symmetry

MDPI

Article

Harmonic–Arithmetic Index: Trees with Maximum Degrees and Comparative Analysis of Antidrugs

Kalpana Ramesh and Shobana Loganathan *

Department of Mathematics, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur 603203, India; kr0274@srmist.edu.in
* Correspondence: shobanal@srmist.edu.in

Abstract: Chemical graph theory connects the analysis of molecular structures with mathematical graph theory, allowing for the prediction of chemical and physical properties through the use of topological indices. Among these, the recently introduced Harmonic–Arithmetic (HA) index, proposed by Abeer M. Albalahi et al. in 2023, offers a novel method to quantify molecular and graph structures. It is defined as $HA(G) = \sum_{\mu\omega \in E(G)} \frac{4d_G(\mu)d_G(\omega)}{(d_G(\mu)+d_G(\omega))^2}$, where $d_G(\mu)$ and $d_G(\omega)$ are degrees of nodes μ and ω in G. In this paper, the HA index examines the bounds for a tree T of order n, with a maximum degree Δ. The application of the HA index extends to QSPR/QSAR analyses, where topological indices play a crucial role in predicting the relationship between molecular structures and physicochemical properties, such as in Parkinson's, disease-related antibiotics by calculating their topological indices and analyzing them using QSPR models. Comparative analyses were performed between linear regression models and curvilinear-approach quadratic and cubic regression models to identify the minimal RMSE and enhance predictive accuracy for physicochemical properties. The results demonstrate that the HA index effectively connects mathematical graph theory with molecular characterization, offering reliable predictions, dependable bounds for tree graphs, and meaningful insights into molecular properties. These findings highlight the HA index's potential as a versatile and innovative tool in advancing chemical graph theory and its applications to real-world problems in chemistry.

Keywords: HA index; pendent node; bound; anti-Parkinson's drugs; correlation coefficient; comparison

check for updates

Academic Editor: Alice Miller

Received: 4 December 2024
Revised: 11 January 2025
Accepted: 13 January 2025
Published: 23 January 2025

Citation: Ramesh, K.; Loganathan, S. Harmonic–Arithmetic Index: Trees with Maximum Degrees and Comparative Analysis of Antidrugs. *Symmetry* **2025**, *17*, 167. https://doi.org/10.3390/sym17020167

Copyright: © 2025 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) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Consider an undirected simple connected graph $G = (V, E)$ comprising precisely n nodes and m edges. The set of nodes adjacent to a node μ in G is denoted as $N_G(\mu)$. The degree of a node μ, denoted as $d_G(\mu)$, represents the count of edges that are incident to the node μ. The maximum degree in graph G is represented as Δ.

Chemical graph theory [1] is a specialized area of computational chemistry [2] that applies graph-theoretical concepts to model the structures of molecules. In this field, atoms are represented as nodes (vertices), and chemical bonds are depicted as edges connecting these nodes. This representation allows chemists to analyze and predict a variety of molecular properties through numerical descriptors known as topological indices (TIs) such as the Wiener index, Randic index, Balaban index and so on. These indices capture crucial information about the molecule's connectivity, pathway lengths, and substructures, providing valuable insights into its overall properties and potential behavior.

Quantitative Structure–Activity Relationship (QSAR) and Quantitative Structure–Property Relationship (QSPR) models use these graph-derived descriptors as input features to predict specific chemical or biological properties. These models apply statistical and machine learning approaches to establish a relationship between the molecular structure and its properties, such as melting points, boiling points, solubility, and toxicity. The use of machine learning algorithms allows complex patterns to be identified in large datasets, improving the accuracy of predictions and aiding in the analysis of new compounds.

The integration of chemical graph theory with QSPR involves constructing molecular graphs and calculating relevant descriptors that will be used as features in predictive models. This process includes data preparation, model training, validation and evaluation, which are all crucial steps to ensure the robustness and generalizability of the model. Once validated, these QSPR models can be applied to predict properties for new, untested molecules, facilitating virtual screening and aiding in the design of new compounds with desired properties. The combined use of chemical graph theory and QSPR enables researchers to streamline the development of new substances and optimize existing ones, offering a powerful tool in modern chemistry and pharmaceutical research. During the 1970s, many degree-based topological indices were proposed as part of chemical graph theory to better understand the structure–activity relationships of molecules. These indices use the degree of each node in a molecular graph to derive numerical values that help predict the chemical and physical behavior of molecules. In [3], the author provides a summary of the key properties of degree-based topological indices and conducts a critical comparative analysis of these indices. This work aims to highlight their distinct characteristics and how they compare in terms of their application and effectiveness in modeling molecular properties.

The Wiener index, introduced by Herman Wiener in 1947 [4], is recognized as the first TI. The index, which measures the sum of shortest path distances between every pair of atoms in a graph representation of molecules, was shown to predict boiling points of alkanes (paraffins). This work established a predictive tool for characterizing and distinguishing chemical properties based on molecular topology, setting a new groundwork in theoretical and computational chemistry. The Randic index, introduced by M. Randic in 1975 [5], is considered one of the first degree-based topological indices. The index measures the product of the degrees of connected atoms; it has been influential in studying the correlation between molecular structure and chemical properties such as boiling points and stability. It is gaining attention both theoretically and practically. Researchers have made numerous modifications to the Randic connectivity index to improve its performance.

Albalahi et al. (2023) [6] introduced and examined the Harmonic–Arithmetic (HA) index, a new graph-theoretical descriptor derived from harmonic and arithmetic means of node degrees, to offer a unique perspective on molecular structure. The authors investigated its properties and established its bounds and applications in studying the structural attributes of molecular systems. The HA index determines its enhanced insights into molecular structures and physicochemical properties, with suggestions for continued research on generalization, enhanced bounds and extending applications.

It is defined as

$$HA(G) = \sum_{\mu\omega \in E(G)} \frac{4d_G(\mu)d_G(\omega)}{(d_G(\mu) + d_G(\omega))^2}$$

where $d_G(\mu)$ and $d_G(\omega)$ are degrees of nodes μ and ω in G.

