Log } K_{ow}$ at 25 °C
Naphthalene	218	80.2	128	31.0	3.5	5.13
Acenaphthylene	280	92.5	152	16.1	3.85	6.34
Acenaphthene	279	93.4	152	3.8	3.92	6.52
Fluorene	295	115	166	1.9	4.11	6.90
Phenanthrene	340	99.2	178	1.1	4.47	7.68
Anthracene	340	215	178	0.045	4.45	7.71
Fluoranthene	384	108	202	0.26	4.90	8.76
Pyrene	404	151	202	0.132	5.18	8.61
Benz[a] anthracene	435	167	228	0.011	5.91	10.28
Chrysene	448	258	228	0.0015	5.79	10.30
Benzo[b] fluoranthene	481	168	252	0.0015	5.78	11.34
Benzo[k] fluoranthene	480	217	252	0.0008	6.11	11.37
Benzo[a]pyrene	495	177	252	0.0038	6.35	11.56
Dibenzo[a,h] anthracene	524	270	278	0.0005	6.75	12.59
Indeno[1,2,3-c,d] pyrene	536	164	276	0.062	6.70	12.55
Benzo[g,h,i] perylene	550	278	276	0.00026	6.90	12.43

3.1. Linear regression model for Priority PAHs.

To investigate the prediction efficacy of the suggested TIs, we employed a linear regression model as given in Eq. (15). Using this model, one can obtain the following results (Tables 13-19).

Table 13. Statistical parameters for the linear QSPR model for CM₁* (G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	83.50601	0.26973	0.9808	0.95926	20.980	354.1603	2.45E-11
MP	16	20.87875	0.1214	0.69361	0.44402	49.326	12.97956	0.00288
MW	16	59.22956	0.12837	0.98032	0.95824	10.113	345.2324	2.91E-11
C _w ^{sat}	16	19.51132	-0.01349	-0.60822	0.32493	6.889	8.21983	0.01243
LogK _{ow}	16	1.81847	0.00291	0.96502	0.92635	0.309	189.6654	1.57E-09
LogK _{oa}	16	1.73397	0.0064	0.97638	0.94997	0.554	285.8466	1.03E-10

Table 14. Statistical parameters for the linear QSPR model for CM₁(G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	62.22717	1.00759	0.98635	0.97095	17.715	502.3893	2.28E-12
MP	16	8.66644	0.46123	0.70942	0.4678	48.260	14.18502	0.00209
MW	16	48.95668	0.47995	0.98675	0.97179	8.312	517.733	1.86E-12
C _w ^{sat}	16	20.67553	-0.0507	-0.61522	0.3341	6.842	8.52577	0.01119
LogK _{ow}	16	1.5809	0.01089	0.97259	0.94208	0.274	244.9742	2.90E-10
LogK _{oa}	16	1.21694	0.02395	0.98333	0.96458	0.466	409.4999	9.17E-12

Table 15. Statistical parameters for the linear QSPR model for CM₂(G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	110.81052	0.19644	0.9721	0.94105	25.237	240.4368	3.28E-10
MP	16	35.99518	0.08653	0.6728	0.41357	50.659	11.5783	0.00429
MW	16	72.36874	0.09339	0.97062	0.93797	12.326	227.8054	4.69E-10
C _w ^{sat}	16	18.01663	-0.00974	-0.59756	0.31115	6.959	7.7754	0.01451
LogK _{ow}	16	2.12171	0.00211	0.95382	0.90333	0.354	141.1611	1.06E-08
LogK _{oa}	16	2.39454	0.00466	0.966	0.92838	0.663	195.4423	1.29E-09

Table 16. Statistical parameters for the linear QSPR model for CF(G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	108.99426	0.09592	0.97263	0.94215	25.001	245.2727	2.87E-10
MP	16	35.17482	0.04226	0.67327	0.41424	50.630	11.6079	0.00425
MW	16	71.52024	0.0456	0.97104	0.93885	12.239	231.2906	4.24E-10
C _w ^{sat}	16	18.12919	-0.00476	-0.59879	0.31274	6.951	7.82573	0.01425
LogK _{ow}	16	2.10231	0.00103	0.9543	0.9043	0.352	142.7469	9.89E-09
LogK _{oa}	16	2.3519	0.00227	0.96647	0.92935	0.659	198.3266	1.17E-09

Table 17. Statistical parameters for the linear QSPR model for CF*(G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	127.43673	0.02539	0.96458	0.92545	28.379	187.2128	1.71E-09
MP	16	45.48722	0.01099	0.65578	0.38934	51.695	10.56367	0.00581
MW	16	80.41593	0.01206	0.96208	0.92028	13.974	174.1478	2.74E-09
C _w ^{sat}	16	17.11549	-0.00125	-0.58964	0.30108	7.010	7.4616	0.01622
LogK _{ow}	16	2.3076	2.72E-04	0.94421	0.88378	0.388	115.071	3.89E-08
LogK _{oa}	16	2.79926	6.01E-04	0.95697	0.90978	0.744	152.2667	6.53E-09

Table 18. Statistical parameters for the linear QSPR model for CHM₁(G).

