

## Novel molecular hybrid geometric-harmonic-Zagreb degree based descriptors and their efficacy in QSPR studies of polycyclic aromatic hydrocarbons

M. Arockiaraj, D. Paul, J. Clement, S. Tigga, K. Jacob & K. Balasubramanian

To cite this article: M. Arockiaraj, D. Paul, J. Clement, S. Tigga, K. Jacob & K. Balasubramanian (2023): Novel molecular hybrid geometric-harmonic-Zagreb degree based descriptors and their efficacy in QSPR studies of polycyclic aromatic hydrocarbons, SAR and QSAR in Environmental Research, DOI: [10.1080/1062936X.2023.2239149](https://doi.org/10.1080/1062936X.2023.2239149)

To link to this article: <https://doi.org/10.1080/1062936X.2023.2239149>



Published online: 04 Aug 2023.



Submit your article to this journal



View related articles



View Crossmark data



# Novel molecular hybrid geometric-harmonic-Zagreb degree based descriptors and their efficacy in QSPR studies of polycyclic aromatic hydrocarbons

M. Arockiaraj <sup>a</sup>, D. Paul <sup>b</sup>, J. Clement <sup>c</sup>, S. Tigga<sup>a</sup>, K. Jacob <sup>c</sup>  
and K. Balasubramanian <sup>d</sup>

<sup>a</sup>Department of Mathematics, Loyola College, Chennai, India; <sup>b</sup>Department of Mathematics, Sri Sairam Institute of Technology, Chennai, India; <sup>c</sup>Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology, Vellore, India; <sup>d</sup>School of Molecular Sciences, Arizona State University, Tempe, AZ, USA

## ABSTRACT

The physicochemical characteristics of polycyclic aromatic compounds critical to environmental modelling such as octanol partition coefficients, solubility, lipophilicity, polarity and several equilibrium constants are functions of their underlying molecular structures, prompting the development of mathematical models to predict such characteristics for which experimental results are difficult to obtain. We propose twelve novel descriptors derived from geometric, harmonic and Zagreb degree-based descriptors and then test the effectiveness of these descriptors on a data set consisting of 55 benzenoid hydrocarbons of environmental importance. Our computations show that the proposed descriptors have a good linear correlation and predictive power when compared to the degree and distance type descriptors. We have also derived the QSPR expressions for four properties of a large series of polycyclic aromatics arising from circumscribing coronenes and show that a scaling factor can be deduced to derive physicochemical properties of such series up to 2D graphene sheets.

## ARTICLE HISTORY

Received 12 June 2023

Accepted 16 July 2023

## KEYWORDS

Hybrid geometric-harmonic-Zagreb descriptors; polycyclic benzenoid hydrocarbons; topological descriptors; QSPR of thermodynamic properties; coronene series

## Introduction

Topological descriptors are mathematical functions of a chemical structure that encode the molecular structural features, and enable to collect, annotate, retrieve, compare, and predict their properties. The structural characterization of molecules with the help of topological descriptors is crucial because the physicochemical properties of molecular structures can be predicted. In molecular modelling, the topological descriptors could be used to correlate with the chemical, physical, biological, medicinal, toxicological and pharmacological characteristics and biological activities of molecular structures. They are extensively used in QSAR and QSPR studies to investigate many properties such as melting point, boiling point, toxicity parameters, chromatographic retention index, vapour pressure, octanol partitions coefficients, and so forth.

In recent years, several studies have been carried out on topological descriptors as they find extensive applications in the development of QSAR and QSPR studies [1–12]. Consequently, there is renewed attention for the advancement and development of new topological descriptors which will further extend the potentiality of QSAR/QSPR approaches and their applications [13–16]. QSAR/QSPR models have been employed in a number of areas to characterize molecular structures, biological networks, and information science to correlate structures or networks with their properties for which the degree and distance measures from graphs have a significant impact. In the context of drug discovery guest-host interactions play critical roles, and thus rapid computations of their properties could result in significant cost saving. These techniques largely depend on quantitative measures, some of which are based on graph distance degree measures that are commonly employed to analyse structural differences of molecular compounds [13,17]. These types of modelling are critical to environmental modelling of pollutants and petroleum products as in such emissions typically a mixture of large number of polycyclic aromatic compounds are found. For example, in a major accidental fire involving rubber tires numerous polycyclic aromatic compounds were found in the emission [18]. Consequently, mathematical modelling techniques are highly desirable for rapid assessment of the potential hazards to the environment so that adequate remediation measures can be taken.

Topological indices and their applicability to circumcised donut benzenoid systems, kekulenes and related drugs have also been considered [19]. The graph of a molecule is denoted by  $G$  in which all the atoms grouped in a set  $V(G)$ , called the vertex set, and the bonds denoted by a set  $E(G)$ , the set of edges of  $G$ . For a vertex  $u$ , the degree  $d_u$  is the number of edges incident to  $u$ . Two types of vertex-degree based Zagreb indices have been well studied due to their diverse applications, and an analogous of the first type Zagreb index is called the forgotten index [20]. The first Zagreb index computes the square of vertex degrees by taking the sum over all vertices, whereas the forgotten index computes the sum of cubes of vertex degrees over all vertices. The additive combination of the forgotten and the second Zagreb indices in the ratio of 1: 2 result in a hyper-Zagreb index and it has received considerable attention [21,22] over the years. In the series, we now take the ratio of 1: 1, and propose two types of new indices called the bi-Zagreb and tri-Zagreb indices, which are stated below:

$$BM(G) = \sum_{u \in V(G)} d_u^2 + \sum_{uv \in E(G)} d_u d_v = \sum_{uv \in E(G)} (d_u + d_v + d_u d_v)$$

$$TM(G) = \sum_{u \in V(G)} d_u^3 + \sum_{uv \in E(G)} d_u d_v = \sum_{uv \in E(G)} (d_u^2 + d_v^2 + d_u d_v)$$

In another perspective, the combination of two topological indices in fractional ratio was introduced in [23] and here we introduce ten new types of topological indices by taking various combinations of geometric, harmonic, bi-Zagreb and tri-Zagreb indices and demonstrate their efficacy.

$$GH(G) = \sum_{uv \in E(G)} \frac{\sqrt{d_u d_v} (d_u + d_v)}{2} \quad (\text{Geometric} - \text{Harmonic})$$

$$GBM(G) = \sum_{uv \in E(G)} \frac{\sqrt{d_u d_v}}{d_u + d_v + d_u d_v} \text{ (Geometric - Bi Zagreb)}$$

$$GTM(G) = \sum_{uv \in E(G)} \frac{\sqrt{d_u d_v}}{d_u^2 + d_v^2 + d_u d_v} \text{ (Geometric - Tri Zagreb)}$$

$$HG(G) = \sum_{uv \in E(G)} \frac{2}{\sqrt{d_u d_v}(d_u + d_v)} \text{ (Harmonic - Geometric)}$$

$$HBM(G) = \sum_{uv \in E(G)} \frac{2}{(d_u + d_v + d_u d_v)(d_u + d_v)} \text{ (Harmonic - Bi Zagreb)}$$

$$HTM(G) = \sum_{uv \in E(G)} \frac{2}{(d_u^2 + d_v^2 + d_u d_v)(d_u + d_v)} \text{ (Harmonic - Tri Zagreb)}$$

$$BMG(G) = \sum_{uv \in E(G)} \frac{d_u + d_v + d_u d_v}{\sqrt{d_u d_v}} \text{ (Bi Zagreb - Geometric)}$$

$$BMH(G) = \sum_{uv \in E(G)} \frac{(d_u + d_v + d_u d_v)(d_u + d_v)}{2} \text{ (Bi Zagreb - Harmonic)}$$

$$TMG(G) = \sum_{uv \in E(G)} \frac{d_u^2 + d_v^2 + d_u d_v}{\sqrt{d_u d_v}} \text{ (Tri Zagreb - Geometric)}$$

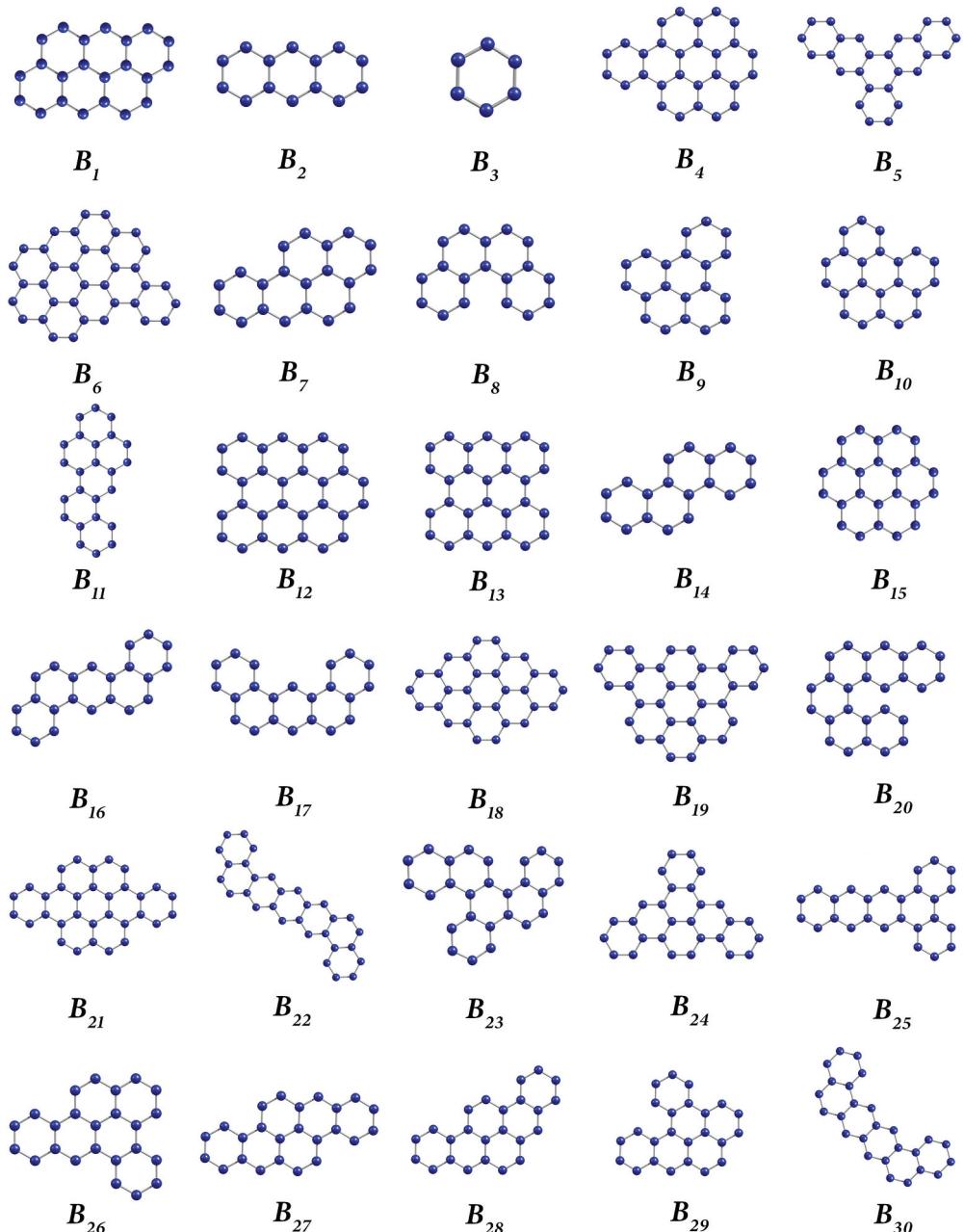
$$TMH(G) = \sum_{uv \in E(G)} \frac{(d_u^2 + d_v^2 + d_u d_v)(d_u + d_v)}{2} \text{ (Tri Zagreb - Harmonic)}$$

We study the effectiveness of the proposed topological descriptors compared to the existing degree and distance types of topological descriptors using a data set consisting of 55 benzenoid hydrocarbons and present the associated linear regression equations for several physicochemical properties along with their implication in predicting physicochemical properties of hydrocarbons.