In [7], the authors present new theoretical results on the HA index that establish bounds and relationships between the HA index and other well-known topological indices, contributing to the understanding of its behavior and applications in analyzing molecular structures. Refs. [7–9] motivated us to provide some new results using the HA index.

The primary objective of this work is to derive a boundary for the HA index of trees based on their number of nodes and maximum degree. In addition, comparative analysis with results from [10] reveals that the H index demonstrated highly significant performance in their study. Similarly, in this paper, the HA index shows highly significant potential, as it achieves the same value as the H index while demonstrating enhanced predictive accuracy. The use of curvilinear models, such as quadratic and cubic regressions, further underscores the HA index's efficacy in minimizing the root mean squared error (RMSE) and achieving superior predictive results. This highlights the HA index's highly significant role in advancing QSPR methodologies, particularly for analyzing Parkinson's disease [11] related antibiotics and supporting drug development efforts.

2. Trees with Bounds on the Harmonic–Arithmetic Index

The following are the terminologies and notations used in this section. In a rooted tree, one particular node is designated as the root. A node with a degree of one is known as a pendant node. A node that is next to a leaf is called a support node. If a node is next to at least two leaves, it is known as a strong support node.

A starlike tree is characterized by having a single node with a degree of at least three, known as the center. In a starlike tree, a leg is defined as the path from the center node to a leaf. When all the legs in a starlike tree are of length one, it is referred to as a star.

Let T represent a rooted tree at node f where $d_T(f) = \Delta$. Additionally, $\mathfrak{T}_{n,\Delta}$ denotes the collection of all starlike trees with a maximal degree Δ and n nodes.

Lemma 1. *If a node in $T \in \mathfrak{T}_{n,\Delta}$ with the exception of f has a degree greater than two, then there is $T^* \in \mathfrak{T}_{n,\Delta}$ such that $HA(T) < HA(T^*)$.*

Proof. Let $d_T(f) = \Delta$ be a root node which contains the maximum degree in T with $N_T(f) = \{f_1, f_2, \ldots f_\Delta\}$. Assume that $d_T(g) \geq 3$ with $N_T(g) = \{g_1, g_2, \ldots g_\zeta\}$ and $d_T(g_\zeta) = k$, such that g_ζ lies between the nodes f and g with the maximum possible distance.

Case 1: g is a strong support node of T.

Let g_1 and g_2 be two pendant nodes of T. Assume that T^* is obtained by joining $T - gg_1$ to the edge g_2g_1 (see Figure 1).

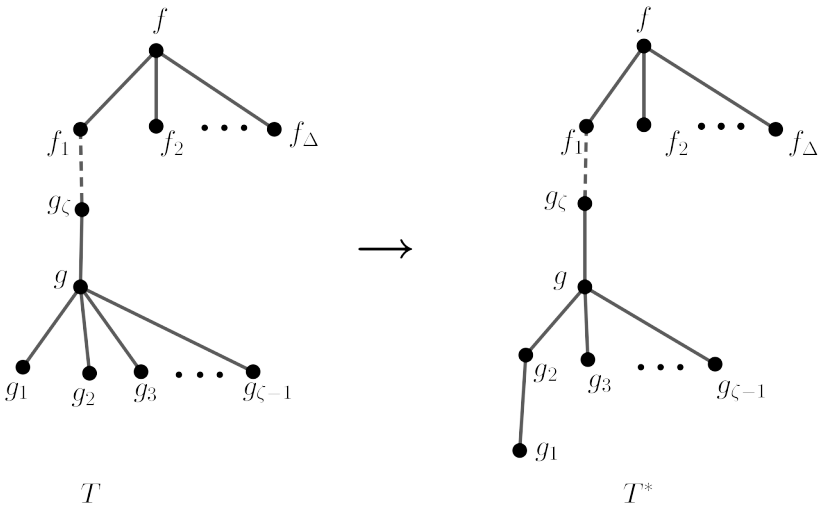


Figure 1. In Case 1, the tree transformation occurs when g is an end-support node.

$$
\begin{aligned}
HA(T) - HA(T^*) &= \sum_{\mu\omega \in E(T)} \frac{4d_T(\mu)d_T(\omega)}{(d_T(\mu)+d_T(\omega))^2} - \sum_{\mu\omega \in E(T^*)} \frac{4d_{T^*}(\mu)d_{T^*}(\omega)}{(d_{T^*}(\mu)+d_{T^*}(\omega))^2} \\
&= \sum_{p=3}^{\zeta-1} \frac{4d_T(g)d_T(g_p)}{(d_T(g)+d_T(g_p))^2} + \frac{4d_T(g)d_T(g_\zeta)}{(d_T(g)+d_T(g_\zeta))^2} + \frac{4d_T(g)d_T(g_1)}{(d_T(g)+d_T(g_1))^2} + \frac{4d_T(g)d_T(g_2)}{(d_T(g)+d_T(g_2))^2} \\
&- \sum_{p=3}^{\zeta-1} \frac{4d_{T^*}(g)d_{T^*}(g_p)}{(d_{T^*}(g)+d_{T^*}(g_p))^2} - \frac{4d_{T^*}(g)d_{T^*}(g_\zeta)}{(d_{T^*}(g)+d_{T^*}(g_\zeta))^2} - \frac{4d_{T^*}(g)d_{T^*}(g_2)}{(d_{T^*}(g)+d_{T^*}(g_2))^2} - \frac{4d_{T^*}(g_1)d_{T^*}(g_2)}{(d_{T^*}(g_1)+d_{T^*}(g_2))^2} \\
&= \sum_{p=3}^{\zeta-1} \frac{4\zeta d_T(g_p)}{(\zeta+d_T(g_p))^2} + \frac{4\zeta k}{(\zeta+k)^2} + 2\left(\frac{4\zeta}{(\zeta+1)^2}\right) - \sum_{p=3}^{\zeta-1} \frac{4(\zeta-1)d_T^*(g_p)}{((\zeta-1)+d_T^*(g_p))^2} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} \\
&- \frac{4(\zeta-1)(2)}{((\zeta-1)+2)^2} - \frac{4(2)(1)}{(2+1)^2} \\
&\leq \frac{8\zeta}{(\zeta+1)^2} + \frac{4\zeta k}{(\zeta+k)^2} - \frac{4k(\zeta-1)}{((\zeta-1)+k)^2} - \frac{8(\zeta-1)}{(\zeta+1)^2} - \frac{8}{9}
\end{aligned}
$$

