

Article

Bounds for General Zagreb Indices and Improved Topological Coindices with QSPR Benchmarking on Octane Isomers

Suha Wazzan ^{1,*}  and Abdu Alameri ² 

¹ Department of Mathematics, Science Faculty, King Abdulaziz University, Jeddah 21589, Saudi Arabia

² Department of Biomedical Engineering, Faculty of Engineering, University of Science and Technology, Sana'a P.O. Box 15201, Yemen; a.alameri2222@gmail.com

* Correspondence: swazzan@kau.edu.sa

Abstract

Topological descriptors play an important role in chemical graph theory and QSPR/QSAR studies by relating molecular structure to measurable physicochemical properties. Among standard benchmark families, octane isomers are frequently used to evaluate the behavior of degree-based descriptors because of their rich branching patterns and well-documented physicochemical data. Although many studies have examined topological indices for octane isomers, comparatively fewer works have focused on topological coindices and derived coindex-based descriptors. In this work, we study several known topological coindices and four derived descriptors, denoted by KJ_1 , KJ_2 , KJ_3 , and KJ_4 , for comparative analysis on the octane-isomer benchmark. We also present a unified treatment of lower and upper bounds for the first and second (α, β) -general Zagreb indices, together with their reduced and expanded variants, in terms of basic graph parameters, such as the minimum degree, maximum degree, order, and size. These bounds cover a range of familiar special cases, including classical Zagreb, forgotten, Sombor, and Randić-type indices, thereby placing several known descriptors within a common framework. For the application part, the original octane-isomer analysis is retained as a controlled benchmark for the proposed descriptors. In addition, an expanded QSPR experiment is added using a Zenodo molecular dataset containing 90 organic compounds and nine physicochemical endpoints. SMILES strings were converted into hydrogen-suppressed molecular graphs, graph-theoretical descriptors were computed, and ordinary least squares, ridge regression, and PLS(2) models were evaluated using an 80:20 train/test split and five-fold cross-validation. The expanded results show strong or useful performance for selected endpoints, especially critical volume, molecular volume, standard Gibbs free energy of formation, and logarithmic water solubility, whereas some temperature-related endpoints remain less stable. The results therefore support the usefulness of degree-based and coindex-based descriptors as compact exploratory QSPR variables while also emphasizing the need for cautious interpretation, redundancy analysis, and external validation on broader chemical families.



Academic Editor: Calogero Vetro

Received: 21 May 2026

Revised: 29 June 2026

Accepted: 30 June 2026

Published: 3 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: topological descriptors; topological indices; topological coindices; Zagreb indices; octane isomers; QSPR

1. Introduction

A topological descriptor is a numerical quantity associated with a molecular graph, where vertices represent atoms and edges represent chemical bonds. Such descriptors,

including topological indices and topological coindices, are widely used in quantitative structure–property/activity relationship (QSPR/QSAR) studies to explore how molecular structure is reflected in physicochemical behavior. According to the International Academy of Mathematical Chemistry, a useful topological parameter should display meaningful association with physicochemical properties for representative molecular families, such as octane isomers.

Octane isomers constitute a standard benchmark family in this context because they exhibit diverse branching patterns while retaining a common molecular formula, which makes them suitable for comparative descriptor studies [1–5]. Topological descriptors have also been successfully used in medicinal chemistry, nanomaterials, and related applied graph-theoretic investigations [6–8]. While the literature on topological indices is extensive, the systematic study of topological coindices remains comparatively limited. Recent QSPR studies have further emphasized the usefulness of topological coindices in chemical applications, including drug-related datasets. In particular, Öztürk Sözen and Eryaşar investigated coindex-based descriptors for breast-cancer-related compounds and for COVID-19 drug candidates, demonstrating their applicability in the analysis of physicochemical properties [9,10]. In addition, Alameri and Alsharafi studied topological indices associated with several new graph operations and discussed their possible applications, further illustrating the continued development of degree-based descriptors and their relevance in applied graph-theoretic investigations [11].

The present manuscript has two complementary aims. First, it develops a unified degree-based framework for the first and second (α, β) -general Zagreb indices and their reduced and expanded forms, emphasizing how many familiar descriptors arise as special cases. Second, it investigates whether selected coindices and four derived descriptors can serve as informative benchmark descriptors for physicochemical properties of octane isomers. Unlike entropy-based descriptor frameworks, our focus here is specifically on Zagreb-type indices and coindices and on the interpretation of their degree-based structure within a QSPR setting. At the same time, we do not claim broad predictive generality from the octane data alone; rather, we regard the application part as an exploratory benchmark study that helps compare the relative behavior of the considered descriptors.

The statistical application part combines two complementary QSPR settings. The octane-isomer dataset is used as a controlled benchmark for comparing the considered coindices and KJ -type descriptors within a structurally homogeneous molecular family. To extend the analysis beyond this small benchmark, a Zenodo dataset containing 90 organic compounds and nine physicochemical endpoints is also examined. In this broader setting, SMILES strings are converted into molecular graphs, descriptors are computed, and model performance is assessed using an external train/test split, five-fold cross-validation, RMSE, MAE, and descriptor-redundancy analysis. Together, these two datasets provide both a classical benchmark comparison and a broader exploratory evaluation of the QSPR relevance of the proposed descriptors.

Throughout this paper we consider a simple, finite and connected graph denoted by $\mathfrak{G}$. The vertex set and edge set of $\mathfrak{G}$ are denoted by $V(\mathfrak{G})$ and $E(\mathfrak{G})$, respectively. For a vertex $\varsigma \in V(\mathfrak{G})$, the number of incident edges is called its degree and is denoted by $\delta_{\mathfrak{G}}(\varsigma)$ (or simply $\delta(\varsigma)$ when there is no ambiguity).

The first and second Zagreb indices were introduced in 1972 [12,13]. For a graph G , they are defined by

$$M_1(G) = \sum_{\varsigma v \in E(G)} [\delta_G(\varsigma) + \delta_G(v)], \quad M_2(G) = \sum_{\varsigma v \in E(G)} \delta_G(\varsigma) \delta_G(v).$$

Došlić [14,15] introduced the corresponding Zagreb coindices, defined as

$$\overline{M}_1(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}(\zeta) + \delta_{\mathfrak{G}}(v)], \quad \overline{M}_2(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} \delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v).$$

The Hyper-Zagreb index was defined by Shirdel et al. [16], and Gutman [17] later introduced its coindex. They are given by

$$HM_1(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta_{\mathfrak{G}}(\zeta) + \delta_{\mathfrak{G}}(v)]^2, \quad \overline{HM}_1(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}(\zeta) + \delta_{\mathfrak{G}}(v)]^2.$$

Furtula and Gutman first introduced the forgotten index (or *F*-index) [18], while the corresponding forgotten coindex was presented by De et al. [19]. They are defined by

$$F(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta_{\mathfrak{G}}^2(\zeta) + \delta_{\mathfrak{G}}^2(v)], \quad \overline{F}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}^2(\zeta) + \delta_{\mathfrak{G}}^2(v)].$$

Gao et al. [20] made important contributions to the study of degree-based indices by:

- Deriving exact formulas for the Zagreb and Hyper-Zagreb indices of carbon nanocones;
- Introducing the second Hyper-Zagreb index, a novel degree-based topological index, given by

$$HM_2(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v)]^2.$$

More recently, Mondal et al. [21] introduced the ND_1 -index (which coincides with the reciprocal Randić index $RR^{-1/2}$), defined as

$$ND_1(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} \sqrt{\delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v)}.$$

Furthermore, the geometric-harmonic index GH was introduced by Usha et al. [22], inspired by the geometric-arithmetic index GA of Vukičević and Furtula [23]. They are defined by

$$GH(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} \frac{(\delta_{\mathfrak{G}}(\zeta) + \delta_{\mathfrak{G}}(v)) \sqrt{\delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v)}}{2}, \quad GA(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} \frac{\sqrt{\delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v)}}{(\delta_{\mathfrak{G}}(\zeta) + \delta_{\mathfrak{G}}(v))/2}.$$

Three novel topological indices were proposed by Alameri et al. [24–26]: the Yemen index, the Yemen coindex and the second Hyper-Zagreb coindex, defined as

$$Y(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta_{\mathfrak{G}}^3(\zeta) + \delta_{\mathfrak{G}}^3(v)], \quad \overline{Y}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}^3(\zeta) + \delta_{\mathfrak{G}}^3(v)],$$

$$\overline{HM}_2(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}(\zeta) \delta_{\mathfrak{G}}(v)]^2.$$

Sethumadhavan et al. [27] studied the *S*-index and its coindex for certain graph operations; they are given by

$$S(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta_{\mathfrak{G}}^4(\zeta) + \delta_{\mathfrak{G}}^4(v)], \quad \overline{S}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta_{\mathfrak{G}}^4(\zeta) + \delta_{\mathfrak{G}}^4(v)].$$

In addition, we introduce four derived *KJ*-type topological descriptors, denoted by KJ_1, KJ_2, KJ_3 and KJ_4 . They are not intended to replace the original indices or coindices, but to test whether selected combinations of edge-based and non-edge-based degree in-

formation can improve benchmark-level descriptor–property associations. The choice of these four quantities is motivated as follows: KJ_1 combines the second Zagreb coindex with the neighborhood-degree contribution ND_1 ; KJ_2 combines fourth-power non-edge degree information with the squared geometric–harmonic descriptor; KJ_3 amplifies the second Hyper-Zagreb coindex, which is sensitive to products of endpoint degrees over non-edges; and KJ_4 amplifies the forgotten coindex, which captures squared-degree contributions over nonadjacent vertex pairs. They are defined by

- $KJ_1(\mathfrak{G}) = \overline{M}_2(\mathfrak{G}) + ND_1(\mathfrak{G})$;
- $KJ_2(\mathfrak{G}) = \overline{S}(\mathfrak{G}) + (GH(\mathfrak{G}))^2$;
- $KJ_3(\mathfrak{G}) = (\overline{HM}_2(\mathfrak{G}))^2$;
- $KJ_4(\mathfrak{G}) = (\overline{F}(\mathfrak{G}))^2$.

These descriptors are introduced for comparative benchmark purposes rather than as universally optimal predictors. In the later sections, we evaluate how they behave relative to classical indices and coindices for the selected octane-isomer properties. Particular attention is paid to whether the derived descriptors provide stronger linear association for some targets and to where their performance remains limited, especially for properties such as $\log P$.

To investigate the connection between molecular structure and physicochemical behavior, we first compute correlation coefficients and simple regression summaries for the octane-isomer benchmark. We then extend the analysis to the larger dataset by using validated multivariate QSPR models and by discussing descriptor redundancy and overfitting control. This two-level application strategy allows the proposed descriptors to be evaluated both in a classical controlled benchmark and in a broader molecular setting.

2. Properties of Topological Indices and Their Boundedness

In this section we collect general degree-based bounds for several Zagreb-type descriptors. The novelty of this part is not the derivation of isolated sharp extremal inequalities for a single index, but rather the organization of ordinary, reduced, expanded, and coindex-related Zagreb-type quantities within a common (α, β) -framework. This unified presentation makes the dependence of each descriptor on the minimum degree, maximum degree, order, and size explicit, and it clarifies how several classical descriptors arise as parameter choices of the same expression.

The resulting bounds should be interpreted as general envelope inequalities. They are useful for the comparison, normalization, and identification of equality behavior. When the upper estimates are expressed in terms of the maximum degree $\Delta(\mathfrak{G})$, the lower and upper envelopes coincide for regular graphs. If $\Delta(\mathfrak{G})$ is replaced by the coarser order bounds $n - 2$, $n - 1$, or n , equality in the order-only estimates generally requires additional extremal structure, for example, the complete graph case. In later applications, this also helps explain why degree-based descriptors may be strongly interrelated when the molecular graphs belong to a structurally homogeneous family, such as octane isomers.

2.1. First (α, β) -General Zagreb Index

The first (α, β) -general Zagreb index is defined by

$$M^{\alpha, \beta}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta(\zeta)^\alpha + \delta(v)^\alpha]^\beta. \quad (1)$$

Several well-known indices are obtained as special cases of (1). To improve citation support and presentation consistency, these specializations are summarized in Table 1.

Table 1. Representative special cases of the first (α, β) -general Zagreb index defined in Equation (1).

Parameters	Descriptor	Specialized Formula	Reference
$\alpha = 1, \beta = 1$	First Zagreb index	$M_1(\mathfrak{G})$	[12,13]
$\alpha = 2, \beta = 1$	Forgotten index	$F(\mathfrak{G})$	[18]
$\alpha = 3, \beta = 1$	Yemen index	$Y(\mathfrak{G})$	[24]
$\alpha = 4, \beta = 1$	S-index	$S(\mathfrak{G})$	[27]
$\alpha = 1, \beta = \frac{1}{2}$	Nirmala index	$N(\mathfrak{G})$	[28]
$\alpha = 1, \beta = -\frac{1}{2}$	Sum-connectivity index	$\sum_{uv \in E(\mathfrak{G})} (d_u + d_v)^{-1/2}$	[29,30]
$\alpha = 2, \beta = \frac{1}{2}$	Sombor index	$SO(\mathfrak{G})$	[31,32]
$\alpha = 3, \beta = \frac{1}{2}$	Dharwad index	$D(\mathfrak{G})$	[33]
$\alpha = 1, \beta = 2$	Hyper-Zagreb index	$HM_1(\mathfrak{G})$	[16]
$\alpha = 2, \beta = 2$	Hyper-forgotten index	$HF(\mathfrak{G})$	[18]

When $\beta = 1$, Equation (1) reduces to the first general Zagreb index, which we denote by

$$M_1^\alpha(\mathfrak{G}) = \sum_{\varsigma v \in E(\mathfrak{G})} [\delta(\varsigma)^\alpha + \delta(v)^\alpha].$$

2.2. Boundedness of the First (α, β) -General Zagreb Index

Let $\mathfrak{G}$ be a graph with vertex set $V(\mathfrak{G})$ and edge set $E(\mathfrak{G})$. For a vertex $\varsigma \in V(\mathfrak{G})$, its degree is denoted by $\delta(\varsigma)$. Let

$$\Delta(\mathfrak{G}) = \max_{\varsigma \in V(\mathfrak{G})} \delta(\varsigma), \quad \delta(\mathfrak{G}) = \min_{\varsigma \in V(\mathfrak{G})} \delta(\varsigma)$$

be the maximum and minimum degree of $\mathfrak{G}$, respectively, and set $m = |E(\mathfrak{G})|$.

For any $\varsigma, v \in V(\mathfrak{G})$ we have

$$2\delta(\mathfrak{G}) \leq \delta(\varsigma) + \delta(v) \leq 2\Delta(\mathfrak{G}). \quad (2)$$

Summing (2) over all edges $\varsigma v \in E(\mathfrak{G})$ yields

$$2m\delta(\mathfrak{G}) \leq \sum_{\varsigma v \in E(\mathfrak{G})} [\delta(\varsigma) + \delta(v)] \leq 2m\Delta(\mathfrak{G}),$$

which gives the lower and upper bounds for the first Zagreb index $M_1(\mathfrak{G})$.