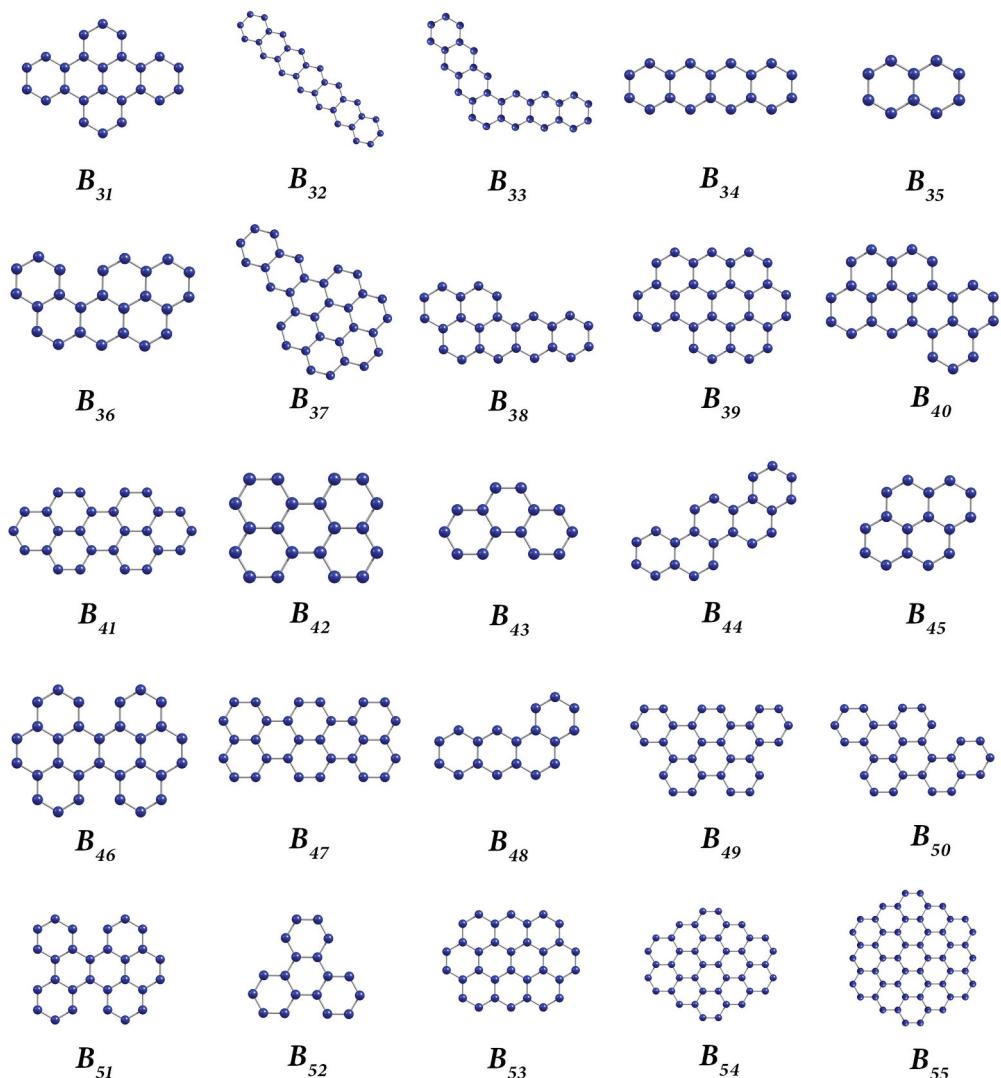
## Computation of hybrid geometric-harmonic-Zagreb descriptors for benzenoid hydrocarbons

In this section, we compute the topological descriptors derived from taking the ratio of geometric, harmonic, bi-Zagreb and tri-Zagreb indices by considering a data set of 55 benzenoid hydrocarbons as shown in Figures 1 and 2. For this purpose, we partition the bonds of hydrocarbons using the notation given below:

$$d_{k,j} = |\{uv \in E(G) : d_u = k, d_v = j\}|$$



**Figure 1.** Benzenoid hydrocarbons: B<sub>1</sub>-Anthanthrene, B<sub>2</sub>-Anthracene, B<sub>3</sub>-Benzene, B<sub>4</sub>-Benzo[a]coronene, B<sub>5</sub>-Benzo[h]pentaphene, B<sub>6</sub>-Benzo[a]ovalene, B<sub>7</sub>-Benzo[a]pyrene, B<sub>8</sub>-Benzo[c]phenanthrene, B<sub>9</sub>-Benzo[e]pyrene, B<sub>10</sub>-Benzo[ghi]perylene, B<sub>11</sub>-Benzo[pqr]picene, B<sub>12</sub>-Benzobisanthene, B<sub>13</sub>-Bisanthene, B<sub>14</sub>-Chrysene, B<sub>15</sub>-Coronene, B<sub>16</sub>-Dibenz[a,h]anthracene, B<sub>17</sub>-Dibenz[a,j]anthracene, B<sub>18</sub>-Dibenz[Bc,Kl]Coronene, B<sub>19</sub>-Dibenzo[a,g]coronene, B<sub>20</sub>-Dibenzo[a,ghi]perylene, B<sub>21</sub>-Dibenzo[a,j]coronene, B<sub>22</sub>-Dibenzo[a,l]pentacene, B<sub>23</sub>-Dibenzo[c,p]chrysene, B<sub>24</sub>-Dibenzo[h,rst]Pentaphene, B<sub>25</sub>-Dibenzo[a,c]tetracene, B<sub>26</sub>-Dibenzo[a,e]pyrene, B<sub>27</sub>-Dibenzo[a,h]pyrene, B<sub>28</sub>-Dibenzo[a,i]pyrene, B<sub>29</sub>-Dibenzo[a,l]pyrene, B<sub>30</sub>-Dibenzo[a,l]tetracene (B<sub>1</sub>, B<sub>2</sub>, B<sub>7</sub>-B<sub>10</sub>, B<sub>14</sub>-B<sub>17</sub> and B<sub>26</sub>-B<sub>29</sub> are reproduced from [11]).



**Figure 2.** Benzenoid hydrocarbons continuation:  $B_{31}$ -Dibenzo[e,l]pyrene,  $B_{32}$ -Heptacene,  $B_{33}$ -Heptaphene,  $B_{34}$ -Naphthacene,  $B_{35}$ -Naphthalene,  $B_{36}$ -Naphtho[1,2-a]pyrene,  $B_{37}$ -Naphtho[2,3-a]coronene,  $B_{38}$ -Naphtho[2,3-a]pyrene,  $B_{39}$ -Naphtho[5,4,3-abc]coronene,  $B_{40}$ -Naphtho[8,1,2-bcd]perylene,  $B_{41}$ -Peropyrene,  $B_{42}$ -Perylene,  $B_{43}$ -Phenanthrene,  $B_{44}$ -Picene,  $B_{45}$ -Pyrene,  $B_{46}$ -Pyreno[4,5-e]pyrene,  $B_{47}$ -Terylene,  $B_{48}$ -Tetraphene,  $B_{49}$ -Tribenzo[b,n,pqr]perylene,  $B_{50}$ -Tribenzo[a,fg,op]tetracene,  $B_{51}$ -Tribenzo[f,ij,no]tetraphene,  $B_{52}$ -Triphenylene,  $B_{53}$ -Ovalene,  $B_{54}$ -Circumpyrene,  $B_{55}$ -Circumcoronene ( $B_{34}, B_{35}, B_{42}, B_{45}, B_{48}$  and  $B_{52}$  are reproduced from [11]).

Graph polynomials are analytical expressions derived from a graph to build various algebraic models to determine graph information [24,25]. The  $M$ -polynomial is one such expression introduced to quantify the degree-based topological indices of chemical

structures using calculus operators [26]. A primary advantage of the  $M$ -polynomial technique is that several topological indices can be computed simultaneously from the expression. The  $M$ -polynomials have been employed extensively in deriving QSAR expressions for toxicity and other properties of antiepileptic and  $\beta$ -dione compounds [27,28].

The  $M$ -polynomial of a graph  $G$  is formally defined as follows:

$$M(G; x, y) = \sum_{k \leq j} d_{kj} x^k y^j$$

The benzenoid hydrocarbons shown in Figures 1, 2 and 3 constitute the data set for the current study and they have partition classes  $d_{2,2}$ ,  $d_{2,3}$  and  $d_{3,3}$  as given in Table 1, thus the above-described  $M$ -polynomial is simplified as follows:

$$M(G; x, y) = d_{2,2}x^2y^2 + d_{2,3}x^2y^3 + d_{3,3}x^3y^3$$

The  $M$ -polynomials of  $B_i$ ,  $1 \leq i \leq 55$ , are presented in Table 1, the numerical quantities for bi-Zagreb, tri-Zagreb, geometric-harmonic, geometric-bi-Zagreb, geometric-tri-Zagreb, harmonic-geometric, harmonic-bi-Zagreb, harmonic-tri-Zagreb, bi-Zagreb-geometric, bi-Zagreb-harmonic, tri-Zagreb-geometric and tri-Zagreb-harmonic descriptors are obtained by constructing suitable differential and integral operators [29,30] with the aid of the MATLAB interface because the manual procedure is tedious to compute the descriptor values which are presented in Table 2. For instance, we show the steps to obtain the  $BMG$  index for anthracene compound by defining its suitable operator as  $D_x^{\frac{1}{2}}S_y^{\frac{1}{2}} + D_y^{\frac{1}{2}}S_x^{\frac{1}{2}} + D_x^{\frac{1}{2}}D_y^{\frac{1}{2}}|_{x=y=1}$ , where

$$D_x^{\frac{1}{2}}(f(x, y)) = \sqrt{x \frac{\partial f(x, y)}{\partial x}} \sqrt{f(x, y)},$$

$$D_y^{\frac{1}{2}}(f(x, y)) = \sqrt{y \frac{\partial f(x, y)}{\partial y}} \sqrt{f(x, y)},$$

$$S_x^{\frac{1}{2}}(f(x, y)) = \sqrt{\int_0^x \frac{f(t, y)}{t} dt} \sqrt{f(x, y)},$$

$$S_y^{\frac{1}{2}}(f(x, y)) = \sqrt{\int_0^x \frac{f(x, t)}{t} dt} \sqrt{f(x, y)}.$$