If $\zeta < k$, then

$$HA(T) - HA(T^*) \leq \frac{8\zeta}{(\zeta+1)^2} + \frac{4\zeta k}{(\zeta+k)^2} - \frac{4k(\zeta-1)}{((\zeta-1)+k)^2} - \frac{8(\zeta-1)}{(\zeta+1)^2} - \frac{8}{9} < 0$$

If $\zeta \geq k$, then

$$\frac{4\zeta k}{(\zeta+k)^2} - \frac{4k(\zeta-1)}{((\zeta-1)+k)^2} \leq \frac{1}{(2\zeta-1)^2}$$

Since $\zeta \geq 3$, then

$$
\begin{aligned}
HA(T) - HA(T^*) &\leq \frac{8\zeta}{(\zeta+1)^2} + \frac{4\zeta k}{(\zeta+k)^2} - \frac{4k(\zeta-1)}{((\zeta-1)+k)^2} - \frac{8(\zeta-1)}{(\zeta+1)^2} - \frac{8}{9} \\
&\leq \frac{8\zeta}{(\zeta+1)^2} + \frac{1}{(2\zeta-1)^2} - \frac{8(\zeta-1)}{(\zeta+1)^2} - \frac{8}{9} < 0.
\end{aligned}
$$

Case 2: g is a support node of T.

Consider g_1 to be a leaf, and $gh_1h_2 \ldots h_\rho$ is a path in a tree T where $\rho \geq 2$ ($h_1 = g_2$). Let T^* be a tree obtained from T by removing the edge gg_1 and adding the path $h_\rho g_1$ (refer to Figure 2).

$$
\begin{aligned}
HA(T) - HA(T^*) &= \sum_{\mu\omega \in E(T)} \frac{4d_T(\mu)d_T(\omega)}{(d_T(\mu)+d_T(\omega))^2} - \sum_{\mu\omega \in E(T^*)} \frac{4d_{T^*}(\mu)d_{T^*}(\omega)}{(d_{T^*}(\mu)+d_{T^*}(\omega))^2} \\
&= \sum_{p=2}^{\zeta-1} \frac{4d_T(g)d_T(g_p)}{(d_T(g)+d_T(g_p))^2} + \frac{4d_T(g)d_{(g_\zeta)}}{(d_T(g)+d_{(g_\zeta)})^2} + \frac{4d_T(g)d_T(g_1)}{(d_T(g)+d_T(g_1))^2} + \frac{4d_T(h_\rho)d_T(h_{\rho-1})}{(d_T(h_\rho)+d_T(h_\rho-1))^2} \\
&- \sum_{p=2}^{\zeta-1} \frac{4d_{T^*}(g)d_{T^*}(g)}{(d_{T^*}(g)+d_{T^*}(g))^2} - \frac{4d_{T^*}(g)d_{T^*}(g_\zeta)}{(d_{T^*}(g)+d_{T^*}(g_\zeta))^2} - \frac{4d_{T^*}(g)d_{T^*}(g_1)}{(d_{T^*}(g)+d_{T^*}(g_1))^2} - \frac{4d_{T^*}(h_\rho)d_{T^*}(h_{\rho-1})}{(d_{T^*}(h_\rho)+d_{T^*}(h_{\rho-1}))^2} \\
&= \sum_{p=2}^{\zeta-1} \frac{4\zeta d_T(g_p)}{(\zeta+d_T(g_p))^2} + \frac{4\zeta k}{(\zeta+k)^2} + \frac{4\zeta}{(\zeta+1)^2} + \frac{8}{9} - \sum_{p=2}^{\zeta-1} \frac{4(\zeta-1)d_{T^*}(g_p)}{((\zeta-1)+d_{T^*}(g_p))^2} \\
&- \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - \frac{8}{9} - 1 \\
&\leq \frac{4\zeta k}{(\zeta+k)^2} + \frac{4\zeta}{(\zeta+1)^2} - \frac{(\zeta-1)k}{((\zeta-1)+k)^2} - 1
\end{aligned}
$$

If $\zeta < k$, then

$$HA(T) - HA(T^*) \leq \frac{4\zeta}{(\zeta+1)^2} + \frac{4\zeta k}{(\zeta+k)^2} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 1 < 0$$

If $\zeta \geq k$, then

$$HA(T) - HA(T^*) \leq \frac{4\zeta}{(\zeta+1)^2} + \frac{4\zeta k}{(\zeta+k)^2} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 1$$
$$\leq \frac{4\zeta}{(\zeta+1)^2} + \frac{1}{(2\zeta-1)^2} - 1 < 0.$$