Physical properties	N	<i>V</i>	μ	R	Adj. R square	SE	F	Sig. F
BP	16	109.88851	0.04853	0.97237	0.94161	25.115	242.9077	3.06E-10
MP	16	35.57879	0.02138	0.67304	0.41391	50.644	11.59349	0.00427

Physical properties	N	V	μ	R	Adj. R square	SE	F	Sig. F
MW	16	71.93797	0.02307	0.97084	0.93842	12.281	229.5983	4.45E-10
C _w ^{sat}	16	18.0737	-0.00241	-0.59819	0.31196	6.955	7.80095	0.01438
LogK _{ow}	16	2.11186	5.22E-04	0.95407	0.90383	0.353	141.976	1.02E-08
LogK _{oa}	16	2.37289	0.00115	0.96624	0.92888	0.661	196.9229	1.23E-09

Table 19. Statistical parameters for the linear QSPR model for CHM₂(G).

Physical properties	N	V	μ	R	Adj. R square	SE	F	Sig. F
BP	16	183.03586	0.00168	0.934	0.86324	38.438	95.68194	1.23E-07
MP	16	74.62548	6.88E-04	0.60149	0.3162	54.703	7.93618	0.01371
MW	16	107.12101	7.95E-04	0.92892	0.8531	18.968	88.11357	2.03E-07
C _w ^{sat}	16	14.07797	-8.06E-05	-0.55554	0.25925	7.216	6.24965	0.02547
LogK _{ow}	16	2.92047	1.79E-05	0.90803	0.81199	0.494	65.78193	1.17E-06
LogK _{oa}	16	4.13969	3.95E-05	0.92223	0.83982	0.992	79.64704	3.75E-07

n = number of pairs, v= intercept, μ= slope, R= correlation coefficient, adj.R square= adjusted R square, SE= standard error, F= F-test, Sig F=significance F(<0.05, consistent).

3.2. Level of predictability.

Mathematically, TIs with higher |R| values (greater than 0.8) are beneficial for the prediction of the physicochemical properties of chemical compounds. Interestingly, All the proposed TIs strongly correlate with BP, MW, LogK_{ow}, and LogK_{oa} of priority PAHs (as highlighted in Table 20). The computed values of TIs for various priority PAHs are displayed in Table 21.

Table 20. Correlation of TIs with physicochemical properties of priority PAHs.

	CM ₁ * _(G)	CM ₁ _(G)	CM ₂ _(G)	CF(G)	CF*(G)	CHM ₁ _(G)	CHM ₂ _(G)
BP	0.98080	0.98635	0.97210	0.97263	0.96458	0.97237	0.93400
MP	0.69361	0.70942	0.67280	0.67327	0.65578	0.67304	0.60149
MW	0.98032	0.98675	0.97062	0.97104	0.96208	0.97084	0.92892
C _w ^{sat}	-0.60822	-0.61522	-0.59756	-0.59880	-0.58964	-0.59820	-0.55554
LogK _{ow}	0.965018	0.97260	0.95382	0.95430	0.94421	0.95407	0.90803
LogK _{oa}	0.97638	0.98333	0.966	0.96647	0.95697	0.96624	0.92223

Table 21. TIs of priority PAHs.

PAHs / TI	CM ₁ * _(G)	CM ₁ _(G)	CM ₂ _(G)	CF(G)	CF*(G)	CHM ₁ _(G)	CHM ₂ _(G)
Naphthalene	540	164	620	1280	4236	2520	39248
Acenaphthylene	810	234	997	2064	7218	4058	82057
Acenaphthene	810	234	997	2064	7218	4058	82057
Fluorene	846	246	1035	2134	7410	4204	82947
Phenanthrene	880	258	1064	2202	7584	4330	82348
Anthracene	868	256	1040	2136	7260	4216	74848
Fluoranthene	1166	330	1483	3060	10974	6026	138567
Pyrene	1152	328	1450	2992	10632	5892	129394
Benz[a] Anthracene	1208	350	1485	3058	10608	6028	118263
Chrysene	1220	352	1509	3124	10932	6142	125889
Benzo[b] Fluoranthene	1508	424	1937	3986	14364	7860	184801
Benzo[k] Fluoranthene	1494	422	1905	3916	13998	7726	174797
Benzo[a]pyrene	1494	422	1903	3918	14022	7724	175313
Dibenzo[a,h] Anthracene	1548	444	1930	3980	13956	7840	161678