From Table 1, we have  $M(B_2; x, y) = f(x, y) = 6x^2y^2 + 8x^2y^3 + 2x^3y^3$  and the operators can be applied as follows:

$$S_y^{\frac{1}{2}}f(x, y) = \frac{6x^2y^2}{\sqrt{2}} + \frac{8x^2y^3}{\sqrt{3}} + \frac{2x^3y^3}{\sqrt{3}}$$

$$D_x^{\frac{1}{2}}S_y^{\frac{1}{2}}f(x, y) = \frac{6\sqrt{2}x^2y^2}{\sqrt{2}} + \frac{8\sqrt{2}x^2y^3}{\sqrt{3}} + \frac{2\sqrt{3}x^3y^3}{\sqrt{3}}$$

Similarly, we have

**Table 1.** Bond partitions of 55 benzenoid hydrocarbons and their *M*-polynomials.

Compounds	Atoms	Bonds	$d_{2,2}$	$d_{2,3}$	$d_{3,3}$	$M(G; x, y)$
Anthanthrene	22	27	6	12	9	$6x^2y^2 + 12x^2y^3 + 9x^3y^3$
Anthracene	14	16	6	8	2	$6x^2y^2 + 8x^2y^3 + 2x^3y^3$
Benzene	6	6	6	0	0	$6x^2y^2$
Benzo[a]coronene	28	35	8	12	15	$8x^2y^2 + 12x^2y^3 + 15x^3y^3$
Benzo[h]pentaphene	26	31	9	14	8	$9x^2y^2 + 14x^2y^3 + 8x^3y^3$
Benzo[a]ovalene	36	46	8	16	22	$8x^2y^2 + 16x^2y^3 + 22x^3y^3$
Benzo[a]pyrene	20	24	7	10	7	$7x^2y^2 + 10x^2y^3 + 7x^3y^3$
Benzo[c]phenanthrene	18	21	8	8	5	$8x^2y^2 + 8x^2y^3 + 5x^3y^3$
Benzo[e]pyrene	20	24	8	8	8	$8x^2y^2 + 8x^2y^3 + 8x^3y^3$
Benzo[ghi]perylene	22	27	7	10	10	$7x^2y^2 + 10x^2y^3 + 10x^3y^3$
Benzo[pqr]picene	24	29	8	12	9	$8x^2y^2 + 12x^2y^3 + 9x^3y^3$
Benzobisanthene	30	38	7	14	17	$7x^2y^2 + 14x^2y^3 + 17x^3y^3$
Bisanthene	28	35	8	12	15	$8x^2y^2 + 12x^2y^3 + 15x^3y^3$
Chrysene	18	21	8	8	5	$8x^2y^2 + 8x^2y^3 + 5x^3y^3$
Coronene	24	30	6	12	12	$6x^2y^2 + 12x^2y^3 + 12x^3y^3$
Dibenz[a,h]anthracene	22	26	8	12	6	$6x^2y^2 + 12x^2y^3 + 6x^3y^3$
Dibenz[a,j]anthracene	22	26	8	12	6	$8x^2y^2 + 12x^2y^3 + 6x^3y^3$
Dibenz[Bc,Kl]Coronene	30	38	6	16	16	$6x^2y^2 + 16x^2y^3 + 16x^3y^3$
Dibenzo[a,g]coronene	32	40	10	12	18	$10x^2y^2 + 12x^2y^3 + 18x^3y^3$
Dibenzo[ghi]perylene	26	32	7	13	12	$7x^2y^2 + 13x^2y^3 + 12x^3y^3$
Dibenzo[a,j]coronene	32	40	10	12	18	$10x^2y^2 + 12x^2y^3 + 18x^3y^3$
Dibenzo[a,l]pentacene	30	36	8	20	8	$8x^2y^2 + 20x^2y^3 + 8x^3y^3$
Dibenzo[c,p]chrysene	26	31	11	11	9	$11x^2y^2 + 11x^2y^3 + 9x^3y^3$
Dibenzo[h,rst]Pentaphene	28	34	10	12	12	$10x^2y^2 + 12x^2y^3 + 12x^3y^3$
Dibenzo[a,c]tetracene	26	31	9	14	8	$9x^2y^2 + 14x^2y^3 + 8x^3y^3$
Dibenzo[a,e]pyrene	24	29	9	10	10	$9x^2y^2 + 10x^2y^3 + 10x^3y^3$
Dibenzo[a,h]pyrene	24	29	8	12	9	$8x^2y^2 + 12x^2y^3 + 9x^3y^3$
Dibenzo[a,i]pyrene	24	29	8	12	9	$8x^2y^2 + 12x^2y^3 + 9x^3y^3$
Dibenzo[a,l]pyrene	24	29	9	10	10	$9x^2y^2 + 10x^2y^3 + 10x^3y^3$
Dibenzo[a,l]tetracene	26	31	8	16	7	$8x^2y^2 + 16x^2y^3 + 7x^3y^3$
Dibenzo[e,l]pyrene	24	29	10	8	11	$10x^2y^2 + 8x^2y^3 + 11x^3y^3$
Heptacene	30	36	6	24	6	$6x^2y^2 + 24x^2y^3 + 6x^3y^3$
Heptaphene	30	36	7	22	7	$7x^2y^2 + 22x^2y^3 + 7x^3y^3$
Naphthacene	18	21	6	12	3	$6x^2y^2 + 12x^2y^3 + 3x^3y^3$
Naphthalene	10	11	6	4	1	$6x^2y^2 + 4x^2y^3 + 1x^3y^3$
naphtho[1,2-a]pyrene	24	29	8	12	9	$8x^2y^2 + 12x^2y^3 + 9x^3y^3$
Naphtho[2,3-a]coronene	32	40	8	16	16	$8x^2y^2 + 16x^2y^3 + 16x^3y^3$
Naphtho[2,3-a]pyrene	24	29	7	14	8	$7x^2y^2 + 14x^2y^3 + 8x^3y^3$
Naphtho[5,4,3-abc]coronene	30	38	7	14	17	$7x^2y^2 + 14x^2y^3 + 17x^3y^3$
Naphtho[8,1,2-bcd]perylene	26	32	8	12	12	$8x^2y^2 + 12x^2y^3 + 12x^3y^3$
Peropyrene	26	32	8	12	12	$8x^2y^2 + 12x^2y^3 + 12x^3y^3$
Perylene	20	24	8	8	8	$8x^2y^2 + 8x^2y^3 + 8x^3y^3$
Phenanthrene	14	16	7	6	3	$7x^2y^2 + 6x^2y^3 + 3x^3y^3$
Picene	22	26	9	10	7	$9x^2y^2 + 10x^2y^3 + 7x^3y^3$
Pyrene	16	19	6	8	5	$6x^2y^2 + 8x^2y^3 + 5x^3y^3$
Pyreno[4,5-e]pyrene	30	37	10	12	15	$10x^2y^2 + 12x^2y^3 + 15x^3y^3$
Terylene	30	37	10	12	15	$10x^2y^2 + 12x^2y^3 + 15x^3y^3$
Tetraphene	18	21	7	10	4	$7x^2y^2 + 10x^2y^3 + 4x^3y^3$
Tribenzo[b,n,pqr]perylene	30	37	11	10	16	$11x^2y^2 + 10x^2y^3 + 16x^3y^3$
Tribenzo[b,n,pqr]perylene	30	37	11	10	16	$11x^2y^2 + 10x^2y^3 + 16x^3y^3$
Tribenzo[a,fg,op]tetracene	28	34	11	10	13	$11x^2y^2 + 10x^2y^3 + 13x^3y^3$
Tribenzo[f,ij,no]tetraphene	28	34	11	10	13	$11x^2y^2 + 10x^2y^3 + 13x^3y^3$
Triphenylene	18	21	9	6	6	$9x^2y^2 + 6x^2y^3 + 6x^3y^3$
Ovalene	32	41	6	16	19	$6x^2y^2 + 16x^2y^3 + 19x^3y^3$
Circumpyrene	42	55	6	20	29	$6x^2y^2 + 20x^2y^3 + 29x^3y^3$
Circumcoronene	54	72	6	24	42	$6x^2y^2 + 24x^2y^3 + 42x^3y^3$

$$D_y^{\frac{1}{2}} S_x^{\frac{1}{2}} f(x, y) = \frac{6\sqrt{2}x^2y^2}{\sqrt{2}} + \frac{8\sqrt{3}x^2y^3}{\sqrt{2}} + \frac{2\sqrt{3}x^3y^3}{\sqrt{3}}$$

$$D_y^{\frac{1}{2}} f(x, y) = 6\sqrt{2}x^2y^2 + 8\sqrt{3}x^2y^3 + 2\sqrt{3}x^3y^3$$

$$D_x^{\frac{1}{2}} D_y^{\frac{1}{2}} f(x, y) = 12x^2y^2 + 8\sqrt{2}\sqrt{3}x^2y^3 + 6x^3y^3$$

By combining the above equations, we get

$$\left[ D_x^{\frac{1}{2}} S_y^{\frac{1}{2}} + D_y^{\frac{1}{2}} S_x^{\frac{1}{2}} + D_x^{\frac{1}{2}} D_y^{\frac{1}{2}} \right] f(x, y) = (6 + 6 + 12)x^2y^2 + \left( \frac{8\sqrt{2}}{\sqrt{3}} + \frac{8\sqrt{3}}{\sqrt{2}} + 8\sqrt{2}\sqrt{3} \right) x^2y^3 + (2 + 2 + 6)x^3y^3$$

Substituting  $x = y = 1$ , we obtain  $BMG(B_2) = 69.9258$ . In a direct way, the topological descriptors can be computed by multiplying each class with its corresponding index function and taking the sum over all the classes. That is, for the above discussed index,  $BMG(B_2) = 6BMG(2, 2) + 8BMG(2, 3) + 2BMG(3, 3)$ , where  $BMG(2, 2) = \frac{8}{\sqrt{4}}$ ,  $BMG(2, 3) = \frac{11}{\sqrt{6}}$  and  $BMG(3, 3) = \frac{15}{\sqrt{9}}$ .

We have listed the experimental data for the 55 benzenoid hydrocarbons in [Table 3](#) for the physicochemical characteristics considered here, namely  $\log P$ , Kovats chromatographic retention index ( $R_I$ ), boiling point ( $BP$ ) and Pitzer's acentric factor ( $\omega$ ) [31–58].