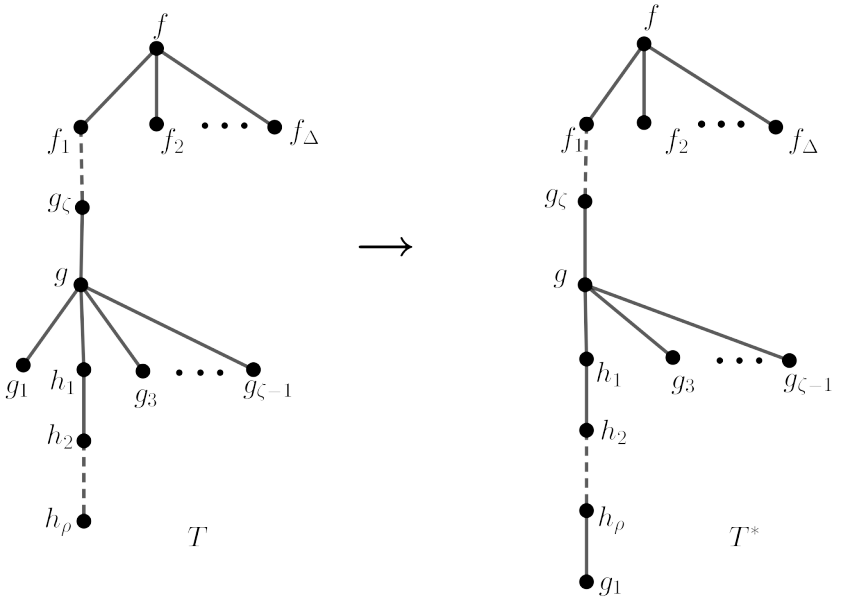


Figure 2. In Case 2, the tree transformation occurs when g is support node.

Case 3: g is not a support node.

Consider two paths $ga_1a_2\ldots a_s$ and $gb_1b_2\ldots b_t$ in a tree T ($s,t \geq 2, a_1 = g_1$ and $b_1 = g_2$). Let T^* be the tree obtained from T by removing the edge gg_1 and adding the path $b_ta_1a_2\ldots a_s$ (refer to Figure 3).

$$HA(T) - HA(T^*) = \sum_{\mu\omega\in E(T)} \frac{4d_T(\mu)d_T(\omega)}{(d_T(\mu)+d_T(\omega))^2} - \sum_{\mu\omega\in E(T^*)} \frac{4d_{T^*}(\mu)d_{T^*}(\omega)}{(d_{T^*}(\mu)+d_{T^*}(\omega))^2}$$
$$= \sum_{p=2}^{\zeta-1} \frac{4d_T(g)d_T(g_p)}{(d_T(g)+d_T(g_p))^2} + \frac{4d_T(g)d_T(g_\zeta)}{(d_T(g)+d_T(g_\zeta))^2} + \frac{4d_T(g)d_T(g_1)}{(d_T(g)+d_T(g_1))^2} + \frac{4d_T(b_t)d_T(b_{t-1})}{(d_T(b_t)+d_T(b_{t-1}))^2}$$
$$- \sum_{p=2}^{\zeta-1} \frac{4d_{T^*}(g)d_{T^*}(g_p)}{(d_{T^*}(g)+d_{T^*}(g_p))^2} - \frac{4d_{T^*}(g)d_{T^*}(g_\zeta)}{(d_{T^*}(g)+d_{T^*}(g_\zeta))^2} - \frac{4d_{T^*}(b_t)d_{T^*}(g_1)}{(d_{T^*}(b_t)+d_{T^*}(g_1))^2} - \frac{4d_{T^*}(b_t)d_{T^*}(b_{t-1})}{(d_{T^*}(b_t)+d_{T^*}(b_{t-1}))^2}$$
$$= \sum_{p=2}^{\zeta-1} \frac{4\zeta d_T(g_p)}{(\zeta+d_T(g_p))^2} + \frac{4\zeta k}{(\zeta+k)^2} + \frac{8\zeta}{(\zeta+2)^2} + \frac{8}{9} - \sum_{p=2}^{\zeta-1} \frac{4\zeta d_{T^*}(g_p)}{(\zeta+d_{T^*}(g_p))^2} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 1 - 1$$
$$\leq \frac{4\zeta k}{(\zeta+k)^2} + \frac{8\zeta}{(\zeta+2)^2} + \frac{8}{9} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 2.$$

If $\zeta < k$, then

$$HA(T) - HA(T^*) \leq \frac{4\zeta k}{(\zeta+k)^2} + \frac{8\zeta}{(\zeta+2)^2} + \frac{8}{9} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 2 < 0$$

If $\zeta \geq k$, then

$$HA(T) - HA(T^*) \leq \frac{4\zeta k}{(\zeta+k)^2} + \frac{8\zeta}{(\zeta+2)^2} + \frac{8}{9} - \frac{4(\zeta-1)k}{((\zeta-1)+k)^2} - 2$$
$$\leq \frac{8\zeta}{(\zeta+2)^2} + \frac{8}{9} + \frac{1}{(2\zeta-1)^2} - 2 < 0.$$

The proof that $HA(T) < HA(T^*)$ is concluded by the aforementioned three cases. □

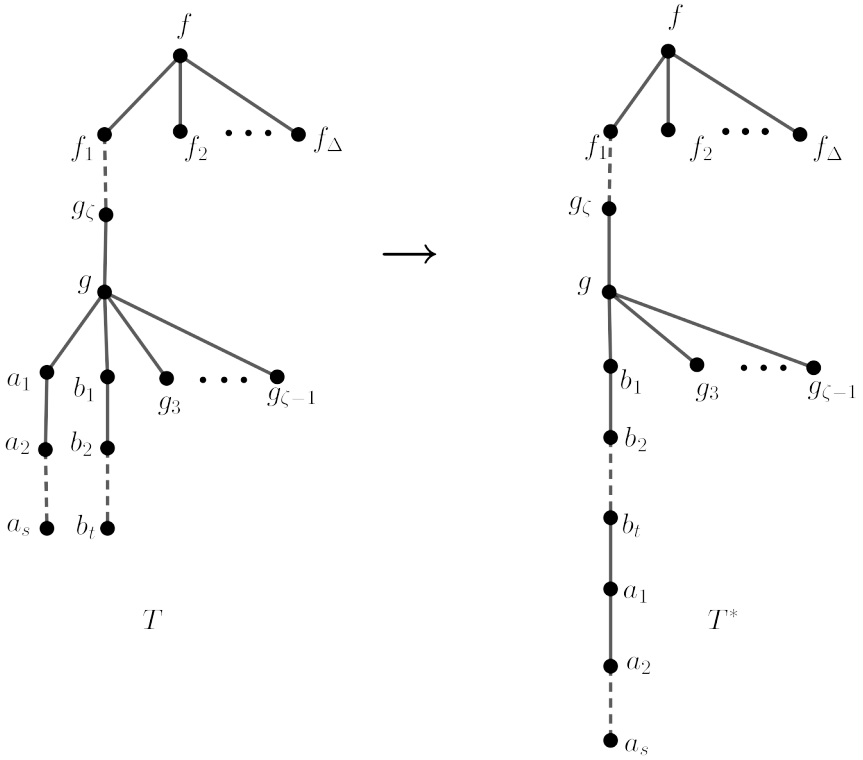


Figure 3. In Case 3, the tree transformation occurs when g is not support node.