Squaring all sides of (2) and summing over all edges $\varsigma v \in E(\mathfrak{G})$ gives

$$m(2\delta(\mathfrak{G}))^2 \leq \sum_{\varsigma v \in E(\mathfrak{G})} [\delta(\varsigma) + \delta(v)]^2 \leq m(2\Delta(\mathfrak{G}))^2,$$

which yields the boundedness of the Hyper-Zagreb index $HM_1(\mathfrak{G})$.

Since, for every $\varsigma \in V(\mathfrak{G})$,

$$\delta(\mathfrak{G})^2 \leq \delta(\varsigma)^2 \leq \Delta(\mathfrak{G})^2,$$

we obtain

$$2\delta(\mathfrak{G})^2 \leq \delta(\varsigma)^2 + \delta(v)^2 \leq 2\Delta(\mathfrak{G})^2,$$

and hence

$$2m\delta(\mathfrak{G})^2 \leq F(\mathfrak{G}) \leq 2m\Delta(\mathfrak{G})^2.$$

Similarly,

$$2m\delta(\mathfrak{G})^3 \leq Y(\mathfrak{G}) \leq 2m\Delta(\mathfrak{G})^3, \quad 2m\delta(\mathfrak{G})^4 \leq S(\mathfrak{G}) \leq 2m\Delta(\mathfrak{G})^4.$$

In general, for any $\alpha > 0$, we have

$$2m \delta(\mathfrak{G})^\alpha \leq M_1^\alpha(\mathfrak{G}) \leq 2m \Delta(\mathfrak{G})^\alpha.$$

For any $\alpha, \beta > 0$,

$$M^{\alpha, \beta}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta(\zeta)^\alpha + \delta(v)^\alpha]^\beta,$$

and therefore

$$m(2 \delta(\mathfrak{G})^\alpha)^\beta \leq M^{\alpha, \beta}(\mathfrak{G}) \leq m(2 \Delta(\mathfrak{G})^\alpha)^\beta.$$

When $\alpha > 0$ and $\beta < 0$, the inequalities reverse:

$$m(2 \Delta(\mathfrak{G})^\alpha)^\beta \leq M^{\alpha, \beta}(\mathfrak{G}) \leq m(2 \delta(\mathfrak{G})^\alpha)^\beta.$$

Thus, the parameters $\delta(\mathfrak{G})$ and $\Delta(\mathfrak{G})$ provide a universal envelope for the first (α, β) -general Zagreb index. In particular, equality is attained in these bounds for regular graphs, where $\delta(\mathfrak{G}) = \Delta(\mathfrak{G})$. In particular, for $\alpha = \beta = 2$, we obtain the boundedness of the Hyper-forgotten index $HF(\mathfrak{G})$.

The above inequalities are intentionally coarse and general: they are not meant to be sharp for every graph class, but rather to provide a unified benchmark that specializes immediately to many familiar descriptors.

In the following subsections, we study the boundedness of the reduced and expanded first (α, β) -general Zagreb indices.

2.3. Reduced First (α, β) -General Zagreb Index

The reduced first (α, β) -general Zagreb index is defined by

$$RM_1^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [(\delta(\zeta) - 1)^\alpha + (\delta(v) - 1)^\alpha]^\beta. \quad (3)$$

When $\alpha \in \{1, 2, 3, 4\}$ and $\beta = 1$, Equation (3) yields, respectively, the reduced first Zagreb index, the reduced forgotten index, the reduced Yemen index and the reduced S -index.

For $\beta = 1$, (3) gives the reduced first general Zagreb index, denoted by $RM_1^{(\alpha)}(\mathfrak{G})$. When $\alpha = 1$ and $\beta \in \mathbb{R}$, (3) defines the reduced general sum-connectivity-type index.

Moreover:

- For $\beta = \frac{1}{2}$, Equation (3) yields the reduced general Sombor index;
- For $\alpha = 1$ and $\beta = \frac{1}{2}$, Equation (3) yields the reduced Nirmala index.

2.4. Boundedness of the Reduced First (α, β) -General Zagreb Index

Let $n = |V(\mathfrak{G})|$ and $m = |E(\mathfrak{G})|$. Recall that $\delta(\mathfrak{G})$ and $\Delta(\mathfrak{G})$ denote the minimum and maximum degree of $\mathfrak{G}$, respectively. Then

$$\delta(\mathfrak{G}) - 1 \leq \delta(\zeta) - 1, \delta(v) - 1 \leq \Delta(\mathfrak{G}) - 1 \leq n - 2.$$

For $\alpha > 0$ this implies

$$(\delta(\mathfrak{G}) - 1)^\alpha \leq (\delta(\zeta) - 1)^\alpha, (\delta(v) - 1)^\alpha \leq (\Delta(\mathfrak{G}) - 1)^\alpha,$$

and hence

$$2(\delta(\mathfrak{G}) - 1)^\alpha \leq (\delta(\zeta) - 1)^\alpha + (\delta(v) - 1)^\alpha \leq 2(\Delta(\mathfrak{G}) - 1)^\alpha. \quad (4)$$

Summing (4) over all edges $\zeta v \in E(\mathfrak{G})$ yields

$$2m(\delta(\mathfrak{G}) - 1)^\alpha \leq RM_1^{(\alpha)}(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) - 1)^\alpha.$$

For $\alpha, \beta > 0$ we obtain

$$m[2(\delta(\mathfrak{G}) - 1)^\alpha]^\beta \leq RM_1^{(\alpha, \beta)}(\mathfrak{G}) \leq m[2(\Delta(\mathfrak{G}) - 1)^\alpha]^\beta. \quad (5)$$

When $\alpha > 0$ and $\beta < 0$, these inequalities are meaningful only under the additional assumption $\delta(\mathfrak{G}) > 1$, because, otherwise, the term $(\delta(\mathfrak{G}) - 1)^\alpha$ vanishes and negative powers are not defined. This restriction is important for molecular applications, since many chemical graphs—including trees such as octane-isomer graphs—contain pendant vertices and hence have a minimum degree of one. Under the assumption $\delta(\mathfrak{G}) > 1$, the inequalities reverse:

$$m[2(\Delta(\mathfrak{G}) - 1)^\alpha]^\beta \leq RM_1^{(\alpha, \beta)}(\mathfrak{G}) \leq m[2(\delta(\mathfrak{G}) - 1)^\alpha]^\beta. \quad (6)$$

Since $\Delta(\mathfrak{G}) - 1 \leq n - 2$, the corresponding order-only upper bounds follow immediately by replacing $\Delta(\mathfrak{G}) - 1$ with $n - 2$. These coarser bounds do not generally merge for every regular graph; equality in the order-only version requires, in addition, $\Delta(\mathfrak{G}) = n - 1$, as in the complete graph case.

Several important reduced indices appear as special cases of (5)–(6):

- $\alpha = \beta = 1$: Bounds for the first reduced Zagreb index $RM_1(\mathfrak{G})$;
- $\alpha = 2, \beta = 1$: Bounds for the reduced forgotten index $RF(\mathfrak{G})$;
- $\alpha = 3, \beta = 1$: Bounds for the reduced Yemen index $RY(\mathfrak{G})$;
- $\alpha = 4, \beta = 1$: Bounds for the reduced S-index $RS(\mathfrak{G})$;
- $\alpha = 2, \beta = \frac{1}{2}$: Bounds for the reduced Sombor index $RSO(\mathfrak{G})$:

$$\sqrt{2}m(\delta(\mathfrak{G}) - 1) \leq RSO(\mathfrak{G}) \leq \sqrt{2}m(\Delta(\mathfrak{G}) - 1);$$

- $\alpha = 1, \beta = -\frac{1}{2}$, with $\delta(\mathfrak{G}) > 1$: Bounds for the reduced sum-connectivity index $RSC(\mathfrak{G})$:

$$\frac{m}{\sqrt{2(\Delta(\mathfrak{G}) - 1)}} \leq RSC(\mathfrak{G}) \leq \frac{m}{\sqrt{2(\delta(\mathfrak{G}) - 1)}};$$

- For general $\alpha > 0$, with $\delta(\mathfrak{G}) > 1$, the reduced inverse Nirmala index satisfies

$$\frac{m}{(2(\Delta(\mathfrak{G}) - 1))^{\alpha/2}} \leq RN^{(\alpha)}(\mathfrak{G}) \leq \frac{m}{(2(\delta(\mathfrak{G}) - 1))^{\alpha/2}}.$$

Accordingly, the reduced-index bounds with positive exponents apply in full generality, whereas the negative-exponent cases should be understood as restricted to graphs with a minimum degree greater than one.

2.5. Expanded First (α, β) -General Zagreb Index

The expanded first (α, β) -general Zagreb index is defined by

$$EM^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} \left[(\delta(\zeta) + 1)^\alpha + (\delta(v) + 1)^\alpha \right]^\beta. \quad (7)$$

For $\beta = 1$ and $\alpha \in \{1, 2, 3, 4\}$, Equation (7) yields, respectively, the expanded first Zagreb index, the expanded forgotten index, the expanded Yemen index and the expanded S-index.

The general form of the expanded Sombor index is obtained by setting $\beta = \frac{1}{2}$. When $\alpha = 1$ and $\beta = \frac{1}{2}$, Equation (7) gives the expanded Nirmala index.

2.6. Boundedness of the Expanded First (α, β) -General Zagreb Index

Let $n = |V(\mathfrak{G})|$ and $m = |E(\mathfrak{G})|$, and let $\delta(\mathfrak{G})$ and $\Delta(\mathfrak{G})$ denote the minimum and maximum degrees of $\mathfrak{G}$, respectively. For $\alpha, \beta > 0$ we have

$$(\delta(\mathfrak{G}) + 1)^\alpha \leq (\delta(\zeta) + 1)^\alpha, (\delta(v) + 1)^\alpha \leq (\Delta(\mathfrak{G}) + 1)^\alpha,$$

and hence

$$2(\delta(\mathfrak{G}) + 1)^\alpha \leq (\delta(\zeta) + 1)^\alpha + (\delta(v) + 1)^\alpha \leq 2(\Delta(\mathfrak{G}) + 1)^\alpha.$$

Raising all terms to the power $\beta > 0$ gives

$$[2(\delta(\mathfrak{G}) + 1)^\alpha]^\beta \leq [(\delta(\zeta) + 1)^\alpha + (\delta(v) + 1)^\alpha]^\beta \leq [2(\Delta(\mathfrak{G}) + 1)^\alpha]^\beta. \quad (8)$$

Summing (8) over all edges $\zeta v \in E(\mathfrak{G})$ yields the bounds for the expanded first (α, β) -general Zagreb index (for $\alpha, \beta > 0$):

$$m[2(\delta(\mathfrak{G}) + 1)^\alpha]^\beta \leq EM^{(\alpha, \beta)}(\mathfrak{G}) \leq m[2(\Delta(\mathfrak{G}) + 1)^\alpha]^\beta. \quad (9)$$

For $\alpha > 0$ and $\beta < 0$, the inequalities reverse:

$$m(2(\Delta(\mathfrak{G}) + 1)^\alpha)^\beta \leq EM^{(\alpha, \beta)}(\mathfrak{G}) \leq m[2(\delta(\mathfrak{G}) + 1)^\alpha]^\beta. \quad (10)$$

As in the previous subsection, these estimates show that the expanded first (α, β) -general Zagreb index is controlled by the extremal degree parameters of the graph. Moreover, equality again occurs for regular graphs, where all vertex degrees coincide, and the lower and upper bounds merge. If one replaces $\Delta(\mathfrak{G}) + 1$ by the coarser bound n , the resulting order-only estimates remain valid, but equality for all regular graphs is no longer guaranteed; it occurs only under stronger conditions, such as $\Delta(\mathfrak{G}) = n - 1$. These bounds are included mainly to complete the parallel treatment of the ordinary, reduced, and expanded forms within one common framework.

As special cases of (9) and (10), we obtain the following bounds for several expanded indices:

$$\begin{aligned} 2m(\delta(\mathfrak{G}) + 1) &\leq EM_1(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) + 1), \\ 2m(\delta(\mathfrak{G}) + 1)^2 &\leq EF(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) + 1)^2, \\ 2m(\delta(\mathfrak{G}) + 1)^3 &\leq EY(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) + 1)^3, \\ 2m(\delta(\mathfrak{G}) + 1)^4 &\leq ES(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) + 1)^4, \\ 2m(\delta(\mathfrak{G}) + 1)^\alpha &\leq EM^{(\alpha)}(\mathfrak{G}) \leq 2m(\Delta(\mathfrak{G}) + 1)^\alpha, \\ \sqrt{2}m(\delta(\mathfrak{G}) + 1) &\leq ESO(\mathfrak{G}) \leq \sqrt{2}m(\Delta(\mathfrak{G}) + 1), \\ \frac{m}{\sqrt{2(\Delta(\mathfrak{G}) + 1)}} &\leq ESC(\mathfrak{G}) \leq \frac{m}{\sqrt{2(\delta(\mathfrak{G}) + 1)}}. \end{aligned}$$

We emphasize that these inequalities are broad envelope bounds rather than sharp extremal characterizations for all graph classes. Their main role in the present manuscript is to place several expanded descriptors into the same degree-based scheme used for the standard and reduced Zagreb-type forms.

3. Second (α, β) -General Zagreb Index

The second (α, β) -general Zagreb index is defined by

$$M_2^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [\delta(\zeta)^\alpha \delta(v)^\alpha]^\beta. \quad (11)$$

For $\beta = 1$, Equation (11) reduces to the second general Zagreb index $M_2^{(\alpha)}(\mathfrak{G})$. In particular, the following special cases are obtained from (11):

- $\alpha = \beta = 1$: The second Zagreb index $M_2(\mathfrak{G})$;
- $\alpha = 2, \beta = 1$: The second Hyper-Zagreb index $HM_2(\mathfrak{G})$;
- $\alpha = 1, \beta = -\frac{1}{2}$: The general Randić index $R^{-1/2}(\mathfrak{G})$;
- $\alpha = 1, \beta = \frac{1}{2}$: The reciprocal Randić index (or inverse product-connectivity index).