## QSPR studies employing the proposed descriptors

QSAR and QSPR are mathematical models that use mathematical tools to correlate the structural information of chemical compounds with their biological activities and properties. They aim to establish quantitative relations between structures and biological activities through topological descriptors. The objective of these models is to identify and investigate the factors that determine the measured properties of chemical compounds. These models facilitate the ability to predict the properties from their structures. Recently, QSAR and QSPR models have received greater attention from the researchers due to their predictive ability, hence many studies have been made to obtain such models. In order to have higher reliability, these models should be robust and statistically significant and should have predictive capabilities with reasonable accuracies [59–61]. Furthermore, these models are very helpful because they reduce the time and expenditure of conducting experimental studies when dealing with a large number of compounds [45,46,62–66]. We now compute the correlation between the physicochemical characteristic values and degree-based descriptor values for the 55 benzenoid hydrocarbons. It is found that all the descriptors have a positive correlation, with the highest value being 0.9972. The scatter diagrams for topological descriptors with the highest correlation are shown in [Figure 3](#) while the complete correlation values are given in [Table 4](#). We now present the regression equations arising from these descriptors to predict the physicochemical characteristics of benzenoid hydrocarbons.

**Table 2.** Topological indices of benzenoid hydrocarbons.

Compounds	<i>B<sub>M</sub></i>	<i>TM</i>	<i>GH</i>	<i>GBM</i>	<i>GTM</i>	<i>HG</i>	<i>HBM</i>	<i>HTM</i>	<i>BMG</i>	<i>BMH</i>	<i>TMG</i>	<i>TMH</i>
Anthanthrene	315	543	178.4847	5.9722	3.5470	4.4596	1.0114	0.6137	122.8888	831	210.0806	1443
Anthracene	166	278	90.9898	3.6814	2.2536	3.0286	0.7104	0.4431	69.9258	406	116.0537	686
Benzene	48	72	24	1.5	1	1.5	0.375	0.25	24	96	36	144
Benzo[ <i>a</i> ]coronene	421	729	240.4847	7.6722	4.5470	5.6263	1.2697	0.7712	160.8888	1133	276.0806	1977
Benzo[ <i>h</i> ]pentaphene	346	590	193.7321	6.9675	4.1938	5.4251	1.2494	0.7685	138.8702	889	234.5940	1529
Benzo[ <i>h</i> ]tetracene	570	994	327.9796	9.9629	5.8405	7.0572	1.5707	0.9418	123.8517	1558	370.1075	2734
Benzo[ <i>a</i> ]ladiene	463	152.2372	5.3768	3.2336	4.1608	0.9567	0.5886	107.9073	702	182.5672	1210	
Benzo[ <i>c</i> ]phenanthrene	227	383	125.9898	4.7814	2.9203	3.8620	0.9020	0.5635	92.9258	573	155.0537	977
Benzo[ <i>e</i> ]pyrene	272	464	152.9898	5.3814	3.2536	4.1953	0.9687	0.6005	107.9258	708	182.0537	1220
Benzo[ <i>gh</i> ]perylene	316	544	179.2372	5.9768	3.5670	4.4941	1.0234	0.6256	122.9073	837	209.5672	1453
Benzo[ <i>gq</i> ]picene	331	567	186.4847	6.4722	3.8804	4.9596	1.1364	0.6971	130.8888	863	222.0806	1491
Benzobisanthene	465	809	266.7321	8.2675	4.8604	5.9251	1.3244	0.7963	175.8702	1262	303.5940	2210
Bisanthene	421	729	240.4847	7.6722	4.5470	5.6263	1.2697	0.7712	160.8888	1133	276.0806	1977
Chrysene	227	383	125.9898	4.7814	2.9203	3.8620	0.9020	0.5635	92.9258	573	155.0537	977
Coronene	360	624	205.4847	6.5722	3.8804	4.7929	1.0780	0.6508	137.8888	966	237.0806	1686
Dibenz[a, <i>h</i> ]anthracene	286	486	159.4847	5.8722	3.5470	4.6263	1.0697	0.6600	115.8888	728	195.0806	1248
Dibenz[ <i>a</i> , <i>h</i> ]anthracene	286	486	159.4847	5.8722	3.5470	4.6263	1.0697	0.6600	115.8888	728	195.0806	1248
Dibenz[B,C]coronene	464	808	265.9796	8.2629	4.8405	5.8906	1.3124	0.7844	175.8517	1256	304.1075	2200
Dibenzola[ <i>g</i> ]coronene	482	834	275.4847	8.7722	5.2137	6.4596	1.4614	0.8915	183.8888	1300	315.0806	2268
Dibenzola[ <i>gh</i> ]perylene	379	655	215.6084	7.0449	4.1760	5.2062	1.1769	0.7135	146.3795	1009.5	250.8373	1757.5
Dibenzolo[ <i>a,j</i> ]coronene	482	834	275.4847	8.7722	5.2137	6.4596	1.4614	0.8915	183.8888	1300	315.0806	2268
Dibenzolo[ <i>c</i> ]pentacene	404	692	226.4745	8.0536	4.8006	6.1549	1.4051	0.8532	161.8146	1038	275.1344	1790
Dibenzolo[ <i>c</i> ]phenylene	344	584	192.3610	6.9995	4.2515	5.5463	1.2875	0.8010	138.3980	883.5	232.3239	1515.5
Dibenzol[ <i>h,rs</i> ]pentaphene	392	672	221.4847	7.5722	4.5470	5.7929	1.3280	0.8174	153.8888	1030	261.0806	1782
Dibenzolo[ <i>c</i> ]tetraacene	346	590	193.7321	6.9675	4.1938	5.4251	1.2494	0.7685	138.8702	889	234.5940	1529
Dibenzola[ <i>e</i> ]elopyrene	332	568	187.2372	6.4768	3.9003	4.9941	1.1484	0.7090	130.9073	869	221.5672	1501
Dibenzola[ <i>e</i> ]elopyrene	331	567	186.4847	6.4722	3.8804	4.9596	1.1364	0.6971	130.8888	863	222.0806	1491
Dibenzola[ <i>gh</i> ]elopyrene	331	567	186.4847	6.4722	3.8804	4.9596	1.1364	0.6971	130.8888	863	222.0806	1491
Dibenzola[ <i>gh</i> ]elopyrene	332	568	187.2372	6.4768	3.9003	4.9941	1.1484	0.7090	130.9073	869	221.5672	1501
Dibenzola[ <i>gh</i> ]elopyrene	345	589	192.9796	6.9629	4.1738	5.3906	1.2374	0.7566	138.8517	883	235.1075	1519
Dibenzolo[ <i>e</i> ]elopyrene	333	569	187.9898	6.4814	3.9203	5.0286	1.1604	0.7209	130.9258	875	221.0537	1511
Heptacene	402	690	224.9694	8.0443	4.7608	6.0859	1.3811	0.8293	161.7775	1026	276.1612	1770
Heptaphene	403	691	225.7219	8.0490	4.7807	6.1204	1.3931	0.8412	161.7961	1032	275.6478	1780
Naphthacene	225	381	124.4847	4.7722	2.8804	3.7929	0.5397	0.8780	92.8888	561	156.0806	957
Naphthalene	107	175	57.4949	5.2907	1.6268	2.2643	0.5427	0.3466	46.9629	251	76.0269	415
Naphtho[1,2- <i>a</i> ]pyrene	331	567	186.4847	6.4722	3.8804	4.9596	1.1364	0.6971	130.8888	863	222.0806	1491
Naphtho[2,3- <i>a</i> ]coronene	480	832	273.9796	8.7629	5.1738	6.3906	1.4374	0.8677	183.8517	1288	316.1075	2248

(Continued)

**Table 2.** (Continued).

Compounds	<i>Bm</i>	<i>Tm</i>	<i>Gh</i>	<i>Gbm</i>	<i>Gtm</i>	<i>Hg</i>	<i>Hbm</i>	<i>Htm</i>	<i>Bmg</i>	<i>Bmh</i>	<i>Tmg</i>	<i>Tmh</i>
Naphtho[2,3- <i>a</i> ]pyrene	330	566	185.7321	6.4675	3.8604	4.9251	1.1244	0.6852	130.8702	857	222.5940	1481
Naphtho[5,4- <i>b</i> ]coronene	465	809	266.7321	8.2675	4.8604	5.9251	1.3244	0.7963	175.8702	1262	303.5940	2210
Naphtho[8,1,2- <i>bcd</i> ]perylene	376	648	213.4847	7.0722	4.2137	5.2929	1.2030	0.7341	145.8888	998	249.0806	1734
Peropyrene	376	648	213.4847	7.0722	4.2137	5.2929	1.2030	0.7341	145.8888	998	249.0806	1734
Perylene	272	464	152.9898	5.3814	3.2536	4.1953	0.9687	0.6005	107.9258	708	182.0537	1220
Phenanthrene	167	279	91.7423	3.6861	2.2735	3.0631	0.7223	0.4550	69.9444	412	115.5403	696
Picene	287	487	160.2372	5.8768	3.5670	4.6608	1.0817	0.6719	115.9073	734	194.5672	1258
Pyrene	211	359	117.9898	4.2814	2.5869	3.3620	0.7770	0.4801	84.9258	541	143.0537	929
Pyrenol[4,5- <i>e</i> ]pyrene	437	753	248.4847	8.1722	4.8804	6.1263	1.3947	0.8545	168.8888	1165	288.0806	2025
Terylene	437	753	248.4847	8.1722	4.8804	6.1263	1.3947	0.8545	168.8888	1165	288.0806	2025
Tetraphene	226	382	125.2372	4.7768	2.9003	3.8274	0.8900	0.5516	92.9073	567	155.5672	967
Tribenzo[ <i>b,n,pqr</i> ]perylene	438	754	249.2372	8.1768	4.9003	6.1608	1.4067	0.8664	168.9073	1171	287.5672	2035
Tribenzo[ <i>a,fg,opq</i> ]tetraene	393	673	222.2372	7.5768	4.5670	5.8274	1.3400	0.8294	153.9073	1036	260.5672	1792
Tribenzof[ <i>ij, noj</i> ]tetraphene	393	673	222.2372	7.5768	4.5670	5.8274	1.3400	0.8294	153.9073	1036	260.5672	1792
Triphenylene	228	384	126.7423	4.7861	2.9402	3.8965	0.9140	0.5754	92.9444	579	154.5403	987
Ovalene	509	889	292.9796	8.8629	5.1738	6.2239	1.3790	0.8214	190.8517	1391	331.1075	2443
Circumpyrene	703	1235	407.4745	11.7536	6.8006	7.9882	1.7467	1.0291	258.8146	1951	452.1344	3443
Circumcoronene	942	1662	548.9694	15.2443	8.7608	10.0859	2.1811	1.2738	341.7775	2646	600.1612	4686