PAHs / TI	CM ₁ * (G)	CM ₁ (G)	CM ₂ (G)	CF(G)	CF*(G)	CHM ₁ (G)	CHM ₂ (G)
Indeno [1,2,3-c,d] pyrene	1780	494	2326	4776	17412	9428	233028
Benzo[g,h,i] Perylene	1780	494	2326	4778	17436	9430	234614

3.3. Graphical analysis.

Though all the suggested indices exhibit a better correlation towards BP, MW, LogK_{ow}, and LogK_{oa} of priority PAHs, the prediction capacity of CM₁(G) was found to be slightly higher than others (as observed in Table 20). Figure 3 reveals pleasantly enough predictive behavior of the CM₁(G) index.

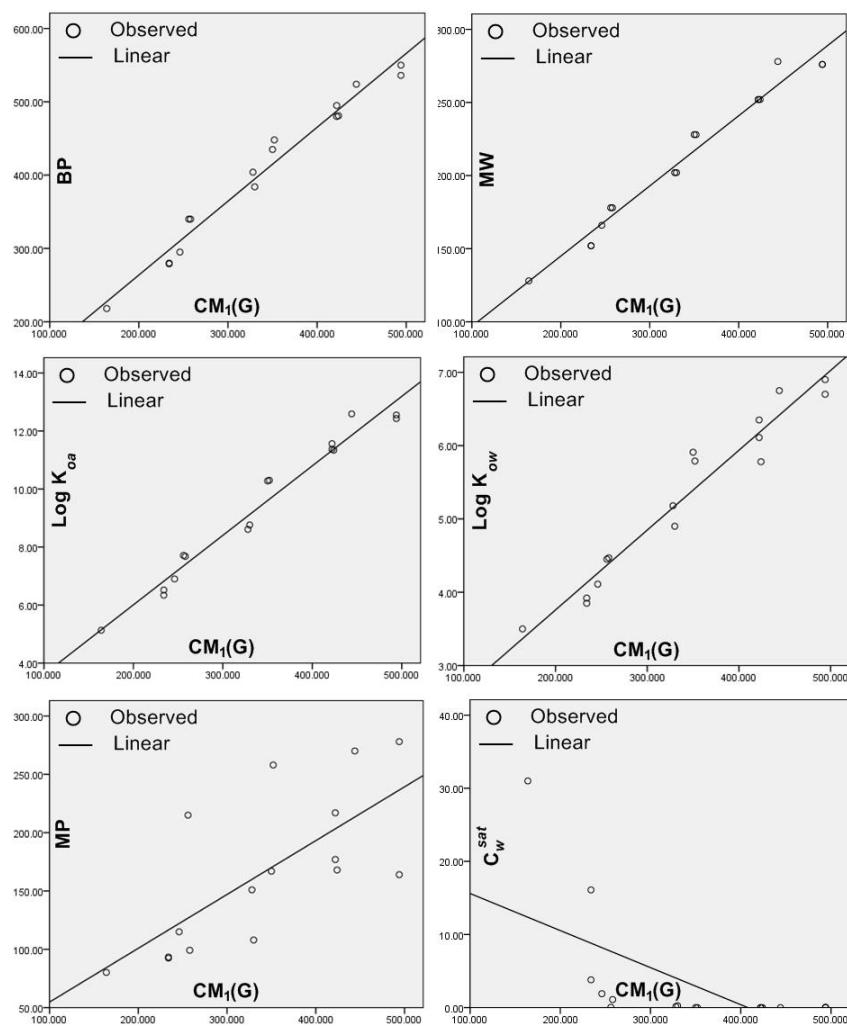


Figure 3. Correlations of physicochemical properties of priority PAHs with CM₁(G).

4. Conclusions

We successfully studied some closed neighborhood TIs. The chemical significance of the indices has been demonstrated with the physicochemical properties of priority PAHs by employing a linear regression model. The findings disclosed that the reported indices strongly correlate with BP, LogK_{ow} MW, and LogK_{oa} of priority PAHs. Moreover, the indices disclose excellent degeneracy properties (exactly 1) to different structural formulae of octane isomers, which indicates their supremacy in discriminating power. From a practical perspective, the

discussed TIs are very useful in QSPR research of novel nanomaterials. They have a wide range of possible applications in the fields of pharmacopeia as well as toxicology.