3.1. Boundedness of the Second (α, β) -General Zagreb Index

Let $\mathfrak{G}$ be a graph with vertex set $V(\mathfrak{G})$, edge set $E(\mathfrak{G})$, $n = |V(\mathfrak{G})|$ and $m = |E(\mathfrak{G})|$. As before, denote the maximum and minimum degrees by

$$\Delta(\mathfrak{G}) = \max\{\delta(\zeta) : \zeta \in V(\mathfrak{G})\}, \quad \delta(\mathfrak{G}) = \min\{\delta(\zeta) : \zeta \in V(\mathfrak{G})\}.$$

From

$$\delta(\mathfrak{G}) \leq \delta(\zeta) \leq \Delta(\mathfrak{G}), \quad \delta(\mathfrak{G}) \leq \delta(v) \leq \Delta(\mathfrak{G}),$$

we obtain

$$\delta(\mathfrak{G})^2 \leq \delta(\zeta) \delta(v) \leq \Delta(\mathfrak{G})^2. \quad (*)$$

Summing $(*)$ over all edges $\zeta v \in E(\mathfrak{G})$ yields

$$m \delta(\mathfrak{G})^2 \leq \sum_{\zeta v \in E(\mathfrak{G})} \delta(\zeta) \delta(v) \leq m \Delta(\mathfrak{G})^2,$$

which gives the boundedness of the second Zagreb index:

$$m \delta(\mathfrak{G})^2 \leq M_2(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^2.$$

Squaring all terms in $(*)$ gives

$$\delta(\mathfrak{G})^4 \leq [\delta(\zeta) \delta(v)]^2 \leq \Delta(\mathfrak{G})^4,$$

and summing over all edges implies the boundedness of the second Hyper-Zagreb index:

$$m \delta(\mathfrak{G})^4 \leq HM_2(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^4.$$

More generally, from $(*)$, for any $\alpha, \beta > 0$, we have

$$\delta(\mathfrak{G})^{2\alpha\beta} \leq [\delta(\zeta)^\alpha \delta(v)^\alpha]^\beta \leq \Delta(\mathfrak{G})^{2\alpha\beta},$$

and hence

$$m \delta(\mathfrak{G})^{2\alpha\beta} \leq \sum_{\zeta v \in E(\mathfrak{G})} [\delta(\zeta)^\alpha \delta(v)^\alpha]^\beta = M_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^{2\alpha\beta}. \quad (12)$$

Therefore, for all $\alpha, \beta > 0$,

$$m \delta(\mathfrak{G})^{2\alpha\beta} \leq M_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^{2\alpha\beta}.$$

Thus, as in the first generalized Zagreb family, the second (α, β) -general Zagreb index is bounded by the extremal degree parameters alone. For regular graphs these bounds coincide, so the lower and upper estimates are attained simultaneously.

Some important special cases of (12) are:

- $\alpha = 2, \beta = 1$:

$$m \delta(\mathfrak{G})^4 \leq HM_2(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^4;$$

- $\alpha = 1, \beta = \frac{1}{2}$:

$$m \delta(\mathfrak{G}) \leq \sum_{\zeta v \in E(\mathfrak{G})} [\delta(\zeta) \delta(v)]^{1/2} \leq m \Delta(\mathfrak{G}),$$

which gives the boundedness of the product-connectivity (reciprocal Randić) index;

- $\alpha = 1, \beta = -\frac{1}{2}$:

$$\Delta(\mathfrak{G})^{-1} \leq [\delta(\zeta) \delta(v)]^{-1/2} \leq \delta(\mathfrak{G})^{-1},$$

so that

$$m \Delta(\mathfrak{G})^{-1} \leq R^{-1/2}(\mathfrak{G}) \leq m \delta(\mathfrak{G})^{-1}.$$

For the second general Zagreb index $M_2^{(\alpha)}(\mathfrak{G})$ (i.e., $\beta = 1$), (12) gives

$$m \delta(\mathfrak{G})^{2\alpha} \leq M_2^{(\alpha)}(\mathfrak{G}) \leq m \Delta(\mathfrak{G})^{2\alpha}, \quad \alpha > 0.$$

If $\alpha > 0$ and $\beta < 0$, then the inequalities reverse:

$$m \Delta(\mathfrak{G})^{2\alpha\beta} \leq M_2^{(\alpha,\beta)}(\mathfrak{G}) \leq m \delta(\mathfrak{G})^{2\alpha\beta}.$$

Since $\Delta(\mathfrak{G}) \leq n - 1 = |V(\mathfrak{G})| - 1$, we obtain

$$m \delta(\mathfrak{G})^2 \leq M_2(\mathfrak{G}) \leq m (n - 1)^2,$$

$$m \delta(\mathfrak{G})^4 \leq HM_2(\mathfrak{G}) \leq m (n - 1)^4.$$

More generally, for $\alpha, \beta > 0$,

$$m \delta(\mathfrak{G})^{2\alpha\beta} \leq M_2^{(\alpha,\beta)}(\mathfrak{G}) \leq m (n - 1)^{2\alpha\beta},$$

and for $\alpha > 0, \beta < 0$,

$$m (n - 1)^{2\alpha\beta} \leq M_2^{(\alpha,\beta)}(\mathfrak{G}) \leq m \delta(\mathfrak{G})^{2\alpha\beta}.$$

As before, these inequalities should be viewed as general envelope bounds rather than sharp extremal results for every graph family. Their role here is to complete the parallel treatment of the first and second Zagreb-type frameworks within a single degree-based scheme.

3.2. Reduced and Expanded (α, β) Second General Zagreb Indices

We now define the reduced and expanded versions of the second (α, β) -general Zagreb index:

1. The reduced second (α, β) -general Zagreb index is defined by

$$RM_2^{(\alpha,\beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [(\delta(\zeta) - 1)^\alpha (\delta(v) - 1)^\alpha]^\beta.$$

2. The expanded second (α, β) -general Zagreb index is defined by

$$EM_2^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} \left[(\delta(\zeta) + 1)^\alpha (\delta(v) + 1)^\alpha \right]^\beta.$$

3.3. Boundedness of the Reduced (α, β) Second General Zagreb Index

Let $n = |V(\mathfrak{G})|$ and $m = |E(\mathfrak{G})|$, and let $\delta(\mathfrak{G})$ and $\Delta(\mathfrak{G})$ denote the minimum and maximum degrees of $\mathfrak{G}$, respectively. For $\alpha, \beta > 0$ we have

$$(\delta(\mathfrak{G}) - 1)^\alpha \leq (\delta(\zeta) - 1)^\alpha \leq (\Delta(\mathfrak{G}) - 1)^\alpha,$$

$$(\delta(\mathfrak{G}) - 1)^\alpha \leq (\delta(v) - 1)^\alpha \leq (\Delta(\mathfrak{G}) - 1)^\alpha,$$

and hence

$$(\delta(\mathfrak{G}) - 1)^{2\alpha} \leq (\delta(\zeta) - 1)^\alpha (\delta(v) - 1)^\alpha \leq (\Delta(\mathfrak{G}) - 1)^{2\alpha}.$$

Summing over all edges $\zeta v \in E(\mathfrak{G})$ and raising to the power β gives

$$m (\delta(\mathfrak{G}) - 1)^{2\alpha\beta} \leq RM_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m (\Delta(\mathfrak{G}) - 1)^{2\alpha\beta}. \quad (13)$$

From (13), the following bounds for some well-known reduced indices are obtained:

- $\alpha = \beta = 1$:

$$m(\delta(\mathfrak{G}) - 1)^2 \leq RM_2(\mathfrak{G}) \leq m(\Delta(\mathfrak{G}) - 1)^2;$$

- $\alpha = 2, \beta = 1$:

$$m(\delta(\mathfrak{G}) - 1)^4 \leq RHM_2(\mathfrak{G}) \leq m(\Delta(\mathfrak{G}) - 1)^4;$$

- $\beta = 1$ (general reduced second Zagreb index):

$$m(\delta(\mathfrak{G}) - 1)^{2\alpha} \leq RM_2^{(\alpha)}(\mathfrak{G}) \leq m(\Delta(\mathfrak{G}) - 1)^{2\alpha}.$$

For $\alpha > 0$ and $\beta < 0$, the negative-power case is meaningful only when $\delta(\mathfrak{G}) > 1$. This restriction should be kept in mind in molecular applications, because many chemical graphs with pendant vertices do not satisfy it. Under the assumption $\delta(\mathfrak{G}) > 1$, the inequalities reverse, and we obtain

$$m (\Delta(\mathfrak{G}) - 1)^{2\alpha\beta} \leq RM_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m (\delta(\mathfrak{G}) - 1)^{2\alpha\beta}.$$

In particular, for $\alpha = 1$ and $\beta = -\frac{1}{2}$,

$$\frac{m}{\Delta(\mathfrak{G}) - 1} \leq RR^{-\frac{1}{2}}(\mathfrak{G}) \leq \frac{m}{\delta(\mathfrak{G}) - 1},$$

which gives the bounds for the reduced reciprocal Randić-type index. As above, replacing $\Delta(\mathfrak{G}) - 1$ by $n - 2$ gives a valid but weaker order-only estimate, whose equality for regular graphs requires the additional condition $\Delta(\mathfrak{G}) = n - 1$.

Likewise, the modified reduced second (α, β) -general Zagreb index is considered only under the same restriction $\delta(\mathfrak{G}) > 1$. For the modified reduced second (α, β) -general Zagreb index we have

$$m (\Delta(\mathfrak{G}) - 1)^{-2\alpha\beta} \leq {}^mRM_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m (\delta(\mathfrak{G}) - 1)^{-2\alpha\beta} \quad \forall \alpha > 0, \beta < 0.$$

3.4. Boundedness of the Expanded (α, β) Second General Zagreb Index

Recall that

$$EM_2^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \in E(\mathfrak{G})} [(\delta(\zeta) + 1)^\alpha (\delta(v) + 1)^\alpha]^\beta.$$

Let $n = |V(\mathfrak{G})|$ and $m = |E(\mathfrak{G})|$, and let $\delta(\mathfrak{G})$ and $\Delta(\mathfrak{G})$ denote the minimum and maximum degrees of $\mathfrak{G}$, respectively. For $\alpha, \beta > 0$ we have

$$\delta(\mathfrak{G}) + 1 \leq \delta(\zeta) + 1, \delta(v) + 1 \leq \Delta(\mathfrak{G}) + 1,$$

and hence

$$(\delta(\mathfrak{G}) + 1)^\alpha \leq (\delta(\zeta) + 1)^\alpha, (\delta(v) + 1)^\alpha \leq (\Delta(\mathfrak{G}) + 1)^\alpha,$$

and therefore

$$(\delta(\mathfrak{G}) + 1)^{2\alpha} \leq (\delta(\zeta) + 1)^\alpha (\delta(v) + 1)^\alpha \leq (\Delta(\mathfrak{G}) + 1)^{2\alpha}.$$

Raising to the power $\beta > 0$ and summing over all edges $\zeta v \in E(\mathfrak{G})$ yields

$$m (\delta(\mathfrak{G}) + 1)^{2\alpha\beta} \leq EM_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m (\Delta(\mathfrak{G}) + 1)^{2\alpha\beta},$$

for all $\alpha, \beta > 0$.

In particular, we obtain the following bounds:

- For $\alpha = 1$ and $\beta = 1$:

$$m(\delta(\mathfrak{G}) + 1)^2 \leq EM_2(\mathfrak{G}) \leq m(\Delta(\mathfrak{G}) + 1)^2.$$

- For $\alpha = 2$ and $\beta = 1$:

$$m(\delta(\mathfrak{G}) + 1)^4 \leq EHM_2(\mathfrak{G}) \leq m(\Delta(\mathfrak{G}) + 1)^4.$$

For $\alpha > 0$ and $\beta < 0$, the inequalities reverse, and we have

$$m (\Delta(\mathfrak{G}) + 1)^{2\alpha\beta} \leq EM_2^{(\alpha, \beta)}(\mathfrak{G}) \leq m (\delta(\mathfrak{G}) + 1)^{2\alpha\beta}.$$

In particular, for the expanded Randić-type index ($\alpha = 1, \beta = -\frac{1}{2}$), we get

$$\frac{m}{\Delta(\mathfrak{G}) + 1} \leq ER^{-\frac{1}{2}}(\mathfrak{G}) \leq \frac{m}{\delta(\mathfrak{G}) + 1}.$$

These expanded bounds mirror the reduced and ordinary cases and complete the parallel hierarchy of second Zagreb-type descriptors considered in this study. Since $\Delta(\mathfrak{G}) + 1 \leq n$, the corresponding order-only bounds remain valid after replacing $\Delta(\mathfrak{G}) + 1$ by n , but equality in those coarser bounds is generally restricted to complete graphs rather than all regular graphs.

The first and second, reduced and expanded, (α, β) -general Zagreb coindices are defined as follows:

$$\overline{M}_1^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} (\delta(\zeta)^\alpha + \delta(v)^\alpha)^\beta,$$

$$\overline{M}_2^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [\delta(\zeta)^\alpha \delta(v)^\alpha]^\beta,$$

$$\overline{RM}_1^{(\alpha, \beta)}(\mathfrak{G}) = \sum_{\zeta v \notin E(\mathfrak{G})} [(\delta(\zeta) - 1)^\alpha + (\delta(v) - 1)^\alpha]^\beta,$$

$$\begin{aligned}\overline{RM}_2^{(\alpha,\beta)}(\mathfrak{G}) &= \sum_{\zeta v \notin E(\mathfrak{G})} [(\delta(\zeta) - 1)^\alpha (\delta(v) - 1)^\alpha]^\beta, \\ \overline{EM}_1^{(\alpha,\beta)}(\mathfrak{G}) &= \sum_{\zeta v \notin E(\mathfrak{G})} [(\delta(\zeta) + 1)^\alpha + (\delta(v) + 1)^\alpha]^\beta, \\ \overline{EM}_2^{(\alpha,\beta)}(\mathfrak{G}) &= \sum_{\zeta v \notin E(\mathfrak{G})} [(\delta(\zeta) + 1)^\alpha (\delta(v) + 1)^\alpha]^\beta.\end{aligned}$$

We record these coindex analogues for completeness, since they parallel the index families studied above and provide a convenient notation for the later application section.

To support our primary findings, we recall the following lemma (see, for example, [15,17,19,24,25,27]).

Lemma 1. *Let $\mathfrak{G}$ be a simple graph with n vertices and m edges. Then*

- $\overline{M}_1(\mathfrak{G}) = 2m(n - 1) - M_1(\mathfrak{G})$;
- $\overline{F}(\mathfrak{G}) = (n - 1)M_1(\mathfrak{G}) - F(\mathfrak{G})$;
- $\overline{Y}(\mathfrak{G}) = (n - 1)F(\mathfrak{G}) - Y(\mathfrak{G})$;
- $\overline{S}(\mathfrak{G}) = (n - 1)Y(\mathfrak{G}) - S(\mathfrak{G})$;
- $\overline{M}_2(\mathfrak{G}) = 2m^2 - M_2(\mathfrak{G}) - \frac{1}{2}M_1(\mathfrak{G})$;
- $\overline{HM}_1(\mathfrak{G}) = 4m^2 + (n - 2)M_1(\mathfrak{G}) - HM_1(\mathfrak{G})$;
- $\overline{HM}_2(\mathfrak{G}) = \frac{1}{2}(M_1(\mathfrak{G})^2 - Y(\mathfrak{G})) - HM_2(\mathfrak{G})$.

Lemma 1 is especially useful in the octane-isomer setting, where several coindices can be computed efficiently from the corresponding indices. This observation also helps explain why some index/coindex comparisons in the application section should be interpreted within the structural constraints of the benchmark dataset.

4. Results and Discussion

In this section, we compute several topological parameters for the graphs of octane isomers (see Figure 1) and examine their linear association with selected physicochemical properties, namely, the acentric factor (AF), motor octane number (MON), entropy (S), molar refraction (MR), and the octanol–water partition coefficient ($\log P$). Our aim is not to claim universal predictive optimality, but rather to benchmark the relative behavior of the considered descriptors on a standard octane-isomer dataset. Accordingly, the statistical discussion is interpreted in a dataset-specific sense, and stronger predictive claims are reserved for the later regression summary and limitations discussion.