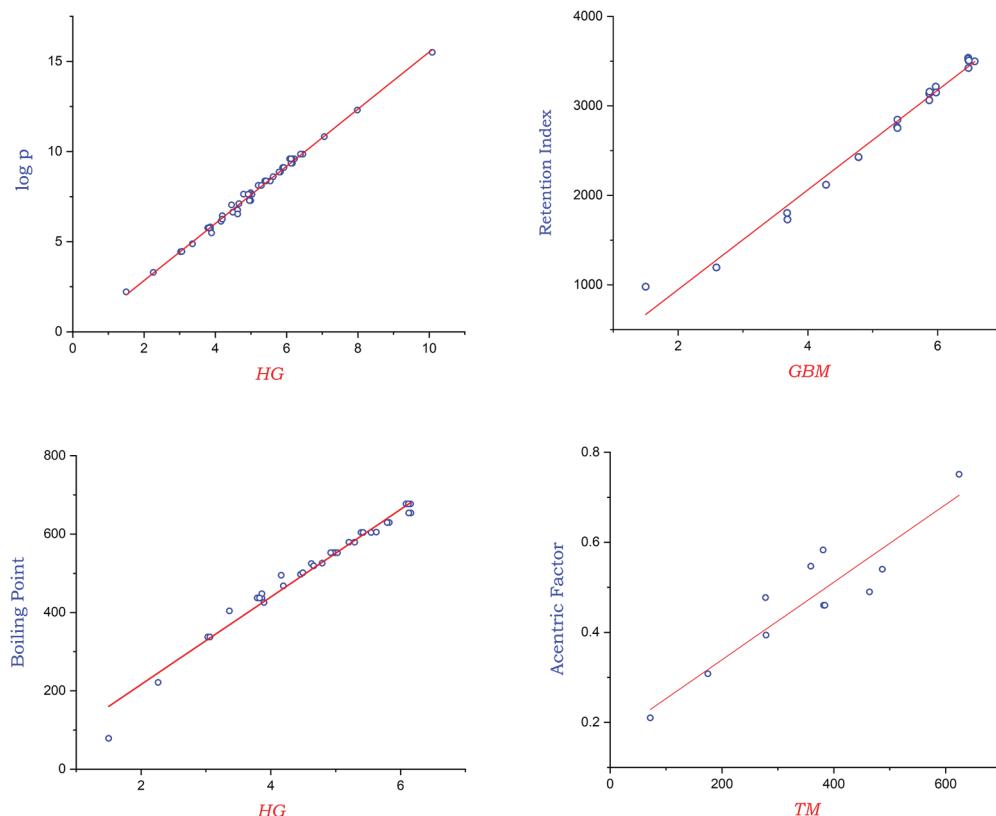
**Table 3.** Experimental physicochemical characteristics of benzenoid hydrocarbons.

Compounds	<i>log P</i>	<i>RI</i>	<i>BP °C</i>	<i>ω</i>	Refs
Anthanthrene	7.04	3215	497.1	—	[31,32,38]
Anthracene	4.45	1804	337.4	0.477	[31–33,38]
Benzene	2.04	979	80	0.21	[31,32,39]
Benzo[a]coronene	8.61	—	604.8	—	[31]
Benzo[h]pentaphene	8.37	—	604.1	—	[31]
Benzo[a]ovalene	10.83	—	—	—	[31]
Benzo[a]pyrene	6.13	2763	495	—	[31,32,38]
Benzo[c]phenanthrene	5.7	2427	436.7	—	[31,32,38]
Benzo[e]pyrene	6.44	2753	467.5	—	[31,32,38]
Benzo[ghi]perylene	6.63	3150	501	—	[31,32,38]
Benzo[pqr]picene	7.63	—	552.3	—	[31]
Benzobisanthene	9.11	—	—	—	[31]
Bisanthene	8.61	—	604.8	—	[31]
Chrysene	5.81	2429	448	0.46	[31,32,38,39]
Coronene	7.64	3498	525.6	0.751	[31–33,38]
Dibenz[a,h]anthracene	6.75	3137	524.7	—	[31,32,38]
Dibenz[a,j]anthracene	6.54	3063	524.7	—	[31,32,38]
Dibenz[B,C,K]Coronene	9.11	—	—	—	[31]
Dibenz[a,g]coronene	9.85	—	—	—	[31]
Dibenz[a,ghi]perylene	8.12	—	579	—	[31]
Dibenz[a,j]coronene	9.85	—	—	—	[31]
Dibenz[a,l]pentacene	9.6	—	677	—	[31]
Dibenzo[c,p]chrysene	8.37	—	604.1	—	[31]
Dibenzo[h,rst]Pentaphene	8.86	—	629.3	—	[31]
Dibenzo[a,c]tetracene	8.37	—	604.1	—	[31]
Dibenzo[a,e]pyrene	7.28	3507	552.3	—	[31,32,38]
Dibenzo[a,h]pyrene	7.28	3537	552.3	—	[31,32,38]
Dibenzo[a,i]pyrene	7.28	3526	552.3	—	[31,32,38]
Dibenzo[a,l]pyrene	7.71	3423	552.3	—	[31,32,38]
Dibenzo[a,l]tetracene	8.37	—	604.1	—	[31]
Dibenzo[e,l]pyrene	7.63	3508	552.3	—	[31,32]
Heptacene	9.6	—	677	—	[31]
Heptaphene	9.6	—	677	—	[31]
Naphthacene	5.76	—	436.7	0.583	[31,33,38]
Naphthalene	3.3	1194	218	0.308	[31–33,38]
Naphtho[1,2-a]pyrene	7.63	—	552.3	—	[31]
Naphtho[2,3-a]coronene	9.85	—	—	—	[31]
Naphtho[2,3-a]pyrene	7.63	—	552.3	—	[31]
Naphtho[5,4,3-abc]coronene	9.11	—	—	—	[31]
Naphtho[8,1,2-bcd]perylene	8.12	—	579	—	[31]
Peropyrene	8.12	—	579	—	[31]
Perylene	6.25	2846	467.5	0.49	[31,32,38,39]
Phenanthrene	4.46	1730	337.4	0.394	[31,32,38,39]
Picene	7.11	3159	519	0.54	[31,32,38,39]
Pyrene	4.88	2119	404	0.547	[31–33,38]
Pyreno[4,5-e]pyrene	9.35	—	653.8	—	[31]
Terylene	9.35	—	653.8	—	[31]
Tetraphene	5.76	—	436.7	0.46	[31,38,39]
Tribenzo[b,n,pqr]perylene	9.35	—	653.8	—	[31]
Tribenzo[a,fg,op]tetracene	8.86	—	629.3	—	[31]
Tribenzo[f,ij,no]tetraphene	8.86	—	629.3	—	[31]
Triphenylene	5.49	—	425	0.46	[31,38,39]
Ovalene	9.6	—	—	—	[31]
Circumpyrene	12.3	—	—	—	[31]
Circumcoronene	15.5	—	—	—	[31]

The general form of linear regression model is represented as  $\mathcal{P} = m\phi + c$ , where  $\mathcal{P}$  = property,  $c$  = intercept,  $m$  = slope and  $\phi$  = topological descriptor. In addition, we use the following notations  $F$  =  $F$  – statistic value of confidence and  $r^2$  = coefficient of determination. Therefore, the linear regression equations for highly correlated descriptors are given below:

1.  $\log P = 1.59(HG) - 0.35$ ,  
 $F = 9785.45$ ,  $r^2 = 0.99$
2.  $RI = 557(GBM) - 165.85$ ,  
 $F = 1155.88$ ,  $r^2 = 0.98$
3.  $P = 111.77(HG) - 7.61$ ,  
 $F = 1692.73$ ,  $r^2 = 0.98$
4.  $\omega = 0.00086(TM) + 0.17$ ,  
 $F = 52.17$ ,  $r^2 = 0.84$

There have been a few studies on the various physicochemical properties of benzenoid hydrocarbons using graph theoretical approaches [47–50]. The QSPR models were analysed for 21 benzenoid hydrocarbons by means of the rigid regression analysis with certain topological indices for  $\log P$ , chromatographic retention index, boiling point and Pitzer's acentric factor [22,51–53,55]. Additionally, other regression models have also been



**Figure 3.** Scatter diagrams for physicochemical properties with proposed descriptors.

**Table 4.** Correlation between physicochemical properties and proposed descriptors.

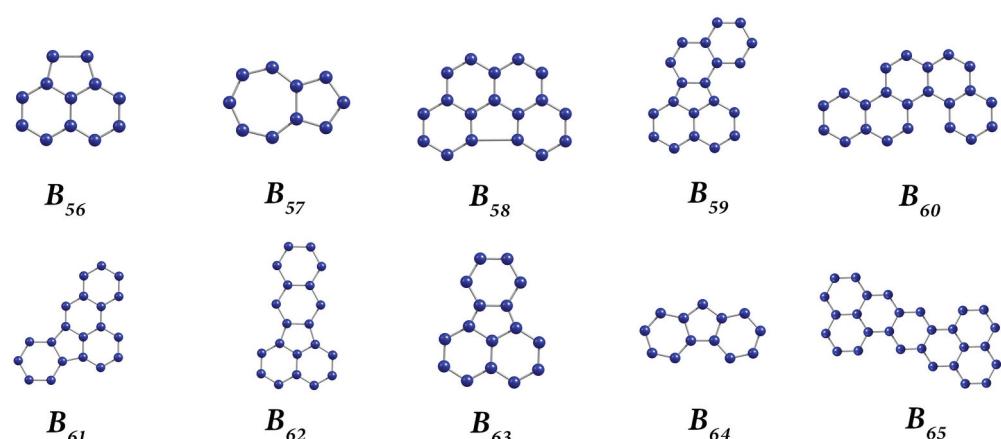
Index	<i>log P</i>	<i>RI</i>	<i>BP</i>	<i>ω</i>
<i>BM</i>	0.9732	0.9886	0.9623	0.9126
<i>TM</i>	0.9702	0.9871	0.9590	0.9160
<i>GH</i>	0.9692	0.9871	0.9569	0.9141
<i>GBM</i>	0.9919	0.9919	0.9832	0.8938
<i>GTM</i>	0.9941	0.9917	0.9842	0.8816
<i>HG</i>	0.9973	0.9876	0.9877	0.8589
<i>HBM</i>	0.9956	0.9836	0.9867	0.8399
<i>HTM</i>	0.9941	0.9785	0.9826	0.8816
<i>BMG</i>	0.9829	0.9911	0.9742	0.9076
<i>BMH</i>	0.9639	0.9851	0.9489	0.9138
<i>TMG</i>	0.9798	0.9898	0.9714	0.9135
<i>TMH</i>	0.9611	0.9836	0.9452	0.9154

proposed by including benzo[c]phenanthrene with a total of 22 benzenoid hydrocarbons [56–58]. The correlation values are obtained by newly proposed descriptors incorporating the structural information of large data set with recent experimental values into the model to enhance their predictive potential. Results from Table 4 shows relationship between hybrid indices and physicochemical properties.