Funding

This research received no external funding.

Acknowledgments

This research has no acknowledgment.

Conflicts of Interest

There are no conflicts of interest.

References

1. Bondy, J.R.; Murty, U.S.R. *Graph theory*. 1st ed.; Springer; New York; USA, 2008; pp. 1–655, <https://www.zib.de/groetschel/teaching/WS1314/BondyMurtyGTWA.pdf>,
2. Trinajstić, N. *Chemical Graph theory*. 2nd ed.; CRC Press; New York, USA, 1992; pp. 1–352; <https://doi.org/10.1201/9781315139111>.
3. Gutman, I. Degree-based topological indices. *Croatica chemica acta* **2013**, *86*, 351–361, <https://doi.org/10.5562/cca2294>.
4. Randic, M. On Characterization of Molecular Branching. *Journal of the American Chemical Society* **1975**, *97*, 6609–6615, <https://doi.org/10.1021/ja00856a001>.
5. Li, X.; Gutman, I. Mathematical aspects of Randic-type molecular structure descriptors. 1st ed.; University of Kragujevac; Serbia, 2006.
6. Gutman, I.; Trinajstić, N. Graph Theory and Molecular Orbitals. Total π -Electron Energy of Alternant Hydrocarbons. *Chemical Physics Letters* **1972**, *17*, 535–538, [https://doi.org/10.1016/0009-2614\(72\)85099-1](https://doi.org/10.1016/0009-2614(72)85099-1).
7. Shirdel, G.H.; Rezapour, H.; Sayadi, A.M. The Hyper-Zagreb Index of Graph Operations. *Iranian Journal of Mathematical Chemistry* **2013**, *4*, 213–220, <https://doi.org/10.22052/IJMC.2013.5294>.
8. Furtula, B.; Gutman, I. A forgotten topological index. *J. Math. Chem.* **2015**, *53*, 184–1190, <https://doi.org/10.1007/s10910-015-0480-z>.
9. Srinivasan, M.S. A Computational Approach on Acetaminophen Drug using Degree-Based Topological Indices and M-Polynomials. *Biointerface Research in Applied Chemistry*, **2022**, *12*, 7249–7266, <https://doi.org/10.33263/BRIAC126.72497266>.
10. Duraisami, M.S.; Parasuraman, K. Computational analysis of some degree based topological indices of cubic structured tungsten trioxide $[l,m,n]$ nanomultilayer. *Nanosystems: Phys. Chem. Math.* **2020**, *11*, 501–509, <https://doi.org/10.17586/2220-8054-2020-11-5-501-509>.
11. Duraisami, M.S.; Anburaj, D.B.; Parasuraman, K. Computing vertex degree-based multiplicative version of topological indices for tungsten trioxide nano multilayer structure in nanotherapeutic anti-cancer activity. *Biointerface Research in Applied Chemistry* **2022**, *12*, 2275–2284, <https://doi.org/10.33263/BRIAC122.22752284>.
12. Gutman, I. Geometric approach to degree based topological indices: Sombor indices. *MATCH Common, Math. Comput. Chem.* **2021**, *86*, 11–16, https://match.pmf.kg.ac.rs/electronic_versions/Match86/n1/match86n1_11-16.pdf.
13. Das, K.C.; Çevik, A.S.; Cangul, I.N.; Shang, Y. On Sombor Index. *Symmetry* **2021**, *13*, <https://doi.org/10.3390/sym13010140>.
14. Gutman, I. Some basic properties of Sombor indices. *Open Journal of Discrete Applied Mathematics* **2021**, *4*, 1–3, <https://doi.org/10.30538/psrp-odam2021.0047>.
15. Kulli, V.R. Sombor indices of certain graph operators. *International Journal of Engineering Sciences and Research Technology* **2021**, *10*, 127–134, <https://doi.org/10.29121/ijesrt.v10.i1.2021.12>.
16. Kulli, V.R. Multiplicative Sombor indices of certain nanotubes. *International Journal of Mathematical Archive* **2021**, *12*.
17. Kulli, V.; Gutman, I. Computation of Sombor Indices of Certain Networks. *International Journal of Applied Chemistry* **2021**, *8*, 1–5, <https://doi.org/10.14445/23939133/IJAC-V8I1P101>.