We compare the obtained results with those reported in earlier studies [1–4,21,23,34–39]. For additional background on the use of topological descriptors in QSPR/QSAR modeling, the interested reader is referred to [5,40–45]. The octane isomers considered in this study form a small and structurally homogeneous benchmark family, since their molecular graphs are trees with the same order and size. This setting provides a useful basis for controlled comparison of degree-based descriptors, while the resulting QSPR observations are best interpreted within the scope of this benchmark family.

Some physicochemical characteristics of the octane isomers considered in this study are summarized in the dataset given in Table 2 [1,5]. Missing values are indicated by “–” and are excluded whenever a correlation or regression model is computed for the corresponding property. We first compute the classical topological indices of the octane isomers, then determine the corresponding topological coindices, and finally evaluate the newly introduced improved topological coindices.

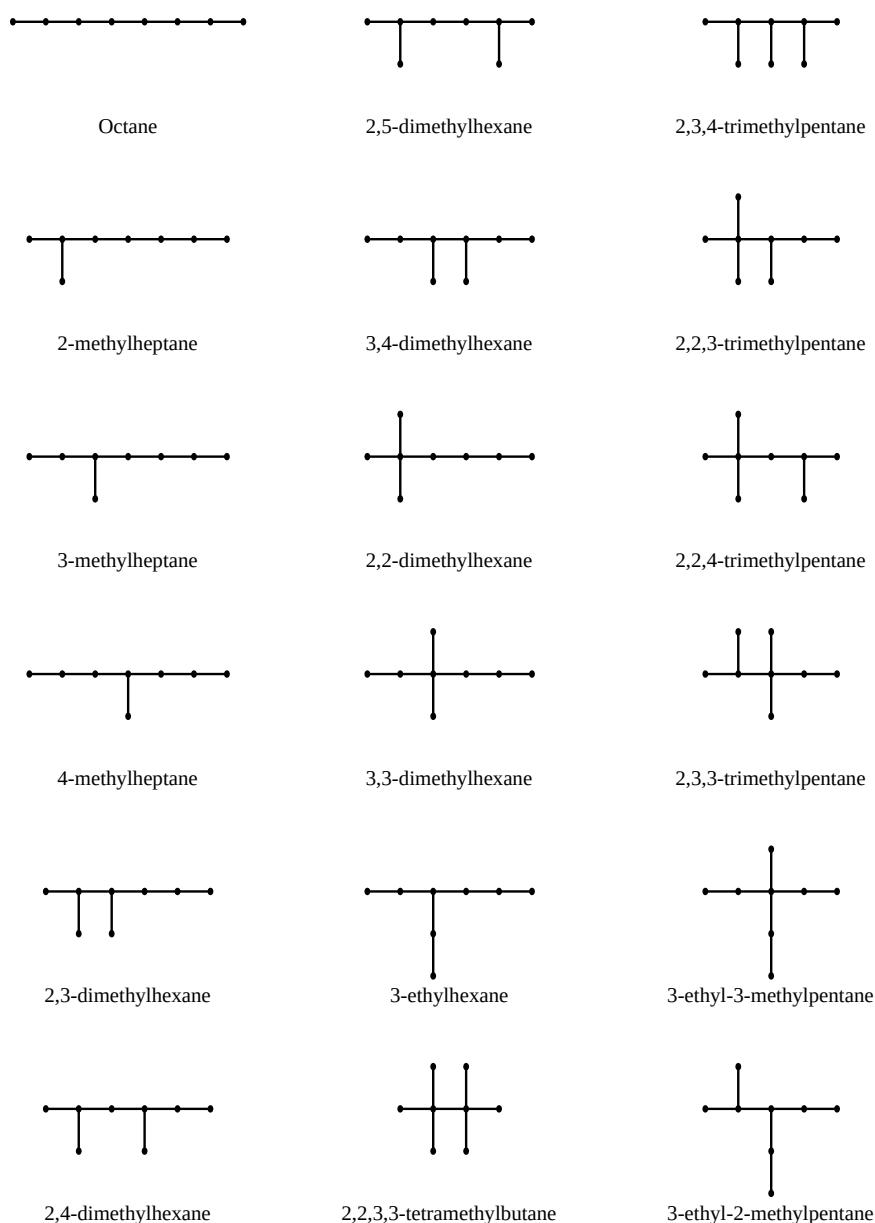


Figure 1. Graphs of the 18 octane isomers considered in the benchmark dataset.

4.1. Correlation Between Specific Topological Indices and the Physicochemical Properties of Octane Isomers

In this subsection, we compute several classical topological indices for the octane isomers and investigate their correlation with the physicochemical properties listed in Table 2 (see also Figure 2). Table 3 reports the Pearson correlation coefficients r between each index and the selected properties. For the AF, entropy, and $\log P$, the sample size is 18; for MR, it is 17 because one value is unavailable; for the MON, it is 16 because two values are unavailable. The corresponding coefficients of determination R^2 for the selected regression models are reported later in the application section.

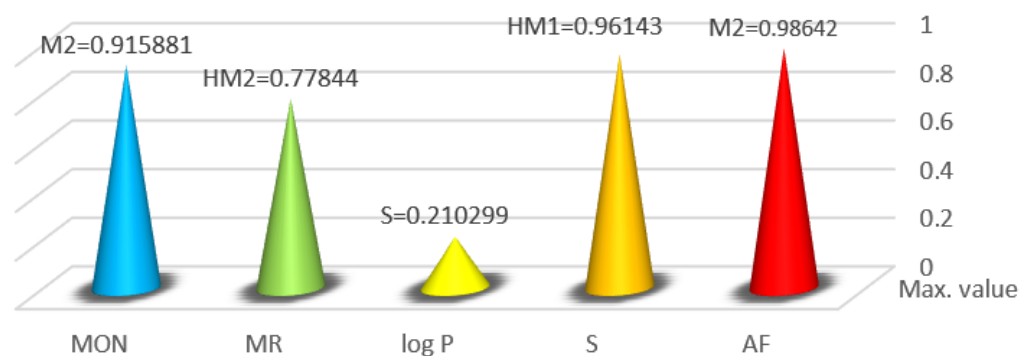


Figure 2. Correlation patterns between selected classical topological indices and the physicochemical properties of octane isomers.

Among the classical indices, the second Zagreb index M_2 shows the strongest linear association with the acentric factor (AF), with $|r| = 0.98642$, corresponding to $R^2 \approx 0.9730$. The same index also gives the strongest association with the motor octane number (MON), with $|r| = 0.915881$, corresponding to $R^2 \approx 0.8388$. In addition, the Hyper-Zagreb index HM_1 yields the highest correlation with the entropy property, with $|r| = 0.96143$ and $R^2 \approx 0.9243$. By contrast, the absolute correlations with $\log P$ are relatively low for all classical indices listed in Table 3, suggesting that this property is less effectively captured by these basic degree-based descriptors within the octane benchmark. Therefore, in the present dataset, the classical indices appear substantially more informative for the AF, MON, and entropy than for the $\log P$ or MR.

Table 2. Physicochemical properties (PCPs) of octane isomers. Missing entries are denoted by “–”.

Molecule	AF	S	$\log P$	MR	MON
n-octane	0.397898	111.67	3.67	39.1922	–
2-methyl-heptane	0.377916	109.84	3.61	39.2316	23.1
3-methyl-heptane	0.371002	111.26	3.61	39.1001	35
4-methyl-heptane	0.371504	109.32	3.61	39.1174	39
3-ethyl-hexane	0.362472	109.43	3.61	38.9441	52.4
2,2-dimethyl-hexane	0.339426	103.42	3.65	39.2525	77.4
2,3-dimethyl-hexane	0.348247	108.02	3.54	38.9808	78.9
2,4-dimethyl-hexane	0.344223	106.98	3.54	39.13	69.9
2,5-dimethyl-hexane	0.356830	105.72	3.54	39.2596	55.7
3,3-dimethyl-hexane	0.322596	104.74	3.65	39.0087	83.4
3,4-dimethyl-hexane	0.340345	106.59	3.54	38.8453	81.7
2-methyl-3-ethyl-pentane	0.332433	106.06	3.54	38.8362	88.1
3-methyl-3-ethyl-pentane	0.306899	101.48	3.65	38.7171	88.7
2,2,3-trimethyl-pentane	0.300816	101.31	3.58	38.9249	99.9
2,2,4-trimethyl-pentane	0.305370	104.09	3.58	39.2617	100
2,3,3-trimethyl-pentane	0.293177	102.06	3.58	38.7617	99.4
2,3,4-trimethyl-pentane	0.317422	102.39	3.48	38.8681	95.9
2,2,3,3-tetramethylbutane	0.255294	93.06	3.62	–	–

Table 3. Topological indices of octane isomers used in the benchmark comparison.

Molecule	M_1	M_2	F	Y	S	HM_1	HM_2
n-octane	26	24	50	98	194	98	88
2-methyl-heptane	28	26	62	148	374	114	106
3-methyl-heptane	28	27	62	148	374	116	121
4-methyl-heptane	28	27	62	148	374	116	121
3-ethyl-hexane	28	28	62	148	374	118	136
2,2-dimethyl-hexane	32	30	92	308	1124	152	124
2,3-dimethyl-hexane	30	30	74	198	554	134	164
2,4-dimethyl-hexane	30	29	74	198	554	132	139
2,5-dimethyl-hexane	30	28	74	198	554	130	124
3,3-dimethyl-hexane	32	32	92	308	1124	156	184
3,4-dimethyl-hexane	30	31	74	198	554	136	179
2-methyl-3-ethyl-pentane	30	31	74	198	554	136	179
3-methyl-3-ethyl-pentane	32	34	92	308	1124	160	220
2,2,3-trimethyl-pentane	34	35	104	358	1304	174	241
2,2,4-trimethyl-pentane	34	32	104	358	1304	168	166
2,3,3-trimethyl-pentane	34	36	104	358	1304	176	262
2,3,4-trimethyl-pentane	32	33	86	248	523	152	257
2,2,3,3-tetramethylbutane	38	40	134	518	2054	214	352

4.2. Relationships Between Specific Topological Coindices and the Physicochemical Properties of Octane Isomers

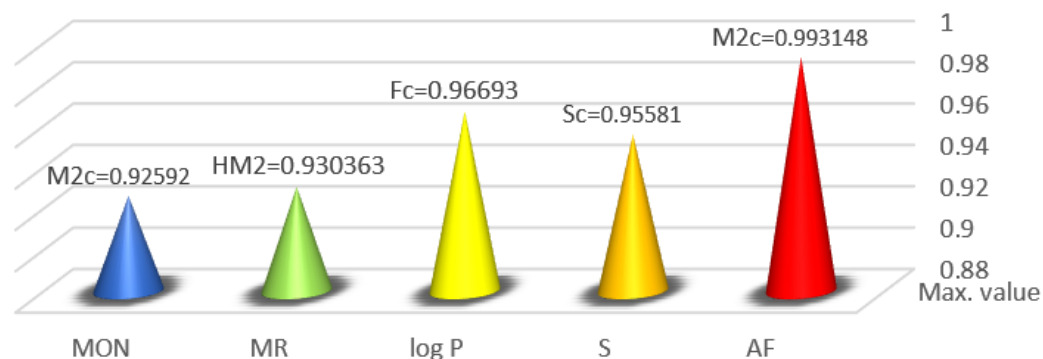
In this subsection, we compute the topological coindices of the octane isomers and analyze their correlation with the physicochemical properties listed in Table 2 (see Figure 3). The values of the coindices are reported in Table 4, and the corresponding Pearson correlation coefficients with the physicochemical properties are given in Table 5.

Table 4. Pearson correlation coefficients r between PCPs and classical topological indices for octane isomers.

PCPs	AF	S	$\log P$	MR	MON
M_1	−0.97306	−0.95429	−0.07933	−0.32377	0.881268
M_2	−0.98642	−0.94169	−0.13239	−0.66878	0.915881
F	−0.96505	−0.95272	0.00550	−0.30398	0.850421
Y	−0.95085	−0.94478	0.08664	−0.28129	0.812371
S	−0.91166	−0.90735	0.21030	−0.22701	0.728204
HM_1	−0.98291	−0.96143	−0.03276	−0.41460	0.889938
HM_2	−0.92284	−0.89933	−0.19840	−0.77844	0.812491

Table 5. Topological coindices of octane isomers used in the benchmark comparison.

Molecule	$\overline{M}_1$	$\overline{M}_2$	$\overline{F}$	$\overline{Y}$	$\overline{S}$	$\overline{HM}_1$	$\overline{HM}_2$
n-octane	72	61	132	252	492	254	201
2-methyl-heptane	70	58	134	286	662	250	212
3-methyl-heptane	70	57	134	286	662	248	197
4-methyl-heptane	70	57	134	286	662	248	197
3-ethyl-hexane	70	56	134	286	662	246	182
2,2-dimethyl-hexane	66	52	132	336	1032	236	234
2,3-dimethyl-hexane	68	53	136	320	832	242	187
2,4-dimethyl-hexane	68	54	136	320	832	244	212
2,5-dimethyl-hexane	68	55	136	320	832	246	227
3,3-dimethyl-hexane	66	50	132	336	1032	232	174
3,4-dimethyl-hexane	68	52	136	320	832	240	172
2-methyl-3-ethyl-pentane	68	52	136	320	832	240	172
3-methyl-3-ethyl-pentane	66	48	132	336	1032	228	138
2,2,3-trimethyl-pentane	64	46	134	370	1202	226	158
2,2,4-trimethyl-pentane	64	49	134	370	1202	232	233
2,3,3-trimethyl-pentane	64	45	134	370	1202	224	137
2,3,4-trimethyl-pentane	66	49	138	354	1213	236	131
2,2,3,3-tetramethylbutane	60	39	132	420	1572	210	111

**Figure 3.** Correlation patterns between selected topological coindices and the physicochemical properties of octane isomers.

Comparing the correlation coefficients in Table 5 with those in Table 3, as well as with the coefficients for several well-known topological indices reported in earlier studies, we obtain the following benchmark-level observations. The second Zagreb coindex $\overline{M}_2$ shows the strongest linear association with the acentric factor (AF), with $|r| = 0.993148$, corresponding to $R^2 \approx 0.9863$. Moreover, $\overline{M}_2$ also exhibits the highest correlation with the motor octane number (MON), with $|r| = 0.92592$, corresponding to $R^2 \approx 0.8573$. Among the coindices, the S -coindex $\overline{S}$ has the largest correlation with the entropy property, where $|r| = 0.95581$ and $R^2 \approx 0.9136$. Finally, the second Hyper-Zagreb coindex $\overline{HM}_2$ is most strongly correlated with the molar refraction (MR), with $|r| = 0.930363$ and $R^2 \approx 0.8656$.

A notable feature of Table 5 is that $\overline{F}$ provides the strongest coindex-level association with $\log P$, with $|r| = 0.96693$, whereas several other coindices remain weak for this

property. This again suggests that $\log P$ is more selective with respect to descriptor choice and is not uniformly well captured by all degree-based coindices in the octane benchmark.