The efficacy of the new indices in correlating with the physicochemical properties considered here can be explained by the underlying mechanisms that govern the physicochemical properties. The properties that we have considered here such as *log P*, boiling point, chromatographic retention index, etc., strongly depend on the extent of branching and the underlying connectivity distances and valence or degree of each vertex. These parameters are effectively included in the newly developed hybrid indices.

#### Comparison between experimental and predicted values

In this section, we test the efficacy of the regression models developed in the preceding section to obtain the physicochemical properties of polycyclic aromatic hydrocarbons



**Figure 4.** Benzenoid hydrocarbons: *B*<sub>56</sub>-Acenaphthylene, *B*<sub>57</sub>-Azulene, *B*<sub>58</sub>-Benzo[ghi]fluoranthene, *B*<sub>59</sub>-Benzol[j]fluoranthene, *B*<sub>60</sub>-Benzo[c]chrysene, *B*<sub>61</sub>-Benzo[b]fluoranthene, *B*<sub>62</sub>-Benzo[k]fluoranthene, *B*<sub>63</sub>-Fluoranthene, *B*<sub>64</sub>- Fluorene, *B*<sub>65</sub>- Heptazethrene.

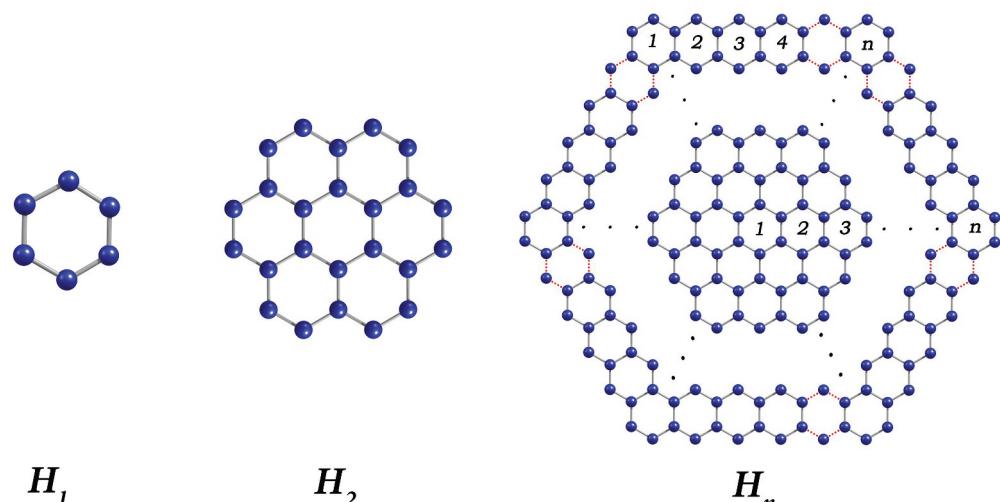
**Table 5.** Predicted physicochemical properties of polycyclic aromatic compounds.

Compounds	log P		RI		BP °C		$\omega$
	Expt.	Predicted	Expt.	Predicted	Expt.	Predicted	
Acenaphthylene	4.26	3.73	1425	1609	270	278.9	0.73
Azulene	3.45	3.25	1326	1277	245	245.5	0.53
Benzog[ghi]fluoranthene	5.66	5.53	2438	2553	432	405.4	1.18
Benzog[j]fluoranthene	–	6.32	2756	2832	480	461.3	1.22
Benzoc[c]chrysene	7.14	7.06	–	3108	–	513.3	1.25
Benzob[b]fluoranthene	–	6.32	–	2832	481	461.3	1.22
Benzok[k]fluoranthene	–	6.27	2761	2829	481	457.4	1.21
Fluoranthene	5.17	5.05	2057	2221	383	372	0.98
Fluorene	4.16	4.12	1557	1748	294	306.8	0.75
Heptazethrene	–	8.75	–	4047	–	632.2	1.69

that were not included in the data set of 55 compounds that were used to develop the QSPR expressions for four properties. For our analysis, we have chosen 10 polycyclic aromatic hydrocarbons which are shown in Figure 4, whose experimental physicochemical properties are available in the literatures [31,38]. The QSPR predicted physicochemical properties of selected hydrocarbons are compared with experiment when experimental results are available in Table 5. As can be seen from Table 5, overall QSPR predicted results are in reasonable agreement with known experimental results for these compounds.

### Generalized predictive measures of circumcoronene series

The molecular framework of circumcoronene series,  $H_n$  is obtained by circumscribing benzenes layers around the benzene ring as illustrated in Figure 5. There exist three classes in the edge partition based on the degree of end vertices of  $H_n$ . The cardinalities of edge partition have been already computed [67,68] which are as follows,  $d_{2,2} = 6$ ,  $d_{2,3} = 12(n - 1)$  and  $d_{3,3} = 9n^2 - 15n + 6$ .

**Figure 5.** Schematic growth of circumcoronene series,  $H_n$ .

By using the cardinality of partition and the explicit formula of the indices one can easily obtain the generalized expression to compute hybrid indices. Further, we have presented generalized expressions for indices that have the best correlation with the physicochemical properties in the following.

$$(1) \ HG(H_n) = n^2 + \frac{41225546728941}{140737488355328}n - \frac{4\sqrt{6}}{5} + \frac{13}{6}$$

$$(2) \ GBM(H_n) = \frac{9}{5} \left[ n^2 - \frac{205057254912529}{112589906842624}n - \frac{20\sqrt{6}}{33} + \frac{3}{2} \right]$$

$$(3) \ TM(H_n) = 3[81n^2 - 59n + 2]$$

By substituting the obtained expression in regression model, we can generalize the physicochemical attributes of  $H_n$ .

$$(1) \ logP(H_n) = 1.5892n^2 + 0.4655n - 0.0210$$

$$(2) \ RI(H_n) = 1002.5994n^2 - 182.6008n - 150.3473$$

$$(3) \ BP(H_n) = 111.7668n^2 + 32.7393n + 15.5331$$

$$(4) \ \omega(H_n) = 0.2095n^2 - 0.15268n + 0.1718$$

**Table 6** displays the boiling points predicted by QSPR compared to available experimental results for different values of  $n$ . It is evident that the QSPR expression for the boiling point not only displays the correct trend but provides very reasonable estimates for circumscribing coronenes. In this context, the domain of validity of the derived expressions needs to be discussed. We note that the expression is clearly not valid to predict boiling point and retention index for  $n = 1$  or benzene because the derived expression is for circumscribing coronene and thus the minimum value of  $n$  must be 2, as shown next to the QSPR expressions. As the expression is quadratic in  $n$  for all properties considered here, it is expected that the validity of the QSPR expression is limited to an upper value of  $n$ . This is derived by either finding the roots of the quadratic equation that scales to the properties of a layer of graphite. For example, the melting process of graphite would cleave the van der Waal's bonds between the adjacent layers of graphene sheets in the graphite materials. Thus, the boiling process which involves a phase transition from the liquid to the gaseous state is suitably modelled by sufficiently large 2D sheets of graphite. Our scaling of the boiling point expression derived for the coronene series then suggests that  $n = 6.4$  would be the optimal scaling that leads to reproducing the observed boiling point of graphite. Hence the validity of the upper limit of the boiling point expression is 6 as shown adjacent to the QSPR expression. Likewise, the measured chromatographic retention indices for

**Table 6.** QSPR predictions for the boiling points of coronene series extrapolated to graphite.

Coronene series	QSPR predicted BP °C	Expt. BP °C
Coronene ( $n = 2$ )	~528.1	525.6 (524.8; Ref. [69])
Circumcoronene ( $n = 3$ )	~1119.7	–
Circumcircumcoronene ( $n = 4$ )	~1934.8	–
3-circumcoronene ( $n = 5$ )	~2973.4	–
4-circumcoronene ( $n = 6$ )	~4235.6	–
Graphite ( $n = 6.4$ )	~4803 ( $n = 6.4$ )	4827

a number of polycyclic aromatics [18,70] were used to arrive at the domain of validity of  $n$  values for the chromatographic retention indices.

The chromatographic retention indices are directly connected to the diffusion constants for the aromatic molecules to diffuse through the chromatographic column material. Consequently, the retention indices are correlated with the molar masses and the extent of branching both of which are correlated quite well with the newly introduced topological indices which provide quantitative measures of these quantities. For example, the QSPR retention index of coronene derived from our expression is 3495 in comparison with the experimental value reported by Grimmer et al. [71] to be 3498 using the polycyclic aromatic hydrocarbon profile analysis, which is in agreement with QSPR. We note that although the QSPR expression for RI of the coronene series as well as the previous general expression predict the retention indices of multi-ring benzenoids reasonably well, RI expression fails for the simplest case of benzene. The coronene series expression for  $n = 1$  yields 669.65 as the retention index which contradicts well known experimental result of 979 for the chromatographic retention index of benzene [72]. Consequently, benzene is an outlier of the QSPR as it does not fit well in a data set of more complex polycyclic aromatics. A simple connectivity index works better in predicting properties of such isolated simple molecules as benzene.

Pitzer's acentric factor ( $\omega$ ) is defined as  $\omega = -\log_{10}(p_r^{sat}) - 1.0$  at reduced temperature  $T_r = 0.7$ , where  $p_r^{sat} = \frac{p^{sat}}{p_c}$  and the ratio is the reduced saturation vapour pressure and  $p^{sat}$  is the saturation vapour pressure. The acentric factor for fluids measures the deviation from sphericity and thus it correlates well with the degree of branching and the extent of topological spread. Several experimental studies have been employed to measure acentric factors. Although the acentric factor can vary from  $-1$  to positive integers, most aromatic compounds have acentric factors below 3. In general, as the molecule becomes less spherical and the number of atoms increases, the acentric factor tends to increase. Leon et al. [33] have carried out extensive studies on acentric factors of mixtures of aromatic compounds of Colombian origin, that is, of SARA (Saturated, Aromatic, Resin, and Asphaltene) fractions of the vacuum residue obtained from the Barrancabermeja refinery. They found that for larger aromatic compounds containing several benzene rings the acentric factors were a bit less than 2. Our QSPR expression for the acentric factor yields a value of 6.8 for 4-circumcoronene which is somewhat large. However, it is not unreasonable in that it suggests very low reduced saturation vapour pressure of  $10^{-4.5}$  implying extremely low extent of vaporization for such a large polycyclic aromatic which is close to 2D-graphene. The low reduced saturation vapour pressure is also consistent with a very large boiling point for graphite. For such reasons we stipulate the domain of validity of our QSPR expression for the acentric factor as  $n \leq 6$  for the coronene series.