Lemma 2. *Let $T \in \mathfrak{T}_{n,\Delta}$ be a starlike tree with $\Delta \geq 3$ and $d_T(f) = \Delta$ be the maximum degree of a node. If there are at least two legs, each with a length greater than one, in T, then there is $T^* \in \mathfrak{T}_{n,\Delta}$, such that $HA(T) > HA(T^*)$.*

Proof. Let T be a starlike tree with a root node of $d_T(f) = \Delta$ and $N_T(f) = f_1, f_2, \ldots f_\Delta$. Assume that $fg_1g_2 \ldots g_\zeta$ and $fz_1z_2 \ldots z_\rho$, $(\zeta, \rho > 1)$ are the paths of two legs such that $g_1 = f_1$ and $z_1 = f_2$. Let T^* be obtained by joining $T - g_1g_2..g_\zeta$ to the edge $z_\rho g_2..g_\zeta$ (refer to Figure 4).

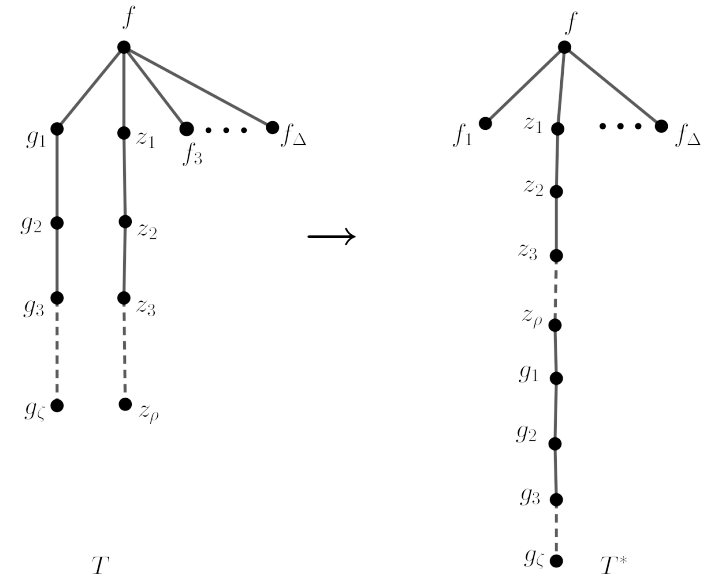


Figure 4. Illustration of transformation.

$$
\begin{aligned}
HA(T) - HA(T^*) &= \sum_{\mu\omega \in E(T)} \frac{4d_T(\mu)d_T(\omega)}{(d_T(\mu)+d_T(\omega))^2} - \sum_{\mu\omega \in E(T^*)} \frac{4d_{T^*}(\mu)d_{T^*}(\omega)}{(d_{T^*}(\mu)+d_{T^*}(\omega))^2} \\
&= \sum_{p=2}^{\Delta} \frac{4d_T(f)d_T(f_p)}{(d_T(f)+d_T(f_p))^2} + \frac{4d_T(g_1)d_T(g_2)}{(d_T(g_1)+d_T(g_2))^2} + \frac{4d_T(f)d_T(f_1)}{(d_T(f)+d_T(f_1))^2} + \frac{4d_T(z_\rho)d_T(z_{\rho-1})}{(d_T(z_\rho)+d_T(z_{\rho-1}))^2} \\
&- \sum_{p=2}^{\Delta} \frac{4d_{T^*}(f)d_{T^*}(f_p)}{(d_{T^*}(f)+d_{T^*}(f_p))^2} - \frac{4d_{T^*}(f)d_{T^*}(f_1)}{(d_{T^*}(f)+d_{T^*}(f_1))^2} - \frac{4d_{T^*}(z_\rho)d_{T^*}(g_2)}{(d_{T^*}(z_\rho)+d_{T^*}(g_2))^2} - \frac{4d_{T^*}(z_\rho)d_{T^*}(z_{\rho-1})}{(d_{T^*}(z_\rho)+d_{T^*}(z_{\rho-1}))^2} \\
&= \frac{4d_T(g_1)d_T(g_2)}{(d_T(g_1)+d_T(g_2))^2} - \frac{4d_{T^*}(z_\rho)d_{T^*}(g_2)}{(d_{T^*}(z_\rho)+d_{T^*}(g_2))^2} + \frac{8\Delta}{(\Delta+2)^2} - \frac{4\Delta}{(\Delta+1)^2} - \frac{1}{9}.
\end{aligned}
$$

If $\rho = 2$, then

$$\frac{4(d_{T^*}(z_\rho))d_{T^*}(g_2)}{(d_{T^*}(z_\rho)+d_{T^*}(g_2))^2} = \frac{4(d_T(g_1))d_T(g_2)}{(d_T(g_1)+d_T(g_2))^2} = \frac{8}{9}$$

If $\rho \geq 3$, then

$$\frac{4(d_{T^*}(z_\rho))d_{T^*}(g_2)}{(d_{T^*}(z_\rho)+d_{T^*}(g_2))^2} = \frac{4(d_T(g_1))d_T(g_2)}{(d_T(g_1)+d_T(g_2))^2} = 1$$

Finally,

$$HA(T) - HA(T^*) = \frac{8\Delta}{(\Delta+2)^2} - \frac{4\Delta}{(\Delta+1)^2} - \frac{1}{9} > 0$$

Thus, $HA(T) > HA(T^*)$ completes the proof. □

3. Main Results

Theorem 1. *If $T \in \mathfrak{T}_{n,\Delta}$ of order $n \geq 3$,*

$$HA(T) \leq (\Delta-1)\frac{4\Delta}{(\Delta+1)^2} + \frac{8\Delta}{(\Delta+2)^2} + (n-\Delta-2) + \frac{8}{9} \quad \textit{if} \quad \Delta < n-1$$

Moreover, the lower bound is attained if and only if a node has a degree greater than two except Δ and the equality is attained if and only if a starlike tree with a maximum of one leg with a length exceeding one is denoted as T.