These findings show that selected coindices provide competitive benchmark-level associations with some physicochemical properties of octane isomers. The interpretation of these results should, however, remain connected to the structure of the dataset. The octane-isomer graphs form a small and closely related family of trees with common order and size, so several coindices are naturally linked to their corresponding indices. Hence, the observed correlation patterns are best viewed as evidence of descriptor behavior within this controlled benchmark setting.

4.3. Correlation Between Improved Topological Coindices and the PCPs of Octane Isomers

In this subsection, we compute several new topological parameters for the octane isomers and analyze their correlations with various physicochemical properties (see Figure 4). These derived descriptors are constructed from previously defined indices and coindices and are included here for comparative benchmark analysis.

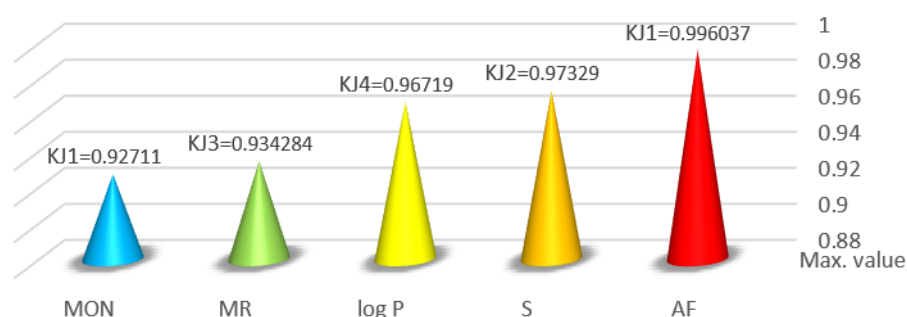


Figure 4. Correlation patterns between selected improved topological descriptors and the physicochemical properties of octane isomers.

The dataset in Table 6 is obtained from the definitions of KJ_1 , KJ_2 , KJ_3 , and KJ_4 . Table 7 reports the corresponding correlation coefficients between the physicochemical properties listed in Table 2 and the improved topological coindices given in Table 6. Comparing Table 7 with Tables 3 and 5, we observe the following strongest benchmark-level associations:

$$|r(KJ_1, AF)| = 0.996037, \quad |r(KJ_1, MON)| = 0.92711, \quad |r(KJ_2, S)| = 0.97329, \\ |r(KJ_3, MR)| = 0.934284, \quad |r(KJ_4, \log P)| = 0.96719.$$

The corresponding coefficients of determination are

$$R^2(KJ_1, AF) \approx 0.9921, \quad R^2(KJ_1, MON) \approx 0.8595, \quad R^2(KJ_2, S) \approx 0.9473, \\ R^2(KJ_3, MR) \approx 0.8729, \quad R^2(KJ_4, \log P) \approx 0.9355.$$

Thus, within the present benchmark dataset, KJ_1 , KJ_2 , KJ_3 , and KJ_4 show the strongest observed associations with the AF, MON, entropy, MR, and $\log P$, respectively. In particular, the gain over the best previously considered descriptors is modest for the AF, MON, MR, and $\log P$, but more noticeable for entropy. This suggests that the derived descriptors may be useful for benchmarking and ranking descriptor behavior, although the improvement is not uniform across all properties. Accordingly, these results should be interpreted as evidence of promising benchmark-level performance rather than as definitive proof that the improved coindices provide broadly applicable predictive models. Their usefulness should be further assessed on larger and chemically more diverse datasets and under stronger validation procedures.

Table 6. Pearson correlation coefficients r between PCPs and topological coindices of octane isomers.

PCPs	AF	S	$\log P$	MR	MON
$\overline{M}_1$	0.973065	0.954294	0.079328	0.323766	−0.88127
$\overline{M}_2$	0.993148	0.955034	0.119290	0.586143	−0.92592
$\overline{F}$	0.116615	0.187354	−0.96693	−0.17771	0.076896
$\overline{Y}$	−0.97158	−0.94308	−0.20972	−0.34656	0.908856
$\overline{S}$	−0.96468	−0.95581	−0.14107	−0.34813	0.876732
$\overline{HM}_1$	0.978812	0.953457	−0.04165	0.548175	−0.86940
$\overline{HM}_2$	0.648131	0.624589	0.106906	0.930363	−0.46794

Table 7. Improved topological coindices of octane isomers used in the benchmark comparison.

Molecule	ND_1	GH	KJ_1	KJ_2	KJ_3	KJ_4
n-octane	23.827	24.423	84.827	1088.483	40,401	17,424
2-methyl-heptane	25.786	27.173	83.786	1400.372	44,944	17,956
3-methyl-heptane	26.559	27.954	83.559	1443.426	38,809	17,956
4-methyl-heptane	26.518	27.954	83.518	1443.426	38,809	17,956
3-ethyl-hexane	27.254	28.735	83.254	1487.700	33,124	17,956
2,2-dimethyl-hexane	29.706	33.607	81.706	2161.430	54,756	17,424
2,3-dimethyl-hexane	29.266	31.637	82.266	1832.900	34,969	18,496
2,4-dimethyl-hexane	28.478	30.885	82.478	1785.883	44,944	18,496
2,5-dimethyl-hexane	27.801	32.248	82.801	1871.934	51,529	18,496
3,3-dimethyl-hexane	31.1296	35.213	81.1296	2271.955	30,276	17,424
3,4-dimethyl-hexane	29.940	32.418	81.940	1882.927	29,584	18,496
2-methyl-3-ethyl-pentane	29.902	32.418	81.902	1882.927	29,584	18,496
3-methyl-3-ethyl-pentane	32.526	36.820	80.526	2387.712	19,044	17,424
2,2,3-trimethyl-pentane	33.880	38.834	79.880	2710.080	24,964	17,956
2,2,4-trimethyl-pentane	31.552	36.537	80.552	2536.952	54,289	17,956
2,3,3-trimethyl-pentane	34.627	39.659	79.627	2774.836	18,769	17,956
2,3,4-trimethyl-pentane	31.907	35.321	80.907	2460.573	17,161	19,044
2,2,3,3-tetramethylbutane	38.749	46.000	77.749	3688.000	12,321	17,424

5. Improved Topological Indices of Octane Isomers

Simple linear regression is a standard tool for examining the relationship between a molecular descriptor and a physicochemical property. In this section, we present exploratory simple linear regression models for the acentric factor (AF), motor octane number (MON), entropy, molar refraction (MR), and the octanol–water partition coefficient ($\log P$). These models are intended as benchmark regressions on the octane-isomer dataset and should not be interpreted as externally validated predictive models.

The regression analyses were carried out in SPSS Version 23 using one descriptor at a time. For the AF, entropy, and $\log P$, the sample size is $n = 18$; for MR, it is $n = 17$; for the MON, it is $n = 16$. To complement the Pearson correlation coefficients reported earlier, Table 8 summarizes the Pearson correlation coefficients for the improved KJ -type

descriptors. The corresponding compact regression summary, including the sample size, absolute correlation coefficient, and coefficient of determination R^2 , is given in Table 9.

Table 8. Pearson correlation coefficients r between PCPs and improved topological coindices of octane isomers.

PCPs	AF	S	log P	MR	MON
KJ_1	0.996037	0.955252	0.110732	0.537683	−0.92711
KJ_2	−0.98053	−0.97329	−0.09063	−0.43452	0.895268
KJ_3	0.593091	0.554915	0.111362	0.934284	−0.43119
KJ_4	0.115016	0.184709	−0.96719	−0.17882	0.079588

Table 9. Compact summary of simple linear regression models for the best-performing coindex and derived KJ -type descriptor in each property.

Property	Descriptor	Sample Size	$ r $	R^2
AF	$\overline{M}_2$	18	0.993148	0.9863
AF	KJ_1	18	0.996037	0.9921
Entropy	$\overline{S}$	18	0.95581	0.9136
Entropy	KJ_2	18	0.97329	0.9473
log P	$\overline{F}$	18	0.96693	0.9350
log P	KJ_4	18	0.96719	0.9355
MR	$\overline{HM}_2$	17	0.930363	0.8656
MR	KJ_3	17	0.934284	0.8729
MON	$\overline{M}_2$	16	0.92592	0.8573
MON	KJ_1	16	0.92711	0.8595

Regression equation for the acentric factor (AF) using the best topological descriptors:

Based on the correlation coefficients in Table 5, the coindex data in Table 4, and the SPSS analysis, the best simple linear regression model involving topological coindices for the AF is (see Figure 5)

$$AF = 0.0066 \overline{M}_2 - 0.0047.$$

For this model, $|r| = 0.993148$ and $R^2 \approx 0.9863$, indicating that the linear fit explains about 98.63% of the observed variance in the AF within this dataset.

Similarly, from the correlation coefficients in Table 7 and the improved-coindex data in Table 6, the best simple linear regression model involving the improved topological coindices for the AF is (see Figure 5)

$$AF = 0.0201 KJ_1 - 1.312.$$

Here, $|r| = 0.996037$ and $R^2 \approx 0.9921$, which represents a modest improvement over the corresponding coindex-based model.

Thus, both $\overline{M}_2$ and KJ_1 show very strong benchmark-level linear association with the AF, with KJ_1 performing slightly better on this dataset. A comparative view of the performance of these descriptors with respect to the AF is shown in Figure 5.

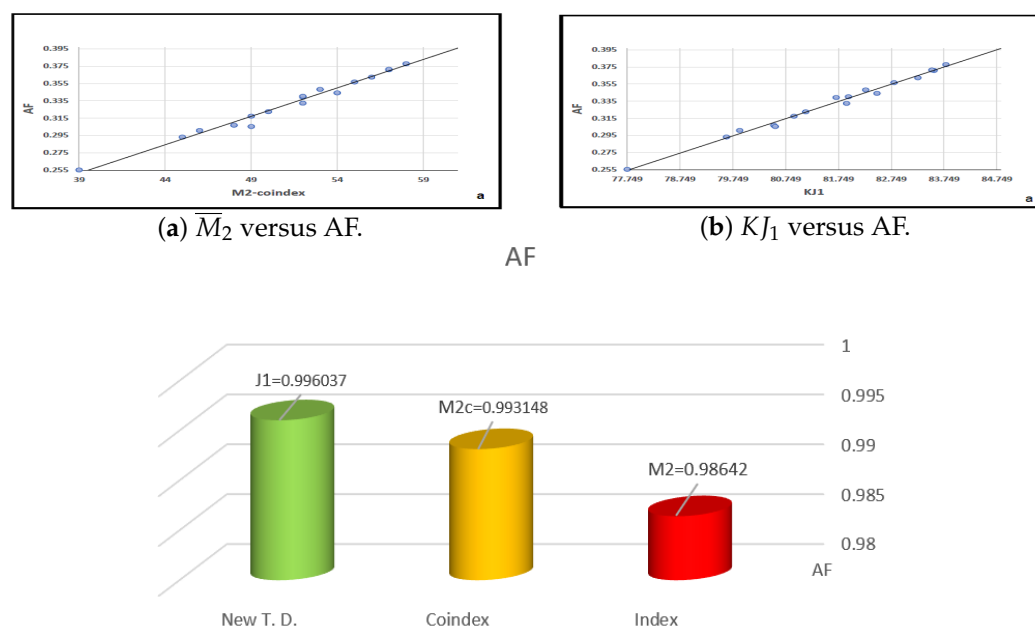


Figure 5. Simple linear regression plots and comparative descriptor performance for the acentric factor (AF) within the octane-isomer benchmark.

Regression equation for entropy using the best topological descriptors:

From the correlation coefficients in Table 5 and the coindex data in Table 4, the best simple linear regression model based on topological coindices for entropy is obtained using SPSS (see Figure 6):

$$S = -0.0158 \bar{S} + 120.13.$$

For this model, $|r| = 0.95581$ and $R^2 \approx 0.9136$.

Furthermore, using the correlation coefficients in Table 7 and the improved-coindex data in Table 6, the best simple linear regression model involving the improved topological coindices for entropy is (see Figure 6)

$$S = -0.0069 KJ_2 + 119.72.$$

For this model, $|r| = 0.97329$ and $R^2 \approx 0.9473$.

Among the properties considered in this study, entropy shows one of the clearest improvements when passing from the best coindex-based model to the best improved-coindex model. Nevertheless, this conclusion remains benchmark-specific and should be interpreted cautiously in the absence of external validation. A comparison of these best-performing descriptors for entropy is displayed in Figure 6.

Regression equation for the octanol–water partition coefficient ($\log P$) using the best topological descriptors:

From the correlation coefficients in Table 5 and the coindex data in Table 4, the best simple linear regression model involving topological coindices to estimate $\log P$ is obtained using SPSS as (see Figure 7)

$$\log P = -0.0275 \bar{F} + 7.28.$$

For this model, $|r| = 0.96693$ and $R^2 \approx 0.9350$.

In addition, from the correlation coefficients in Table 7 and the improved-coindex data in Table 6, the best simple linear regression model involving the improved topological coindices for $\log P$ is (see Figure 7)

$$\log P = -0.0001 KJ_4 + 5.4318.$$

For this model, $|r| = 0.96719$ and $R^2 \approx 0.9355$.

The two models have almost identical R^2 values, so the improvement provided by KJ_4 over $\bar{F}$ is negligible in practical terms for this benchmark. This also helps explain why $\log P$ should not be described as substantially better predicted by the improved coindex in comparison with the best classical coindex. A comparative illustration of the most suitable descriptors for $\log P$ is provided in Figure 7.

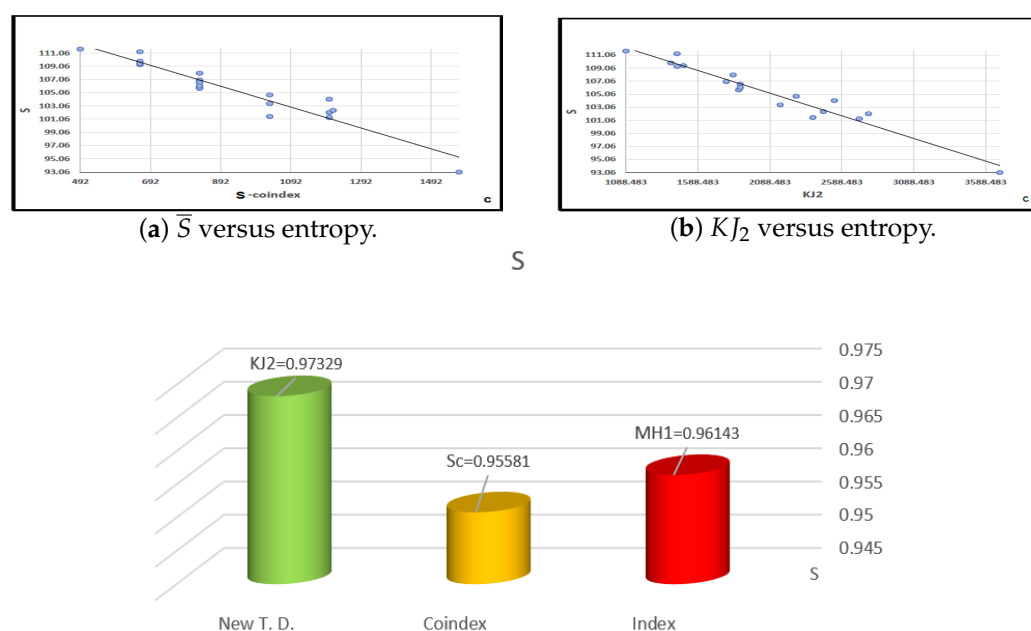


Figure 6. Simple linear regression plots and comparative descriptor performance for entropy within the octane-isomer benchmark.