There are a number of limitations to the QSPR expressions developed here. First, they work only for the congeneric set of compounds considered to derive the QSPR expressions. That is, the expressions seem to work for a series of polycyclic aromatic compounds with fused rings such that there are no cavities or holes. For example, the QSPR expressions are not applicable to kekulenes, septulenes, octulenes, holey nanographenes and so forth. The techniques also fail for the simplest molecules such as benzene and naphthalene as they become outliers when one considers of polycyclic aromatics containing a larger number of rings that were included in the congeneric set to develop the QSPR expressions.

## Conclusion

In this study, we have introduced new variations of the Zagreb descriptors, which we called the bi-Zagreb and tri-Zagreb descriptors. Furthermore, we have developed novel descriptors by taking the ratio combination of geometric, harmonic, bi-Zagreb and tri-Zagreb which in essence constitute hybrid geometric-harmonic-Zagreb descriptors and we have tested their efficacy to predict the physicochemical properties of polycyclic benzenoid hydrocarbons. We found that the newly developed hybrid descriptors are efficacious compared to the existing indices. We have also presented the linear regression QSAPR expressions for a number of physic chemical properties. Although the developed hybrid indices performed quite well for some of the physicochemical properties such as the octanol partition coefficients, boiling point, chromatographic retention indices etc., our attempts to correlate these indices with enthalpies of formation in the gas phase were not successful. It would be highly desirable to have robust and yet reliable QSPR relations for the enthalpy of formations of polycyclic aromatic compounds as at present we have more data from *ab initio* Gaussian-3 computations and experimental studies [44] on enthalpies compared to a couple of decades ago. Our QSPR expressions for four physicochemical properties obtained for circumscribing coronene series suggest that 2D-sheets of graphene are adequately modelled by fourth circumscribing around coronene. These features could also facilitate measures of critical sizes of 2D-sheets of graphene required for modelling such 2D-sheets for computing their properties through first principles computations.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

The author(s) reported there is no funding associated with the work featured in this article.

## ORCID

- M. Arockiaraj  <http://orcid.org/0000-0002-4457-1046>  
D. Paul  <http://orcid.org/0000-0003-1832-310X>  
J. Clement  <http://orcid.org/0000-0001-7797-6456>  
K. Jacob  <http://orcid.org/0000-0002-6407-7790>  
K. Balasubramanian  <http://orcid.org/0000-0002-9612-8783>

## References

- [1] H. González-Díaz, S. Vilar, L. Santana, and E. Uriarte, *Medicinal chemistry and bioinformatics current trends in drugs discovery with networks topological indices*, Curr. Top. Med. Chem. 7 (2007), pp. 1015–1029. doi:[10.2174/156802607780906771](https://doi.org/10.2174/156802607780906771).
- [2] D. Maji, G. Ghorai, M.K. Mahmood, and M.A. Alam, *On the inverse problem for some topological indices*, J. Math. 2021 (2021), pp. 1–8. doi:[10.1155/2021/9411696](https://doi.org/10.1155/2021/9411696).

- [3] S. Hosamani, D. Perigidad, S. Jamagoud, Y. Maled, and S. Gavade, *QSPR analysis of certain degree based topological indices*, J. Stat. Appl. Pro. 6 (2017), pp. 361–371. doi:10.18576/jsap/060211.
- [4] J. Devillers and A.T. Balaban, *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.
- [5] S. Hayat and M. Imran, *Computation of topological indices of certain networks*, Appl. Math. Comput. 240 (2014), pp. 213–228.
- [6] X. Zhang, H.M. Awais, M. Javaid, and M.K. Siddiqui, *Multiplicative Zagreb indices of molecular graphs*, J. Chem. 2019 (2019), pp. 1–19. doi:10.1155/2019/6720616.
- [7] M. Arockiaraj, S.R.J. Kavitha, K. Balasubramanian, and J.-B. Liu, *On certain topological indices of octahedral and icosahedral networks*, IET Control Theory Appl. 12 (2018), pp. 215–220. doi:10.1049/iet-cta.2017.0671.
- [8] S. Hayat and S. Khan, *Quality testing of spectrum-based valency descriptors for polycyclic aromatic hydrocarbons with applications*, J. Mol. Struct. 1228 (2021), pp. 129789. doi:10.1016/j.molstruc.2020.129789.
- [9] I. Gutman, *Degree-based topological indices*, Croat. Chem. Acta 86 (2013), pp. 351–361. doi:10.5562/cca2294.
- [10] M.I. Stankevich, I.V. Stankevich, and N.S. Zefirov, *Topological indices in organic chemistry*, Russ. Chem. Rev. 57 (1988), pp. 191–208. doi:10.1070/RC1988v057n03ABEH003344.
- [11] D. Paul, M. Arockiaraj, K. Jacob, and J. Clement, *Multiplicative versus scalar multiplicative degree based descriptors in QSAR/QSPR studies and their comparative analysis in entropy measures*, Eur. Phys. J. Plus 138 (2023), pp. 323. doi:10.1140/epjp/s13360-023-03920-7.
- [12] M. Arockiaraj, A.B. Greeni, and A.R.A. Kalaam, *Linear versus cubic regression models for analyzing generalized reverse degree based topological indices of certain latest Corona treatment drug molecules*, Int. J. Quantum Chem. 123 (2023), pp. e27136. doi:10.1002/qua.27136.
- [13] K. Balasubramanian, *Computational and Artificial Intelligence Techniques for Drug Discovery and Administration, Reference Module in Biomedical Sciences*, Elsevier, Amsterdam, 2021.
- [14] S. Hayat and M. Imran, *On degree based topological indices of certain nanotubes*, J. Comput. Theor. Nanosci. 12 (2015), pp. 1–7.
- [15] T. Došlić, B. Furtula, A. Graovac, I. Gutman, S. Moradi, and Z. Yarahmadi, *On vertex-degree-based molecular structure descriptors*, MATCH Commun. Math. Comput. Chem. 66 (2011), pp. 613–626.
- [16] M. Arockiaraj, S.R.J. Kavitha, K. Balasubramanian, and I. Gutman, *Hyper-Wiener and Wiener polarity indices of silicate and oxide frameworks*, J. Math. Chem. 56 (2018), pp. 1493–1510. doi:10.1007/s10910-018-0881-x.
- [17] K. Balasubramanian, *Mathematical and computational techniques for drug discovery: Promises and developments*, Curr. Top. Med. Chem. 18 (2018), pp. 2774–2799. doi:10.2174/1568026619666190208164005.
- [18] Z. Wang, K. Li, P. Lambert, and C. Yang, *Identification, characterization and quantitation of pyrogenic polycyclic aromatic hydrocarbons and other organic compounds in tire fire products*, J. Chromatogr. A 1139 (2007), pp. 14–26. doi:10.1016/j.chroma.2006.10.085.
- [19] M. Arockiaraj, J. Clement, and K. Balasubramanian, *Topological indices and their applications to circumcised donut benzenoid systems, kekulenes and drugs*, Polycyclic Aromat. Compd. 40 (2017), pp. 280–303. doi:10.1080/10406638.2017.1411958.
- [20] B. Furtula and I. Gutman, *A forgotten topological index*, J. Math. Chem. 53 (2015), pp. 1184–1190. doi:10.1007/s10910-015-0480-z.
- [21] G.H. Shirdel, H. Rezapour, and A.M. Sayadi, *The hyper Zagreb index of graph operations*, Iranian J. Math. Chem. 4 (2013), pp. 213–220.
- [22] G.V. Rajasekharaiyah and U.P. Murthy, *Hyper-Zagreb indices of graphs and its applications*, J. Algebra Comb. Discrete Struct. Appl. 8 (2021), pp. 9.
- [23] D. Vukičević and B. Furtula, *Topological index based on the ratios of geometrical and arithmetical means of end vertex degrees of edges*, J. Math. Chem. 46 (2009), pp. 1369–1376. doi:10.1007/s10910-009-9520-x.

- [24] H. Hosoya, *On some counting polynomials in chemistry*, Discrete Appl. Math. 19 (1988), pp. 239–257. doi:[10.1016/0166-218X\(88\)90017-0](https://doi.org/10.1016/0166-218X(88)90017-0).
- [25] G.G. Cash, *Relationship between the Hosoya polynomial and the hyper-Wiener index*, Appl. Math. Let. 15 (2002), pp. 893–895. doi:[10.1016/S0893-9659\(02\)00059-9](https://doi.org/10.1016/S0893-9659(02)00059-9).
- [26] E. Deutsch and S. Klavžar, *M-polynomial and degree-based topological indices*, Iran. J. Math. Chem. 6 (2015), pp. 93–102.
- [27] M.I. Huilgol, V. Sriram, and K. Balasubramanian, *Structure-activity relations for antiepileptic drugs through omega polynomials and topological indices*, Mol. Phys. 120 (2022), pp. e1987542.
- [28] M.I. Huilgol, V. Sriram, H.J. Udupa, and K. Balasubramanian, *Computational studies of toxicity and properties of  $\beta$ -diketones through topological indices and M/NM-polynomials*, Comput. Theor. Chem. 1224 (2023), pp. 114108. doi:[10.1016/j.comptc.2023.114108](https://doi.org/10.1016/j.comptc.2023.114108).
- [29] M. Ajmal, W. Nazeer, M. Munir, S.M. Kang, and C.Y. Jung, *The M-polynomials and topological indices of toroidal polyhex network*, Int. J. Math. Anal. 11 (2017), pp. 305–315. doi:[10.12988/ijma.2017.7119](https://doi.org/10.12988/ijma.2017.7119).
- [30] F. Afzal, S. Hussain, D. Afzal, and S. Razaq, *Some new degree based topological indices via M polynomial*, J. Inf. Optim. Sci. 41 (2020), pp. 1061–1076.
- [31] ChemSpider, <http://www.chemspider.com/>.
- [32] NIST Chemistry Webbook, <https://webbook.nist.gov/>.
- [33] A. Yovani León, M. Parra, and J.L. Grosso, *Estimation of critical properties of typically Colombian vacuum residue Sara fractions*, CT&F- Ciencia, Tecnología Y Futuro 3 (2008), pp. 129–142. doi:[10.29047/01225383.467](https://doi.org/10.29047/01225383.467).
- [34] K.S. Pitzer, D.Z. Lippmann, R.F. Curl, C.M. Huggins, and D.E. Petersen, *The volumetric and thermodynamic properties of fluids II. Compressibility factor, vapor pressure and entropy of vaporization*, J. Am. Chem. Soc. 77 (1955), pp. 3433–3440. doi:[10.1021/ja01618a002](https://doi.org/10.1021/ja01618a002).
- [35] R.A. Alberty and A.K. Reif, *Standard chemical thermodynamic properties of polycyclic aromatic hydrocarbons and their isomer groups I. Benzene series*, J. Phys. Chem. Ref. Data 17 (1988), pp. 241–253.
- [36] R.A. Alberty, M.B. Chung, and A.K. Reif, *Standard chemical thermodynamic properties of polycyclic aromatic hydrocarbons and their isomer groups. II. Pyrene series, naphthopyrene series, and coronene series*, J. Phys. Chem. Ref. Data 18 (1989), pp. 77–109. doi:[10.1063/1.555826](https://doi.org/10.1063/1.555826).
- [37] R.A. Alberty, M.B. Chung, and A.K. Reif, *Standard chemical thermodynamic properties of polycyclic aromatic hydrocarbons and their isomer groups. III. Naphthocoronene series, ovalene series, and first members of some higher series*, J. Phys. Chem. Ref. Data 19 (1990), pp. 349–370. doi:[10.1063/1.555852](https://doi.org/10.1063/1.555852).
- [38] C.M. White, *Prediction of the boiling point, heat of vaporization, and vapor pressure at various temperatures for polycyclic aromatic hydrocarbons*, J. Chem. Eng. Data 31 (1986), pp. 198–203. doi:[10.1021/je00044a020](https://doi.org/10.1021/je00044a020).
- [39] H. Wang and M. Frenklach, *Transport properties of polycyclic aromatic hydrocarbons for flame modeling*, Combust. Flame 96 (1994), pp. 163–170.
- [40] M.M. Ferreira, *Polycyclic aromatic hydrocarbons: A QSPR study*, Chemosphere 44 (2001), pp. 125–146.
- [41] V.H. Alvarez and M.D. Saldana, *Modeling solubility of polycyclic aromatic compounds in sub-critical water*, Ind. Eng. Chem. Res. 50 (2011), pp. 11396–11405. doi:[10.1021/ie201019b](https://doi.org/10.1021/ie201019b).
- [42] Q. Wang, Q. Jia, and P. Ma, *Prediction of the acentric factor of organic compounds with the positional distributive contribution method*, J. Chem. Eng. Data 57 (2012), pp. 169–189. doi:[10.1021/je200971z](https://doi.org/10.1021/je200971z).
- [43] M. Perry and C. White, *Correlations of molecular connectivity with critical volume and acentricity*, AIChE J. 33 (1987), pp. 146–151. doi:[10.1002/aic.690330117](https://doi.org/10.1002/aic.690330117).
- [44] T.C. Allison and D.R. Burgess, *First-principles prediction of enthalpies of formation for polycyclic aromatic hydrocarbons and derivatives*, J. Phys. Chem. A. 119 (2015), pp. 11329–11365. doi:[10.1021/acs.jpca.5b07908](https://doi.org/10.1021/acs.jpca.5b07908).