Proof. Assume that if $\Delta = 2$ then T^* is a path graph and

$$HA(T^*) = (\Delta-1)\frac{4\Delta}{(\Delta+1)^2} + \frac{8\Delta}{(\Delta+2)^2} + (n-\Delta-2) + \frac{8}{9} \tag{1}$$

as required.

Let $T \in \mathfrak{T}_{n,\Delta}$ be a tree with $HA(T) \leq HA(T^*)$ for $T^* \in \mathfrak{T}_{n,\Delta}$. By the choice of T from Lemma 1, T is a starlike tree with centre f. Now, assume that T has only one leg with a length exceeding one; then,

$$HA(T) = (\Delta-1)\frac{4\Delta}{(\Delta+1)^2} + \frac{8\Delta}{(\Delta+2)^2} + (n-\Delta-2) + \frac{8}{9} \tag{2}$$

Hence, the proof is complete. □

Theorem 2. *Consider $T \in \mathfrak{T}_{n,\Delta}$ of order $n \geq 3$,*

$$HA(T) \geq (\Delta - 1)\frac{4\Delta}{(\Delta+1)^2} + \frac{8\Delta}{(\Delta+2)^2} + (n - \Delta - 2) + \frac{8}{9} \quad \textit{if} \quad \Delta < n - 1$$

Moreover, the upper bound is attained if and only if a node has a degree greater than one at least for two legs and the equality is attained if and only if a starlike tree with a maximum of one leg with a length exceeding one is denoted as T.

Proof. Assume that if $\Delta = 2$ then T is a path graph and Equation (1) holds. Let $T^* \in \mathfrak{T}_{n,\Delta}$ be a tree with $HA(T) \geq HA(T^*)$ for $T \in \mathfrak{T}_{n,\Delta}$. By the choice of T^* from Lemma 2, T^* is denoted as a starlike tree with centre f. Now, assume that T^* has exactly only one leg with a length exceeding one; then, Equation (2) holds.

Hence, the proof is complete. □

Observation: For any tree $T \in \mathfrak{T}_{n,\Delta}$ of order $n \geq 3$,

$$HA(T) \geq \frac{4\Delta^2}{(\Delta+1)^2} \quad \text{if} \quad \Delta = n - 1$$

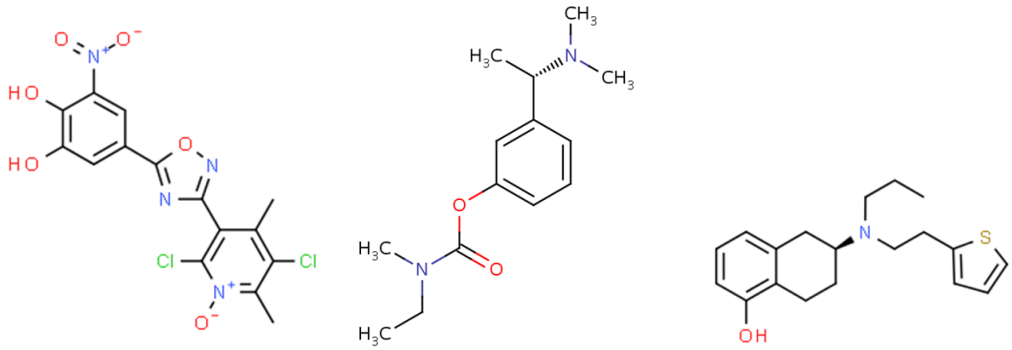
and the equality is attained if and only if T is a star.

4. Application: Chemical Connection of HA Index

The QSPR analysis of the Harmonic–Arithmetic (HA) topological index can be applied to a variety of drugs and demonstrates the physical and chemical properties that are highly correlated with the topological index of anti-Parkinson's drugs. Also, correlation coefficients with the Harmonic index in [10] and HA index are compared.

This section concentrates on the computation of a Harmonic–Arithmetic (HA) topological index for 18 Parkinson's disease antibiotics, namely Opicapone, Rivastigmine, Rotigotine, Istradefylline, Safinamide, Trihexyphenidyl, Tolcapone, Ropinirole, Procyclidine, Pramipexole, Piribedil, Pergolide, Levodopa, Entacapone, Carbidopa, Biperiden, Apomorphine and Amantadine. The QSPR analysis of topological indices is examined using a curvilinear regression model.

The molecular structures (see Figure 5) and physicochemical property data along with the HA index values of the listed drugs are presented in Table 1. Both the physicochemical property data and molecular structures were obtained from ChemSpider.