Regression equation for molar refraction (MR) using the best topological descriptors:

Using the correlation coefficients in Table 5, together with the coindex data in Table 4 and the SPSS software, the best simple linear regression model based on topological coindices for estimating MR is (see Figure 8)

$$MR = 0.0052 \overline{HM}_2 + 38.057.$$

For this model, $|r| = 0.930363$ and $R^2 \approx 0.8656$.

Moreover, from the correlation coefficients in Table 7 and the improved-coindex data in Table 6, the best simple linear regression model involving the improved topological coindices for MR is (see Figure 8)

$$MR = 0.00001 KJ_3 + 38.519.$$

For this model, $|r| = 0.934284$ and $R^2 \approx 0.8729$.

The improvement from $\overline{HM}_2$ to KJ_3 is present but modest, and both descriptors yield fairly similar benchmark-level performance for MR. The comparison of the best descriptors for MR is summarized in Figure 8.

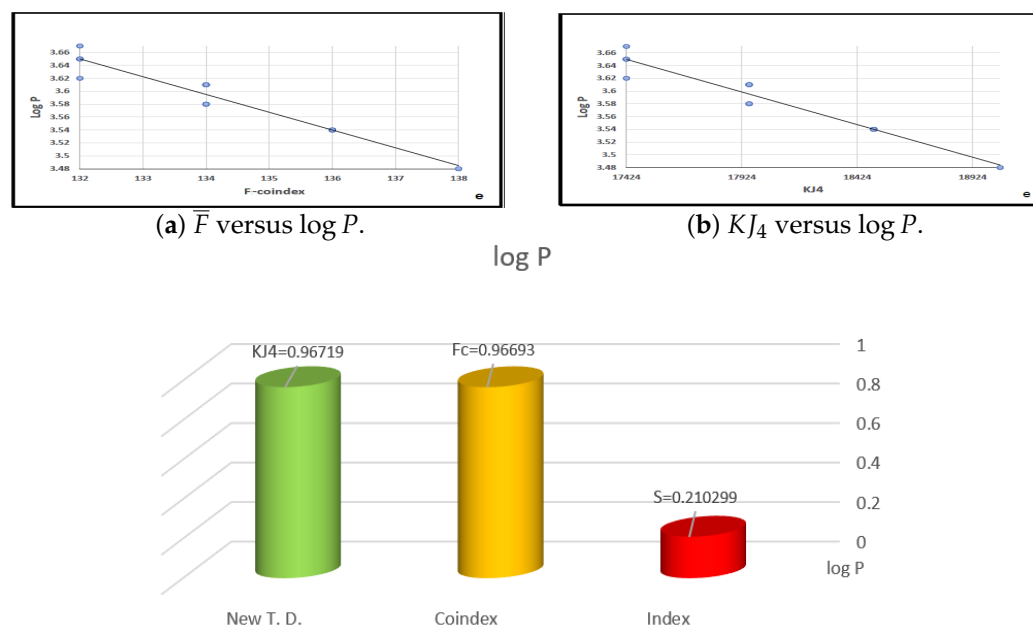


Figure 7. Simple linear regression plots and comparative descriptor performance for $\log P$ within the octane-isomer benchmark.

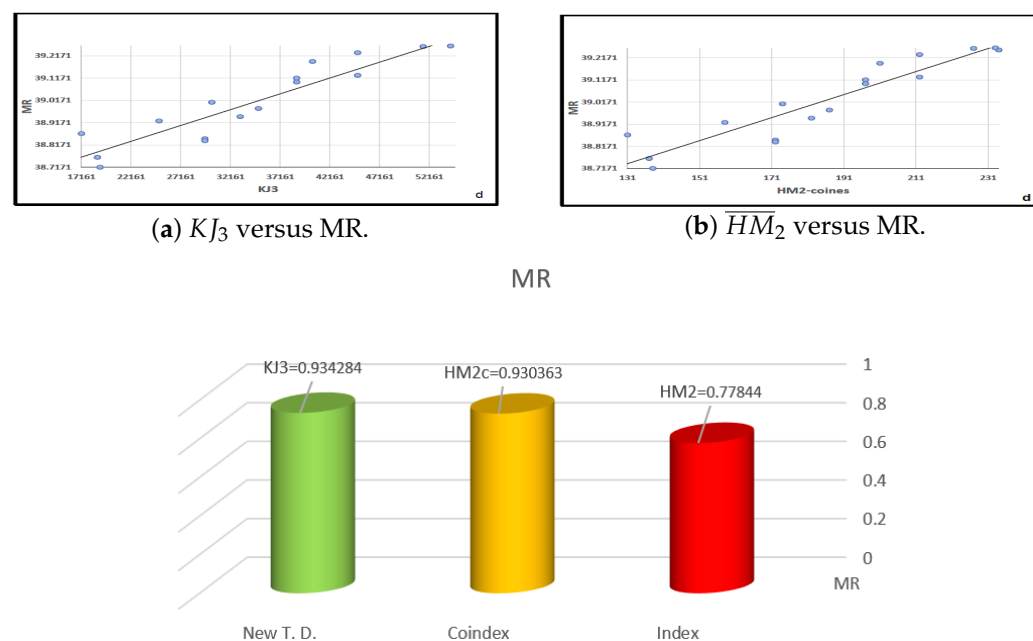


Figure 8. Simple linear regression plots and comparative descriptor performance for molar refraction (MR) within the octane-isomer benchmark.

Regression equation for the motor octane number (MON) using the best topological descriptors:

From the correlation coefficients in Table 5, the coindex data in Table 4, and the SPSS analysis, the best simple linear regression model involving topological coindices for MON is (see Figure 9)

$$MON = -5.7485 \overline{M}_2 + 372.31.$$

For this model, $|r| = 0.92592$ and $R^2 \approx 0.8573$.

Similarly, from the correlation coefficients in Table 7 and the improved-coindex data in Table 6, the best simple linear regression model involving the improved topological coindices for the MON is (see Figure 9)

$$MON = -17.254 KJ_1 + 1485.5.$$

For this model, $|r| = 0.92711$ and $R^2 \approx 0.8595$.

As with the AF, the improved descriptor KJ_1 performs slightly better than the best coindex model, but the gain is small. Accordingly, the MON analysis supports comparative usefulness within the benchmark dataset rather than a strong claim of general predictive superiority. A comparison of the best-performing descriptors for the MON is presented in Figure 9.

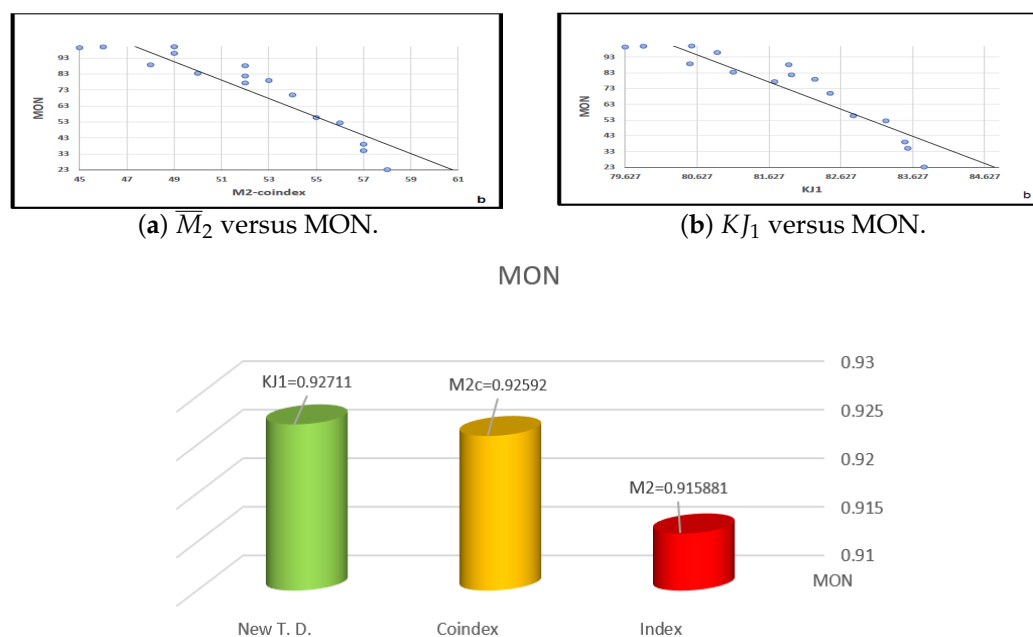


Figure 9. Simple linear regression plots and comparative descriptor performance for the motor octane number (MON) within the octane-isomer benchmark.

Overall, the regression analysis confirms the same general pattern seen in the correlation study: some coindices and improved coindices provide strong linear benchmark fits for selected properties, but the improvements are often modest and remain limited to the present octane-isomer dataset. In particular, entropy shows the clearest improvement from a coindex-based model to an improved-coindex model, whereas the gain for the log P , MR, and MON is comparatively small.

6. QSPR Modeling on an Expanded Chemo Dataset

The octane-isomer analysis in the previous section is useful as a controlled benchmark, but it is necessarily limited by the small number of isomers and by the structural homogeneity of alkane trees. To provide a broader exploratory validation, we added an expanded QSPR experiment using a molecular dataset obtained from Zenodo (DOI: <https://doi.org/10.5281/zenodo.18694026>) [46]. The dataset contains 90 organic compounds represented by SMILES strings, InChI identifiers, molecular formulae, and nine physicochemical endpoints. The numerical values were used as reported in the source dataset.

6.1. Dataset and Molecular Graph Construction

Each SMILES string was converted into a hydrogen-suppressed molecular graph using RDKit [47]. In this graph representation, vertices correspond to heavy atoms and edges correspond to covalent bonds. If a disconnected representation was encountered, the largest heavy-atom fragment was retained before descriptor calculation. All 90 SMILES strings were successfully parsed, and no compound was removed during graph construction. The endpoints considered in the present QSPR experiment are summarized in Table 10. The target variables include molecular volume, critical constants, thermodynamic properties, water solubility, and phase-transition temperatures.

Table 10. Summary of the expanded molecular dataset used for QSPR modeling. Units are those reported in the source dataset.

Property	Description	N	Min.	Max.	Mean	SD
McVol	Molecular volume	90	99.04	269.03	171.07	35.31
Vc	Critical volume	90	0.310	1.018	0.642	0.140
ΔfG°	Standard Gibbs free energy of formation	90	−2460.30	75.24	−417.09	452.68
$\Delta_{fus}H^\circ$	Enthalpy of fusion	90	16.69	44.82	26.80	6.340
$\log_{10} WS$	Logarithmic water solubility	90	−5.980	−1.170	−3.149	1.107
Pc	Critical pressure	90	1504.65	6349.11	2995.43	759.92
Tboil	Boiling temperature	90	521.20	1083.39	707.52	109.86
Tc	Critical temperature	90	750.72	1383.53	929.06	124.35
Tfus	Fusion/melting temperature	90	314.75	739.24	448.80	86.14

6.2. Computed Graph-Theoretical Descriptors

For each molecular graph, we computed the classical degree-based descriptors, the corresponding coindices, and the four proposed composite descriptors introduced earlier in this paper. The calculated descriptor set included graph-size variables, Zagreb-type indices, higher-order degree-sum indices, square-root degree indices, coindices over non-adjacent vertex pairs, and the proposed *KJ*-type descriptors. The descriptor groups are listed in Table 11. In total, 30 descriptor columns were generated; one zero-variance descriptor was automatically excluded during model fitting, leaving 29 descriptors for the regression models.

Table 11. Graph-theoretical descriptor groups used in the expanded QSPR experiment.

Descriptor Group	Descriptors Included
Graph-size and degree variables	$n, m, \delta_{\min}, \delta_{\max}$, average degree
Classical degree-based indices	$M_1, M_2, F, Y, S, HM_1, HM_2, ND_1, GH$, Nirmala, Randić, sum-connectivity, Sombor
Topological coindices	$\overline{M}_1, \overline{M}_2, \overline{F}, \overline{Y}, \overline{S}, \overline{HM}_1, \overline{HM}_2, \overline{N}$
Proposed composite descriptors	$KJ_1 = \overline{M}_2 + ND_1, KJ_2 = \overline{S} + GH^2, KJ_3 = (\overline{HM}_2)^2, KJ_4 = (\overline{F})^2$

6.3. Modeling Protocol and Validation Metrics

Three transparent regression methods were used: ordinary least squares (OLS), ridge regression, and partial least squares with two latent components, denoted by PLS(2). All models were implemented in a standardized pipeline in which descriptors were scaled using the standard scaler before model fitting. Ridge regression used internal cross-validation over 80 logarithmically spaced regularization parameters from 10^{-4} to 10^4 . Model fitting and validation were performed with scikit-learn [48]. For each target, the dataset was split into training and external test subsets with an 80:20 ratio, giving 72 training compounds and 18 test compounds. In addition, five-fold cross-validation with shuffling was applied to the

full available dataset to obtain cross-validated predictions. This validation strategy follows the general QSPR practice of combining goodness-of-fit measures with predictive error measures and cross-validation, as recommended in descriptor-based QSPR studies [49].

The following metrics were reported:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (14)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (15)$$

The cross-validated coefficient $Q_{5\text{-fold}}^2$ was computed using the same R^2 formula, but with the predictions obtained from five-fold cross-validation. Reporting both R_{test}^2 and $Q_{5\text{-fold}}^2$ is important because the external split can be sensitive to the particular division of a moderate-sized dataset, whereas cross-validation gives a complementary estimate of stability.

6.4. QSPR Performance on the Expanded Dataset

The best-performing model for each endpoint, selected according to the external test-set R^2 , is shown in Table 12. The strongest results were obtained for the critical volume V_c and molecular volume McVol. For V_c , ridge regression achieved $R_{test}^2 = 0.884$ and $Q_{5\text{-fold}}^2 = 0.953$, with a test RMSE of 0.035. For McVol, ridge regression achieved $R_{test}^2 = 0.845$ and $Q_{5\text{-fold}}^2 = 0.869$. Good test-set performance was also obtained for ΔfG° and $\log_{10} WS$, with $R_{test}^2 = 0.838$ and 0.717, respectively. The critical pressure P_c showed moderate performance. In contrast, the phase-transition temperatures T_{boil} , T_c , and T_{fus} exhibited weaker external test-set behavior, despite some high cross-validated scores for ridge regression. Therefore, these endpoints should be interpreted cautiously and should not be presented as strongly validated predictive targets.

Table 12. The best QSPR model for each physicochemical endpoint on the expanded dataset. The best model was selected by the external test-set R^2 .