18. Milovanovic, I.; Milovanovic, E.; Matejic, M. On some mathematical properties of Sombor indices. *Bull. Int. Math. Virtual Inst.* **2021**, *11*, 341–353, <https://doi.org/10.7251/BIMVI2102341M>.
19. Redzepovic, I. Chemical applicability of Sombor indices. *J. Serb. Chem. Soc.* **2021**, *86*, 445-457, <https://doi.org/10.2298/JSC201215006R>.
20. Reti, T.; Doslic, T.; Ali, A. On the Sombor index of graphs. *Contributions of Mathematics* **2021**, *3*, 11–18, <https://doi.org/10.47443/cm.2021.0006>.
21. Mumtaz, H.B.; Javaid, M.; Awais, H.M.; Bonyah, E. Topological Indices of Pent-Heptagonal Nanosheets via M-Polynomials. *Journal of Mathematics* **2021**, *2021*, 1–13, <https://doi.org/10.1155/2021/4863993>.
22. Anwar Saleh.; Sophia Shalini, G. B.; Dhananjayamurthy, B. V. The Reduced Neighborhood Topological Indices and RNM – Polynomial for the Treatment of COVID-19. *Biointerface Research in Applied Chemistry* **2021**, *11*, 11817–11832, <https://doi.org/10.33263/BRIAC114.1181711832>.
23. Mondal, S.; De, N.; Siddiqui, M. K.; Pal, A. Topological Properties of Para-Line Graph of Some Convex Polytopes Using Neighborhood M-Polynomial. *Biointerface Research in Applied Chemistry* **2021**, *11*, 9915–9927, <https://doi.org/10.33263/BRIAC113.99159927>.
24. Mondal, S.; De, N.; Pal, A. On Neighbourhood Zagreb index of product graphs. *J. Mol. Struct.* **2021**, *1223*, 129210, <https://doi.org/10.1016/j.molstruc.2020.129210>.
25. Shanmukha, M.C.; Usha, A.; Basavarajappa, N.S.; Shilpa, K.C. Graph entropies of porous Graphene using topological indices. *Computational and Theoretical Chemistry* **2021**, *1197*, 113142, <https://doi.org/10.1016/j.comptc.2021.113142>.
26. Mondal, S.; De, N.; Pal, A. QSPR analysis of some novel neighborhood degree based topological descriptors, *Complex & Intelligent Systems* **2021**, *7*, 977–966, <https://doi.org/10.1007/s40747-020-00262-0>.
27. Kulli, V.R. Neighbourhood Dakshayani indices. *International Journal of Mathematical Archive* **2019**, *10*, 23–31.
28. Srinivasan, M.S.; Helen .R. Computing some Novel Closed Neighborhood Degree- Based Topological Indices of Graphene Structures. *Biointerface Research in Applied Chemistry* **2023**, *13*, 1–15, <https://doi.org/10.33263/BRIAC131.092>.
29. Srinivasan, M.S.; Helen, R. On bounds for certain closed neighbourhood topological indices. *Advances and Applications in Mathematical Sciences* **2022**, (Accepted).
30. Randić, M.; Trinajstić, N. In search for graph invariants of chemical interest. *J. Mol. Struct.* **1993**, *300*, 551–571, [https://doi.org/10.1016/0022-2860\(93\)87047-D](https://doi.org/10.1016/0022-2860(93)87047-D).
31. Konstantinova, E.V. The Discrimination Ability of Some Topological and Information Distance Indices for Graphs of Unbranched Hexagonal Systems. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 54–57, <https://doi.org/10.1021/ci9502461>.
32. Bonchev, D.; Mekenyan, O.; Trinajstić, N. Isomer discrimination by topological information approach. *J. Comput. Chem.* **1981**, *2*, 127–148, <https://doi.org/10.1002/jcc.540020202>.
33. Fouchecourt, M.O.; Arnold, M.; Berny, P.; Videmann, B.; Rether, B.; Riviere, J.L. Assessment of the bioavailability of PAHs in rats exposed to a polluted soil by natural routes: induction of EROD activity and DNA adducts and PAH burden in both liver and lung. *Environmental Research Section A* **1999**, *80*, 330–339, <https://doi.org/10.1006/enrs.1998.3932>.
34. WHO, International Agency for Research on Cancer, 2006.
35. Tamilarasi, C.; Simon Raj, F. QSPR Analysis of Novel Indices with Priority Polycyclic Aromatic Hydrocarbons (PAHs). *Turkish Journal of Computer and Mathematics Education* **2021**, *12*, 3992–3999, <https://turcomat.org/index.php/turkbilmat/article/view/5110>.