(**a**) (**b**) (**c**)

Figure 5. *Cont.*

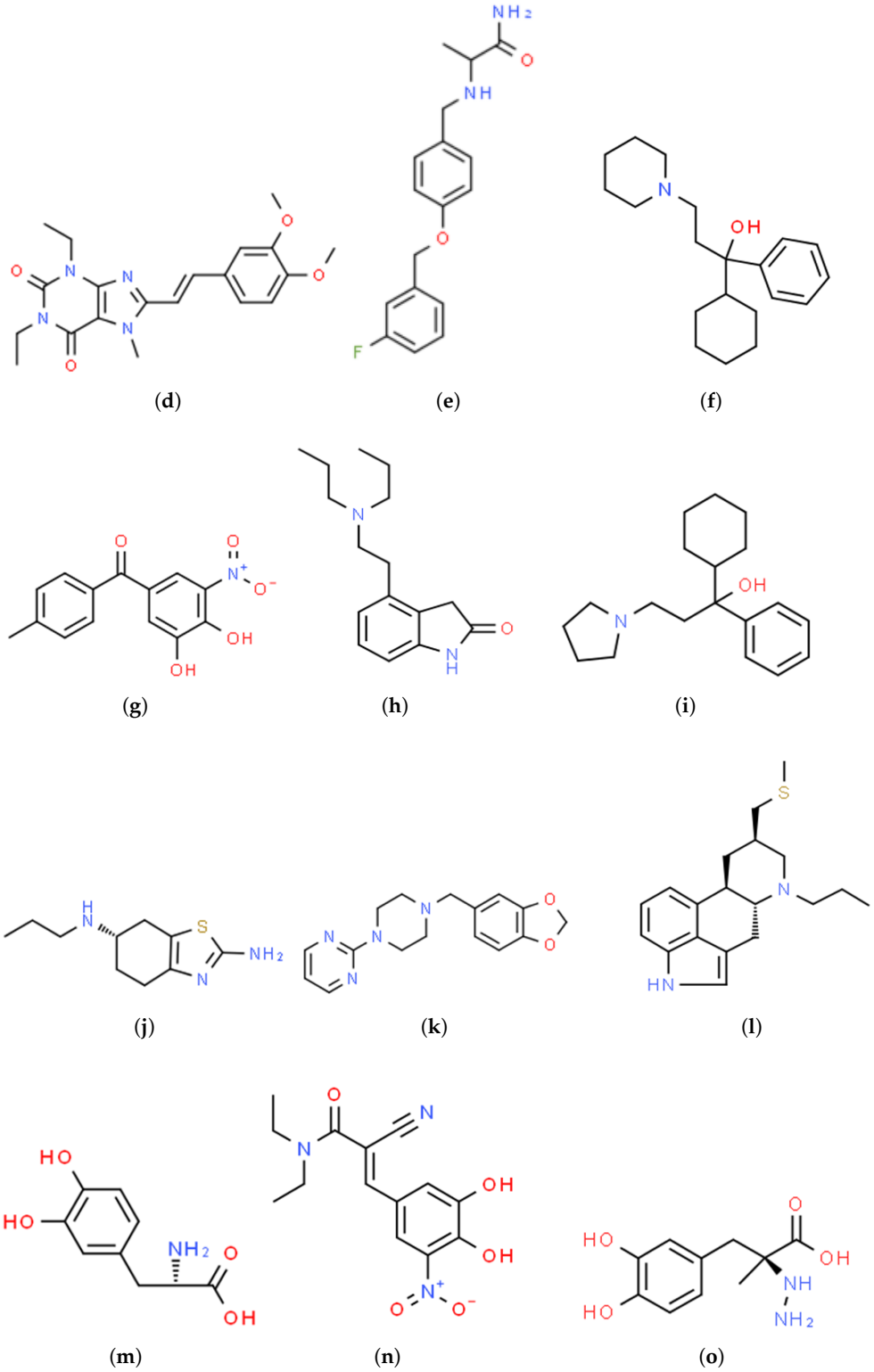
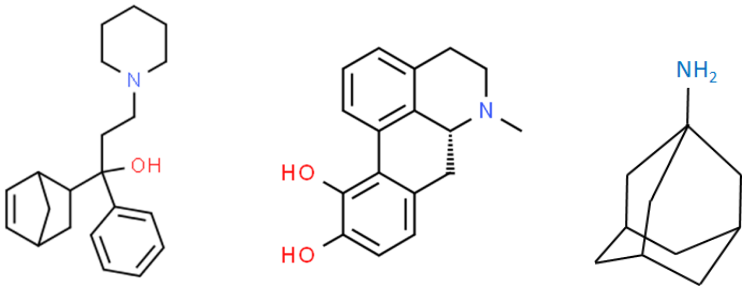


Figure 5. *Cont.*



(**p**) (**q**) (**r**)

Figure 5. Molecular structures: anti-Parkinson's drugs. (**a**) Opicapone; (**b**) Rivastigmine; (**c**) Rotigotine; (**d**) Istradefylline; (**e**) Safinamide; (**f**) Trihexyphenidyl; (**g**) Tolcapone; (**h**) Ropinirole; (**i**) Procyclidine; (**j**) Pramipexole; (**k**) Piribedil; (**l**) Pergolide; (**m**) Levodopa; (**n**) Entacapone; (**o**) Carbidopa; (**p**) Biperiden; (**q**) Apomorphine; (**r**) Amantadine.

Table 1. Physicochemical properties and HA index value of anti-Parkinson's drugs.

Anti-Drugs	B.P.	E.	F.P.	M.R.	P.	M.V.	HA(G)
Opicapone	701.1	106.3	377.8	92.5	36.7	229.3	26.43
Rivastigmine	316.2	55.8	145	73.1	29	241.2	16.36
Rotigotine	470.1	76.1	238.1	94.6	37.5	272.4	23.20
Istradefylline	601	89.4	317.3	105.6	41.8	309.2	28.33
Safinamide	476.7	74	242.1	83.3	33	254.2	21.52
Trihexyphenidyl	447.9	74.4	211	91.8	36.4	289.7	23.21
Tolcapone	485.6	78	205.7	71.2	28.2	192.5	19.18
Ropinirole	410.5	66.3	202	78.4	31.1	250.2	19.13
Procyclidine	433.5	72.7	205.7	87.2	34.6	271.9	22.21
Pramipexole	378	62.6	182.4	60.3	23.9	180.5	14.42
Piribedil	469.4	73.2	237.7	81.5	32.3	228.7	24.44
Pergolide	491.3	75.8	250.9	97.4	38.6	279.6	24.30
Levodopa	448.4	74.5	225	49.3	19.5	134.3	12.51
Entacapone	526.6	83.1	272.3	79.1	31.4	219.2	18.96
Carbidopa	528.7	84.6	273.5	57.5	22.8	159.3	14.09
Biperiden	462.1	76.2	224.5	94.1	37.3	281.7	25.05
Apomorphine	473.4	76.5	268.8	77.9	30.9	205.6	21.93
Amantadine	225.7	46.2	96	45.7	18.1	141.8	11.95

4.1. Regression Models

The Harmonic–Arithmetic (HA) topological index is modelled using six physical properties of the eighteen anti-Parkinson's drugs: boiling point (BP), enthalpy (E), flash point (FP), molar refractivity (MR), polarity (P) and molar volume (MV) (refer to Figure 5).