Target	Model	N	Train/Test	Features	R_{train}^2	R_{test}^2	$Q_{5\text{-fold}}^2$	$RMSE_{test}$	MAE_{test}	$RMSE_{5\text{-fold}}$
McVol	Ridge	90	72/18	29	0.980	0.845	0.869	10.01	6.833	12.70
Pc	OLS	90	72/18	29	0.862	0.649	0.572	434.36	341.16	494.19
Tboil	OLS	90	72/18	29	0.950	0.113	0.213	58.37	39.25	96.89
Tc	OLS	90	72/18	29	0.964	0.155	0.297	64.24	41.40	103.71
Tfus	Ridge	90	72/18	29	0.858	−0.009	0.738	55.22	42.54	43.82
Vc	Ridge	90	72/18	29	0.970	0.884	0.953	0.035	0.025	0.030
$\log_{10} WS$	Ridge	90	72/18	29	0.806	0.717	0.762	0.568	0.441	0.537
ΔfG°	Ridge	90	72/18	29	0.967	0.838	0.829	159.03	110.33	186.37
$\Delta_{fus}H^\circ$	OLS	90	72/18	29	0.954	0.262	0.796	3.546	2.233	2.847

Figure 10 compares the external test-set R^2 values with the five-fold cross-validated Q^2 values for the selected best models. This comparison shows that some endpoints, such as V_c , McVol, ΔfG° , and $\log_{10} WS$, have reasonably consistent validation behavior. However, for T_{boil} , T_c , and T_{fus} , the gap between R_{test}^2 and $Q_{5\text{-fold}}^2$ indicates sensitivity to the external split and possible limitations of the descriptor-only linear models for these properties.

The cross-validated observed-versus-predicted plots for the best models are presented in Figure 11. The plots visually confirm the strong associations for V_c , McVol, and ΔfG° , while showing larger scatter for the less stable temperature-related endpoints. These results

support the use of the proposed descriptors as exploratory QSPR variables, but they also show that descriptor performance depends strongly on the physicochemical endpoint.

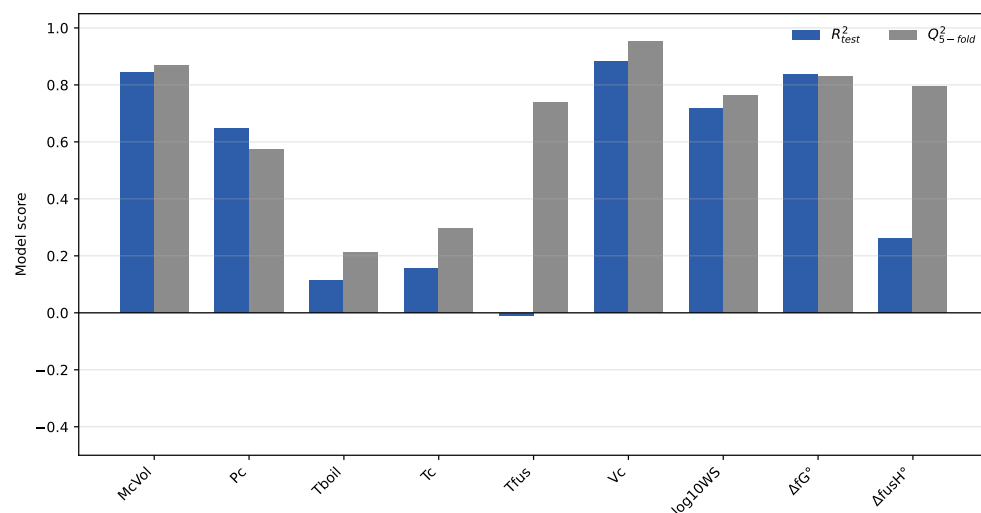


Figure 10. External test-set and five-fold cross-validated performance of the best model selected for each physicochemical endpoint. The discrepancy between R^2_{test} and Q^2_{5-fold} highlights the importance of reporting both external-split and cross-validation metrics.

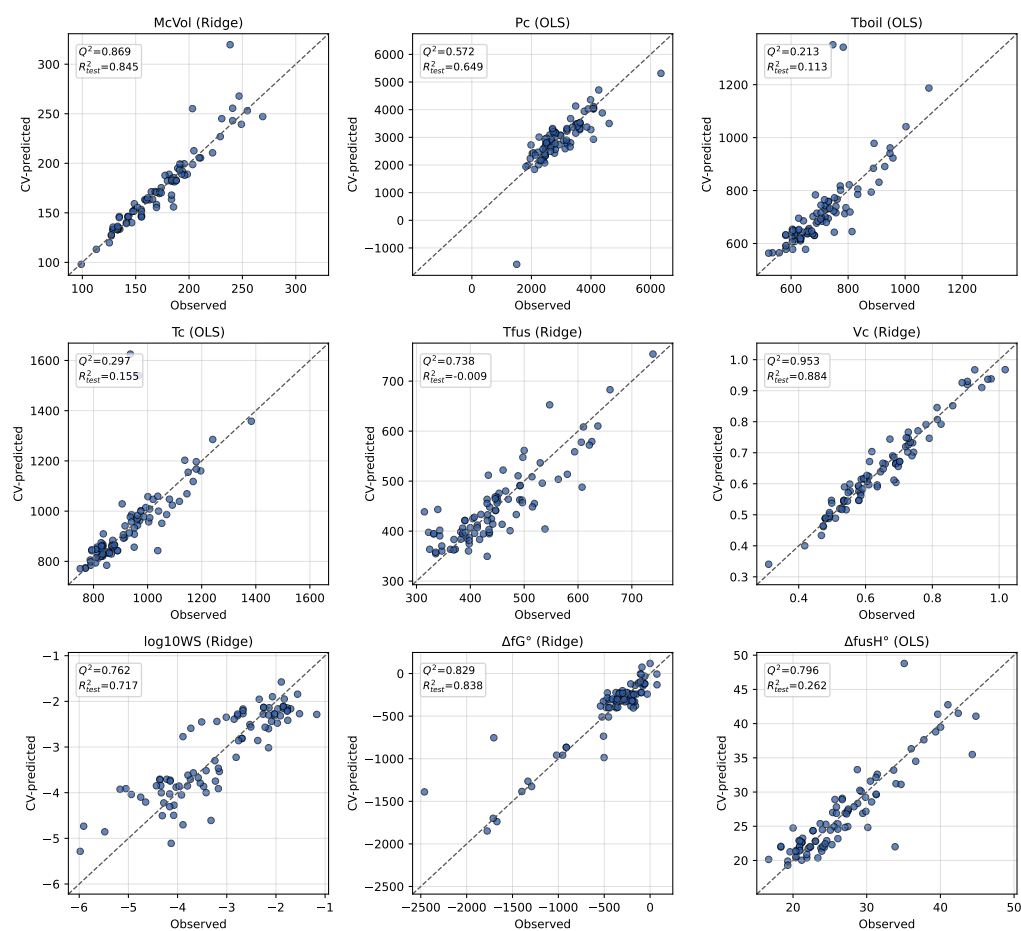


Figure 11. Observed versus five-fold cross-validated predicted values for the best model selected for each endpoint in the expanded Chemo dataset. The dashed line represents the ideal prediction line.

For completeness, Table 13 reports the full comparison among the OLS, ridge, and PLS(2) models. Ridge regression was frequently competitive, especially for end-

points affected by descriptor collinearity. OLS produced high training scores for several targets, but the cross-validated and external test metrics show that this did not always translate into stable predictive performance. This observation is consistent with the known redundancy of degree-based molecular descriptors.

Table 13. Full model comparison for the expanded QSPR experiment.

Target	Model	R^2_{train}	R^2_{test}	Q^2_{5-fold}	$RMSE_{test}$	MAE_{test}	$RMSE_{5-fold}$	MAE_{5-fold}
McVol	Ridge	0.980	0.845	0.869	10.01	6.833	12.70	6.889
McVol	OLS	0.984	0.829	0.684	10.51	7.075	19.73	8.332
McVol	PLS(2)	0.920	0.790	0.894	11.67	8.007	11.42	8.242
Pc	OLS	0.862	0.649	0.572	434.36	341.16	494.19	317.22
Pc	Ridge	0.784	0.643	0.725	437.91	335.65	396.46	297.07
Pc	PLS(2)	0.586	0.354	0.513	589.16	493.47	527.39	414.24
Tboil	OLS	0.950	0.113	0.213	58.37	39.25	96.89	44.94
Tboil	Ridge	0.929	−0.004	0.822	62.09	42.99	46.10	34.04
Tboil	PLS(2)	0.868	−0.418	0.763	73.79	57.00	53.18	41.72
Tc	OLS	0.964	0.155	0.297	64.24	41.40	103.71	44.09
Tc	Ridge	0.944	0.083	0.843	66.93	43.80	49.02	34.52
Tc	PLS(2)	0.891	−0.308	0.796	79.93	59.93	55.82	42.53
Tfus	Ridge	0.858	−0.009	0.738	55.22	42.54	43.82	32.57
Tfus	OLS	0.905	−0.067	−0.236	56.80	43.13	95.23	44.35
Tfus	PLS(2)	0.765	−0.257	0.657	61.64	50.39	50.19	39.22
Vc	Ridge	0.970	0.884	0.953	0.035	0.025	0.030	0.024
Vc	OLS	0.974	0.870	0.799	0.037	0.027	0.062	0.032
Vc	PLS(2)	0.895	0.853	0.875	0.039	0.028	0.049	0.036
$\log_{10} WS$	Ridge	0.806	0.717	0.762	0.568	0.441	0.537	0.431
$\log_{10} WS$	OLS	0.857	0.676	0.071	0.608	0.459	1.061	0.604
$\log_{10} WS$	PLS(2)	0.770	0.674	0.743	0.610	0.504	0.558	0.459
ΔfG°	Ridge	0.967	0.838	0.829	159.03	110.33	186.37	103.18
ΔfG°	PLS(2)	0.956	0.819	0.927	168.51	116.52	121.83	90.39
ΔfG°	OLS	0.978	0.811	0.491	171.99	116.26	321.16	123.94
$\Delta fusH^\circ$	OLS	0.954	0.262	0.796	3.546	2.233	2.847	1.868
$\Delta fusH^\circ$	Ridge	0.909	0.209	0.692	3.672	2.442	3.497	2.088
$\Delta fusH^\circ$	PLS(2)	0.868	−0.018	0.786	4.165	3.024	2.917	2.236

6.5. Descriptor Redundancy and Overfitting Control

Because the descriptors considered here are all degree-based or coindex-based, substantial collinearity is expected. To assess descriptor redundancy, we computed the absolute Pearson correlation matrix among the calculated descriptors. Figure 12 shows the lower triangular redundancy matrix, and Table 14 lists the strongest descriptor-pair correlations. Several pairs are almost perfectly correlated; for example, $\overline{M}_2$ and KJ_1 have an absolute correlation of 0.99998, reflecting the fact that $KJ_1 = \overline{M}_2 + ND_1$ is dominated by the coindex component in this dataset. Similarly, strong correlations occur between $\overline{M}_1$, $\overline{HM}_1$, and $\overline{N}$ and between square-root degree descriptors, such as ND_1 and the Nirmala index.

The redundancy analysis explains why regularized or latent-variable approaches, such as ridge regression and PLS(2), are useful in this setting. It also prevents overinterpretation of individual descriptor effects: when several descriptors carry nearly the same information, high model performance should be attributed to the descriptor family as a whole rather than to a single isolated index. Accordingly, the expanded QSPR results are presented as exploratory validation of degree-based and coindex-based descriptor usefulness, not as evidence that the proposed KJ -type descriptors are universally superior.

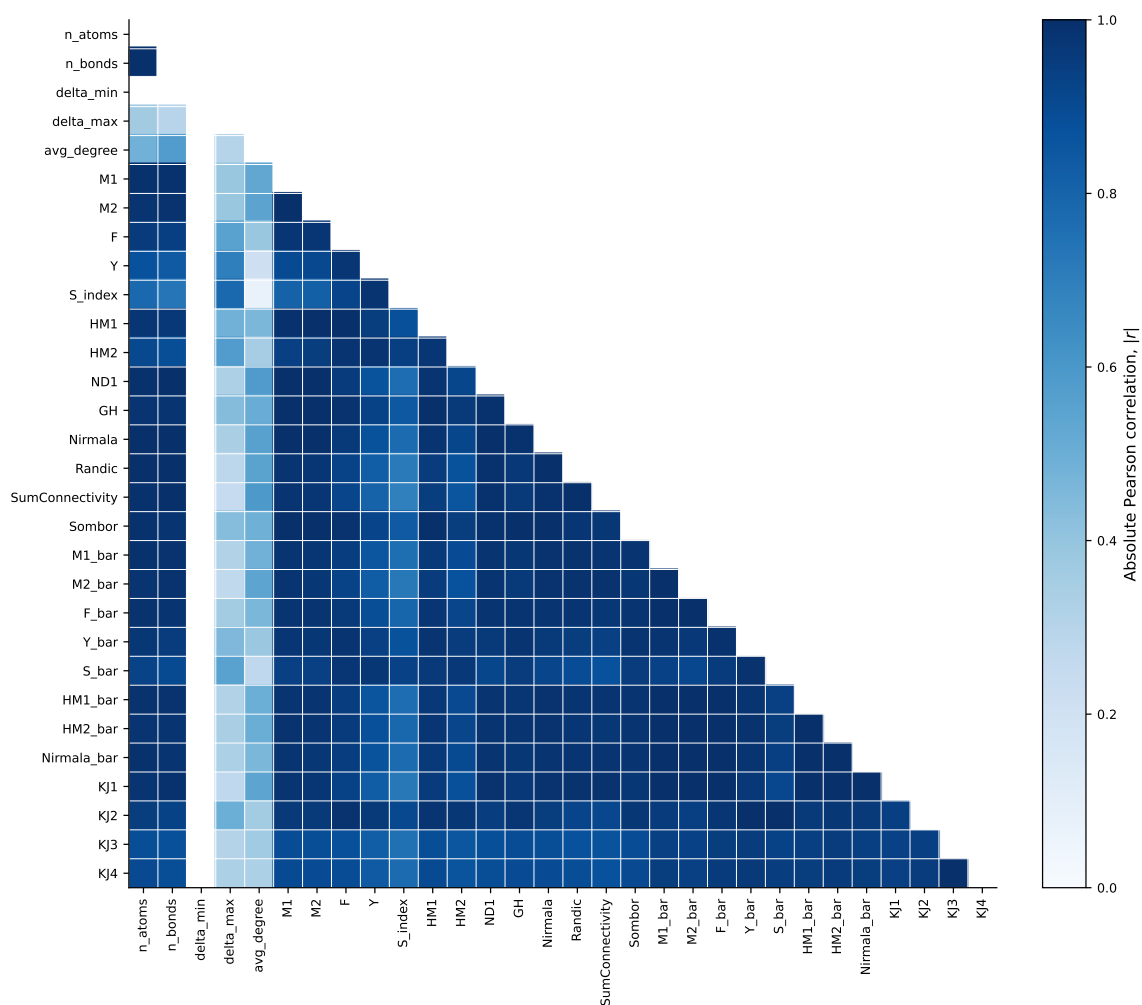


Figure 12. Descriptor-redundancy analysis based on absolute Pearson correlation coefficients among the computed graph-theoretical descriptors. Only the lower triangular part is shown to avoid duplicated information; darker cells indicate stronger linear dependence between descriptor pairs.