- [45] J. Verma, V.M. Khedkar, and E.C. Coutinho, *3D-QSAR in drug design - a review*, *Curr. Top. Med. Chem.* 10 (2010), pp. 95–115. doi:[10.2174/156802610790232260](https://doi.org/10.2174/156802610790232260).
- [46] S. Kwon, H. Bae, J. Jo, and S. Yoon, *Comprehensive ensemble in QSAR prediction for drug discovery*, *BMC Bioinf.* 20 (2019), pp. 1–12. doi:[10.1186/s12859-019-3135-4](https://doi.org/10.1186/s12859-019-3135-4).
- [47] A. Jahanbani, S.M. Sheikholeslami, and R. Khoeilar, *On the energy of benzenoid hydrocarbons*, *Polycyclic Aromat. Compd.* 42 (2021), pp. 5204–5216. doi:[10.1080/10406638.2021.1933103](https://doi.org/10.1080/10406638.2021.1933103).
- [48] I. Gutman and S. Klavžar, *Bounds for the Schultz molecular topological index of benzenoid systems in terms of the Wiener index*, *J. Chem. Inf. Comput. Sci.* 37 (1997), pp. 741–744. doi:[10.1021/ci9700034](https://doi.org/10.1021/ci9700034).
- [49] B. Lučić, N. Trinajstić, and B. Zhou, *Comparison between the sum-connectivity index and product-connectivity index for benzenoid hydrocarbons*, *Chem. Phys. Lett.* 475 (2009), pp. 146–14. doi:[10.1016/j.cplett.2009.05.022](https://doi.org/10.1016/j.cplett.2009.05.022).
- [50] R. Kanabur and V. Shigehalli, *QSPR analysis of degree-based topological indices with physical properties of benzenoid hydrocarbons*, *Gen. Lett. Math.* 2 (2017), pp. 5–169.
- [51] Z. Yu, S. Zhou, and T. Tian, *Inverse sum indeg reciprocal status index and co-index of graphs*, *Circuits Syst. Signal Process* 42 (2022), pp. 1–21.
- [52] A.A. Khabyah, S. Zaman, A.N. Koam, A. Ahmad, and A. Ullah, *Minimum Zagreb eccentricity indices of two-mode network with applications in boiling point and benzenoid hydrocarbons*, *Mathematics* 10 (2022), pp. 1393. doi:[10.3390/math10091393](https://doi.org/10.3390/math10091393).
- [53] M.K. Jamil, M. Imran, and K.A. Sattar, *Novel face index for benzenoid hydrocarbons*, *Mathematics* 8 (2020), pp. 312. doi:[10.3390/math8030312](https://doi.org/10.3390/math8030312).
- [54] J.A. Bumpus, *Gas-phase heat of formation values for buckminsterfullerene ( $C_{60}$ ), C<sub>70</sub> fullerene ( $C_{70}$ ), corannulene, coronene, sumanene, and other polycyclic aromatic hydrocarbons calculated using density functional theory (M06 2x) coupled with a versatile inexpensive group-equivalent approach*, *J. Phys. Chem. A* 122 (2018), pp. 6615–6632. doi:[10.1021/acs.jpca.8b03321](https://doi.org/10.1021/acs.jpca.8b03321).
- [55] H.S. Ramane and A.S. Yalnaik, *Status connectivity indices of graphs and its applications to the boiling point of benzenoid hydrocarbons*, *J. Appl. Math. Comput.* 55 (2016), pp. 609–627. doi:[10.1007/s12190-016-1052-5](https://doi.org/10.1007/s12190-016-1052-5).
- [56] S. Hayat, S. Khan, M. Imran, and J.B. Liu, *Quality testing of distance-based molecular descriptors for benzenoid hydrocarbons*, *J. Mol. Struct.* 1222 (2020), pp. 128927. doi:[10.1016/j.molstruc.2020.128927](https://doi.org/10.1016/j.molstruc.2020.128927).
- [57] M. Chamua, J. Buragohain, A. Bharali, and M.E. Nazari, *Predictive ability of neighborhood degree sum-based topological indices of polycyclic aromatic hydrocarbons*, *J. Mol. Struct.* 1270 (2022), pp. 133904. doi:[10.1016/j.molstruc.2022.133904](https://doi.org/10.1016/j.molstruc.2022.133904).
- [58] P. Sarkar, N. De, and A. Pal, *On some topological indices and their importance in chemical sciences: A comparative study*, *Eur. Phys. J. Plus.* 137 (2022), pp. 195. doi:[10.1140/epjp/s13360-022-02431-1](https://doi.org/10.1140/epjp/s13360-022-02431-1).
- [59] A. Tropsha, P. Gramatica, and V.K. Gombar, *The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models*, *Comb. Sci.* 22 (2003), pp. 69–77. doi:[10.1002/qsar.200390007](https://doi.org/10.1002/qsar.200390007).
- [60] K. Roy, *Advances in QSAR Modeling: Applications in Pharmaceutical, Chemical, Food, Agricultural and Environmental Sciences*, Springer, Cham, Switzerland, 2017.
- [61] X. Zhang, M.K. Siddiqui, S. Javed, L. Sherin, F. Kausar, and M.H. Muhammad, *Physical analysis of heat for formation and entropy of Ceria oxide using topological indices*, *Comb. Chem. High Throughput Screen.* 25 (2022), pp. 441–450. doi:[10.2174/1386207323999201001210832](https://doi.org/10.2174/1386207323999201001210832).
- [62] S.A.K. Kirmani, P. Ali, and F. Azam, *Topological indices and QSPR/QSAR analysis of some antiviral drugs being investigated for the treatment of COVID-19 patients*, *Int. J. Quantum Chem.* 121 (2021), pp. e26594. doi:[10.1002/qua.26594](https://doi.org/10.1002/qua.26594).
- [63] K. Roy, S. Kar, and R.N. Das, *QSAR/QSPR modeling: Introduction*, Springer Briefs in Mol. Sci. (2015), pp. 1–36.
- [64] F. Abedin, B. Roughton, P. Spencer, Q. Ye, and K. Camarda, *Computational molecular design of water compatible dentin adhesive system*, *Int. Symp. Process. Eng.* 37 (2015), pp. 2081–2086.

- [65] Z. Raza, *The expected values of arithmetic bond connectivity and geometric indices in random phenylene chains*, Heliyon 6 (2020), pp. 44–59. doi:[10.1016/j.heliyon.2020.e04479](https://doi.org/10.1016/j.heliyon.2020.e04479).
- [66] X. Zhang, X. Wu, S. Akhter, M. Jamil, J.B. Liu, and M. Farahani, *Edge-version atom-bond connectivity and geometric arithmetic indices of generalized bridge molecular graphs*, Symmetry 10 (2018), pp. 751–786. doi:[10.3390/sym10120751](https://doi.org/10.3390/sym10120751).
- [67] I. Gutman, S.J. Cyvina, and V. Ivanov-Petrović, *Topological properties of circumcoronenes*, Z. Naturforsch. A. 53 (1998), pp. 699–70. doi:[10.1515/zna-1998-0810](https://doi.org/10.1515/zna-1998-0810).
- [68] M. Arockiaraj, J. Clement, and K. Balasubramanian, *Analytical expressions for topological properties of polycyclic benzenoid networks*, J. Chemom. 30 (2016), pp. 682–697.
- [69] P. Liu, Y. Li, S.M. Sarathy, and W.L. Roberts, *Gas-to-liquid phase transition of PAH at flame temperatures*, J. Phys. Chem. A. 124 (2020), pp. 3896–3903. doi:[10.1021/acs.jpca.0c01912](https://doi.org/10.1021/acs.jpca.0c01912).
- [70] M. Piao, S. Chu, M. Zheng, and X. Xu, *Characterization of the combustion products of polyethylene*, Chemosphere 39 (1999), pp. 1497–1512. doi:[10.1016/S0045-6535\(99\)00054-5](https://doi.org/10.1016/S0045-6535(99)00054-5).
- [71] G. Grimmer and H. Böhnke, *Polycyclic aromatic hydrocarbon profile analysis of high-protein foods, oils, and fats by gas chromatography*, J. Ass. Offic. Anal. Chem. 58 (1975), pp. 725–733.
- [72] S.V. Kurbatova, E.E. Finkelshtein, E.A. Kolosova, A.V. Kartashev, and S.N. Yashkin, *Structural analogy method in studies of adamantanes*, J. Struct. Chem. 45 (2004), pp. 144–150. doi:[10.1023/B:JORY.0000041513.82837.4e](https://doi.org/10.1023/B:JORY.0000041513.82837.4e).