Table 14. Highest descriptor-pair absolute Pearson correlations in the expanded dataset.

Descriptor 1	Descriptor 2	$ r $
M2_bar	KJ1	1.0000
M1_bar	Nirmala_bar	0.9996
M1_bar	HM1_bar	0.9994
ND1	Nirmala	0.9994
Randic	SumConnectivity	0.9984
M1	Sombor	0.9984
M2	GH	0.9984
HM1_bar	Nirmala_bar	0.9984
F_bar	HM1_bar	0.9983
F_bar	Nirmala_bar	0.9982
HM1	GH	0.9980
n_bonds	Nirmala	0.9980
n_bonds	SumConnectivity	0.9979
M1	Nirmala	0.9979
F_bar	HM2_bar	0.9978

6.6. Interpretation of the Expanded QSPR Analysis

The expanded Chemo experiment addresses the main statistical limitation of the octane-isomer benchmark by increasing the number of compounds from 18 to 90 and by including several different physicochemical endpoints. The results indicate that graph-theoretical descriptors can provide useful information for selected size-related and thermodynamic properties, especially V_c , McVol, ΔfG° , and $\log_{10} WS$. At the same time, the weaker behavior observed for some temperature endpoints shows that degree-based descriptors alone do not capture all factors controlling these properties. Consequently, the proposed descriptors should be viewed as compact, interpretable structural descriptors that can support QSPR modeling, particularly when used with validation and redundancy analysis.

This expanded analysis strengthens the application part of this manuscript in two ways. First, it supplements the octane-isomer benchmark with a larger and chemically more diverse dataset. Second, it reports external test-set and cross-validation metrics, including the RMSE and MAE, thereby reducing the risk of relying only on simple correlations or unvalidated regression fits. Nevertheless, the conclusions remain appropriately cautious: the models are descriptor-based exploratory QSPR models, and broader predictive claims would require larger datasets, external validation on independent chemical families, and comparison with additional physicochemical and electronic descriptors.

7. Conclusions

In this work, we studied Zagreb-type descriptors from theoretical and applied perspectives. On the theoretical side, we presented a unified collection of degree-based bounds for the first and second (α, β) -general Zagreb indices and for their reduced and expanded variants. These inequalities express the considered descriptors in terms of basic graph parameters, such as the order, size, minimum degree, and maximum degree. They therefore provide a common framework for comparing several classical and generalized indices and for identifying how their extremal behavior is controlled by the degree range of the graph.

On the octane-isomer benchmark, the computed correlations and simple linear regression models showed that several coindices and derived descriptors have strong benchmark-level associations with selected physicochemical properties. In particular, $\overline{M}_2$ and KJ_1 were effective for the AF and MON, $\overline{S}$ and KJ_2 were effective for entropy, $\overline{HM}_2$ and KJ_3 were effective for MR, and $\overline{F}$ and KJ_4 were effective for $\log P$. The largest relative improvement from a coindex-based model to a derived-descriptor model was observed for entropy, whereas the gains for the AF, MON, MR, and especially $\log P$ were modest. Hence, the octane results should be interpreted as controlled benchmark evidence rather than as broad predictive validation.

Furthermore, we added an expanded QSPR analysis based on 90 organic compounds obtained from Zenodo and nine physicochemical endpoints. In this analysis, molecular graphs were constructed from SMILES strings, 30 graph-theoretical descriptors were generated, and OLS, ridge regression, and PLS(2) models were evaluated using an 80:20 train/test split and five-fold cross-validation. The strongest and most stable results were observed for the critical volume, molecular volume, standard Gibbs free energy of formation, and logarithmic water solubility. For example, ridge regression gave strong external and cross-validated performance for V_c and McVol, while the results for some temperature-related endpoints were less stable and require cautious interpretation.

The descriptor-redundancy analysis showed that several degree-based and coindex-based descriptors are highly correlated. This is expected because many of them are built from the same vertex-degree information. Consequently, individual descriptor effects should not be overinterpreted, and high model performance should be attributed to the

descriptor family as a whole rather than to a single isolated index. Overall, the results support the usefulness of topological coindices and KJ -type descriptors as compact and interpretable exploratory QSPR variables. Future work may focus on sharper bounds for specific graph classes, larger external molecular datasets, additional physicochemical and electronic descriptors, and more rigorous validation strategies.

Author Contributions: Conceptualization: S.W. and A.A.; Methodology: A.A.; Validation: S.W.; Investigation: S.W. and A.A.; Resources: A.A.; Writing—Original Draft: A.A.; Writing—Review and Editing: S.W. All authors have read and agreed to the published version of this manuscript.

Funding: This project was funded by KAU Endowment (WAQF) at King Abdulaziz University, Jeddah, Saudi Arabia. The authors therefore acknowledge, with thanks, WAQF and the Deanship of Scientific Research (DSR) for their technical and financial support.

Data Availability Statement: The octane-isomer data used in the benchmark part are cited at the relevant places within the text. The expanded QSPR dataset used in Section 6 was obtained from Zenodo, DOI: <https://doi.org/10.5281/zenodo.18694026>, and the numerical values were used as reported by the source dataset. Scripts and codes used for descriptor computation and model benchmarking are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Ediz, S.; Çiftçi, İ.; Taş, Z.; Cancan, M.; Farahani, M.R.; Aldemir, M.Ş. A note on QSPR analysis of total Zagreb and total Randić indices of octanes. *Eurasian Chem. Commun.* **2021**, *3*, 139–145.
2. Shanmukha, M.C.; Basavarajappa, N.S.; Anilkumar, K.N. Predicting physico-chemical properties of octane isomers using QSPR approach. *Malaya J. Math.* **2020**, *8*, 104–116. [\[CrossRef\]](#)
3. Ediz, S. Predicting some physicochemical properties of octane isomers: A topological approach using ev -degree and ve -degree Zagreb indices. *Int. J. Syst. Sci. Appl. Math.* **2017**, *2*, 87–92.
4. Kuanar, M.; Kuanar, S.K.; Mishra, B.K.; Gutman, I. Correlation of line graph parameters with physicochemical properties of octane isomers. *Indian J. Chem. A* **1999**, *38A*, 525–528.
5. Alsharafi, M.; Zeren, Y. Prodeg type Shannon graph entropies with closed forms bounds and QSPR modeling. *Sci. Rep.* **2026**. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Das, K.C.; Gutman, I.; Furtula, B. Survey on geometric-arithmetic indices of graphs. *MATCH Commun. Math. Comput. Chem.* **2011**, *65*, 595–644.
7. Alameri, A.; Alsharafi, M. *Topological Indices Types in Graphs and Their Applications*; Generis Publishing: Chisinau, Moldova, 2021.
8. Alsharafi, M.; Alameri, A. The F -index and coindex of V-phenylenic nanotubes and nanotorus and their molecular complement graphs. *Nanosyst. Phys. Chem. Math.* **2021**, *12*, 263–270. [\[CrossRef\]](#)
9. Öztürk Sözen, E.; Eryaşar, E. An algebraic approach to calculate some topological coindices and QSPR analysis of some novel drugs used in the treatment of breast cancer. *Polycycl. Aromat. Compd.* **2024**, *44*, 2226–2243.
10. Öztürk Sözen, E.; Eryaşar, E. QSPR analysis of some drug candidates investigated for COVID-19 via new topological coindices. *Polycycl. Aromat. Compd.* **2024**, *44*, 1291–1308.
11. Alameri, A.Q.S.; Al-Sharafi, M.S.Y. Topological indices of some new graph operations and their possible applications. *Asian J. Probab. Stat.* **2021**, *13*, 16–30. [\[CrossRef\]](#)
12. Gutman, I.; Trinajstić, N. Graph theory and molecular orbitals. Total π -electron energy of alternant hydrocarbons. *Chem. Phys. Lett.* **1972**, *17*, 535–538.
13. Gutman, I.; Rušćić, B.; Trinajstić, B.; Wilcox, C.F., Jr. Graph theory and molecular orbitals. XII. Acyclic polyenes. *J. Chem. Phys.* **1975**, *62*, 3399–3405. [\[CrossRef\]](#)
14. Došlić, T. Vertex-weighted Wiener polynomials for composite graphs. *Ars. Math. Contemp.* **2008**, *1*, 66–80. [\[CrossRef\]](#)
15. Ashrafi, A.; Došlić, T.; Hamzeh, A. The Zagreb coindices of graph operations. *Discret. Appl. Math.* **2010**, *158*, 1571–1578. [\[CrossRef\]](#)
16. Shirdel, G.H.; Rezapour, H.; Sayadi, A.M. The hyper-Zagreb index of graph operations. *Iran. J. Math. Chem.* **2013**, *4*, 213–220. [\[CrossRef\]](#)
17. Gutman, I. On hyper-Zagreb index and coindex. *Bull. Acad. Serbe Sci. Arts Cl. Sci. Math. Natur. Sci. Math.* **2017**, *42*, 1875–1890.
18. Furtula, B.; Gutman, I. A forgotten topological index. *J. Math. Chem.* **2015**, *53*, 1184–1190. [\[CrossRef\]](#)
19. De, N.; Nayeem, S.; Abu, M.; Pal, A. The F -coindex of some graph operations. *SpringerPlus* **2016**, *5*, 221. [\[CrossRef\]](#) [\[PubMed\]](#)

20. Gao, W.; Farahani, M.R.; Siddiqui, M.K. On the first and second Zagreb and first and second hyper-Zagreb indices of carbon nanocones $CNC_k[n]$. *J. Comput. Theor. Nanosci.* **2016**, *13*, 7475–7482. [CrossRef]
21. Mondal, S.; Dey, A.; De, N.; Pal, A. QSPR analysis of some novel neighbourhood degree-based topological descriptors. *Complex Intell. Syst.* **2021**, *7*, 977–996. [CrossRef]
22. Usha, A.; Shanmukha, M.C.; Kumar, K.N.A.; Shilpa, K.C. Comparison of novel index with geometric-arithmetic and sum-connectivity indices. *J. Math. Comput. Sci.* **2021**, *11*, 5344–5360.
23. Vukičević, D.; Furtula, B. Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges. *J. Math. Chem.* **2009**, *46*, 1369–1376. [CrossRef]
24. Alameri, A.; Al-Naggar, N.; Al-Rumaima, M.; Alsharafi, M. Y-index of some graph operations. *Int. J. Appl. Eng. Res.* **2020**, *15*, 173–179.
25. Alameri, A.; Al-Rumaima, M.; Almazah, M. Y-coindex of graph operations and its applications of molecular descriptors. *J. Mol. Struct.* **2020**, *1221*, 128754. [CrossRef]
26. Alsharafi, M.; Zeren, Y.; Alameri, A. The second Hyper-Zagreb index of complement graphs and its applications of some nanostructures. *Asian J. Probab. Stat.* **2021**, *15*, 54–75.
27. Sethumadhavan, N.; Gokulathilagan, K. S-coindex of some graph operations and its properties and conjugated polymers. *Glob. J. Eng. Technol. Adv.* **2023**, *16*, 140–149. [CrossRef]
28. Kulli, V.R. Nirmala index. *Int. J. Math. Trends Technol.* **2021**, *67*, 8–12. [CrossRef]
29. Xing, R.; Zhou, B.; Trinajstić, N. Sum-connectivity index of molecular trees. *J. Math. Chem.* **2010**, *48*, 583–591. [CrossRef]
30. Wang, S.; Zhou, B.; Trinajstić, N. On the sum-connectivity index. *Filomat* **2011**, *25*, 29–42. [CrossRef]
31. Gutman, I. Geometric approach to degree-based topological indices: Sombor indices. *MATCH Commun. Math. Comput. Chem.* **2021**, *86*, 11–16.
32. Das, K.C.; Çevik, A.S.; Cangul, I.N.; Shang, Y. On Sombor index. *Symmetry* **2021**, *13*, 140. [CrossRef]
33. Kulli, V.R. Dharwad indices. *Int. J. Eng. Sci. Res. Technol.* **2021**, *10*, 17–21. [CrossRef]
34. Shanmukha, M.C.; Usha, A.; Siddiqui, M.K.; Asare-Tuah, A. Novel degree-based topological descriptors of carbon nanotubes. *J. Chem.* **2021**, *2021*, 3734185. [CrossRef]
35. Doley, A.; Buragohain, J.; Bharali, A. Inverse sum indeg status index of graphs and its applications to octane isomers and benzenoid hydrocarbons. *Chemom. Intell. Lab. Syst.* **2020**, *203*, 104059. [CrossRef]
36. Wang, G.; Yan, L.; Zaman, S.; Zhang, M. The connective eccentricity index of graphs and its applications to octane isomers and benzenoid hydrocarbons. *Int. J. Quantum Chem.* **2020**, *120*, e26334. [CrossRef]
37. Aldemir, M.S.; Ediz, S.; Yamaç, K. Angular geometric indices. *MATI* **2019**, *1*, 52–57.
38. Fath-Tabar, G.; Furtula, B.; Gutman, I. Correlation of domination parameters with physicochemical properties of octane isomers. *Appl. Math. Nonlinear Sci.* **2016**, *1*, 345–352. [CrossRef]
39. Fath-Tabar, G.; Furtula, B.; Gutman, I. A new geometric-arithmetic index. *J. Math. Chem.* **2010**, *47*, 477–486.
40. Al-Fahemi, J.H.; Albis, N.A.; Gad, E.A. QSPR models for octane number prediction. *J. Theor. Chem.* **2014**, *2014*, 520652. [CrossRef]
41. Alsharafi, M.; Zeren, Y. Degree-based Inverse Prodeg indices linking graph theory and cheminformatics with sharp results and QSPR modeling. *Sci. Rep.* **2026**. [CrossRef] [PubMed]
42. Gao, W.; Siddiqui, M.K.; Imran, M.; Jamil, M.; Farahani, M.R. Forgotten topological index of chemical structure in drugs. *Saudi Pharm. J.* **2016**, *24*, 258–264. [CrossRef] [PubMed]
43. Bilal, M.; Jamil, M.K.; Waheed, M.; Alameri, A. Three topological indices of two new variants of graph products. *Math. Probl. Eng.* **2021**, *2021*, 7724177. [CrossRef]
44. Liu, J.B.; Wang, C.; Wang, S.; Wang, B. Zagreb indices and multiplicative Zagreb indices of Eulerian graphs. *Bull. Malays. Math. Sci. Soc.* **2019**, *42*, 67–78.
45. Shao, Z.; Siddiqui, M.K.; Muhammad, M.H. Computing Zagreb indices and Zagreb polynomials for symmetrical nanotubes. *Symmetry* **2018**, *10*, 244. [CrossRef]
46. Alsharafi, M.; Zeren, Y. Degree-Based Prodeg Indices in Graph Theory and Cheminformatics with Sharp Results and QSPR Modeling: Data and Codes. *Zenodo* **2026**. [CrossRef]
47. RDKit. RDKit: Open-Source Cheminformatics. Available online: <https://www.rdkit.org> (accessed on 26 February 2026).
48. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
49. Brezovnik, S.; Tratnik, N.; Žigert Pleteršek, P. Weighted Wiener indices of molecular graphs with application to alkenes and alkadienes. *Mathematics* **2021**, *9*, 153. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.