The following regression models are considered for our comparative study:

$$Z = \beta + \alpha_1[TI_1] \quad \text{(Equation for linear regression)} \tag{3}$$

$$Z = \beta + \alpha_1[TI_2] + \alpha_2[TI_2^2] \quad \text{(Equation for quadratic regression)} \tag{4}$$

$$Z = \beta + \alpha_1[TI_3] + \alpha_2[TI_3^2] + \alpha_3[TI_3^3] \quad \text{(Equation for cubic regression)} \tag{5}$$

where Z, β, $\alpha_i (i = 1,2,3)$ and $TI_i (i = 1,2,3)$ are the physical properties of the drug (dependent variable), constant, regression coefficient and topological index (independent variable).

4.2. Regression Models for Harmonic–Arithmetic Index HA(G)

(i) Harmonic–Arithmetic index $HA(G)$ of linear regression model is

BP = 193.557 + 13.24$[HA(G)]$

E = 42.734 + 1.5699$[HA(G)]$

FP = 71.17 + 7.8829$[HA(G)]$

MR = 11.578 + 3.3007$[HA(G)]$

P = 4.5884 + 1.3085$[HA(G)]$
MV = 47.177 + 8.965$[HA(G)]$

(ii) The Harmonic–Arithmetic index $HA(G)$ of the quadratic regression model is
BP = 421.193 − 11.451$[HA(G)]$ + 0.629$[HA(G)]^2$
E = 72.521 − 1.6611$[HA(G)]$ + 0.0823$[HA(G)]^2$
FP = 248.11 − 11.31$[HA(G)]$ + 0.4891$[HA(G)]^2$
MR = −17.392 + 6.443$[HA(G)]$ − 0.0801$[HA(G)]^2$
P = −7.1538 + 2.5822$[HA(G)]$ − 0.0325$[HA(G)]^2$
MV = −96.811 + 24.583$[HA(G)]$ − 0.398$[HA(G)]^2$

(iii) The Harmonic–Arithmetic index $HA(G)$ of the cubic regression model is
BP = −1485.1 + 296.4$[HA(G)]$ − 15.2$[HA(G)]^2$ + 0.2614$[HA(G)]^3$
E = −124.39 + 30.139$[HA(G)]$ − 1.5528$[HA(G)]^2$ + 0.027$[HA(G)]^3$
FP = −771.55 + 153.36$[HA(G)]$ − 7.978$[HA(G)]^2$ + 0.1398$[HA(G)]^3$
MR = −133.23 + 25.151$[HA(G)]$ − 1.042$[HA(G)]^2$ + 0.0159$[HA(G)]^3$
P = −53.032 + 9.9914$[HA(G)]$ − 0.4134$[HA(G)]^2$ + 0.0063$[HA(G)]^3$
MV = −530.44 + 94.613$[HA(G)]$ − 3.9988$[HA(G)]^2$ + 0.0595$[HA(G)]^3$

It was noted that the curvilinear regression model demonstrated significantly superior RMSE and r in the models that exhibited a decrease in RMSE values, an increase in r values and a p-value less than 0.005, which indicates strong statistical significance in a QSPR model, thereby concluding that it is a superior regression model. In Tables 2–4, the correlation between the HA topological index and the six physical properties of anti-Parkinson's drugs is represented by the statistical parameters. The curvilinear variation between MR and the HA index is illustrated in Figure 6a; the curvilinear variation between P and the HA index is illustrated in Figure 6b; and the curvilinear variation between MV and the HA index is illustrated in Figure 6c.

Table 2. Statistical parameters for linear regresstion model: QSPR analysis of HA of a graph.

Physical Properties	r	r^2	F	SE	MSE	$RMSE$	P	Indicator
BP	0.407	0.638	10.963	80.550	5767.401	75.943	0.001	Sig
E	0.598	0.358	8.907	10.595	99.783	9.989	0.009	Sig
FP	0.622	0.387	10.109	49.941	2216.988	47.084	0.006	Sig
MR	0.954	0.910	162.524	5.215	24.176	4.917	0.000	Sig
P	0.954	0.910	161.483	2.07	8.968	2.995	0.000	Sig
MV	0.841	0.707	38.534	29.090	752.213	27.427	0.000	Sig

Statistical parameters: r—correlation coefficient; r^2—coefficient of determination; F—ratio of two variables; SE—standard error; MSE—mean squared error; $RMSE$—root mean squared error; P—probability value; Sig—significant.

Table 3. Statistical parameters for quadratic regresstion model: QSPR analysis of HA of a graph.

Physical Properties	r	r^2	F	SE	MSE	$RMSE$	P	Indicator
BP	0.652	0.426	5.555	81.855	5583.452	74.723	0.016	Sig
E	0.615	0.378	4.556	10.768	96.634	9.830	0.028	Sig
FP	0.646	0.418	5.384	50.270	2105.862	45.890	0.017	Sig
MR	0.960	0.921	87.943	5.044	21.198	4.604	0.000	Sig
P	0.960	0.921	87.905	2.000	3.335	1.826	0.000	Sig
MV	0.858	0.735	20.835	28.537	678.624	26.050	0.000	Sig

Table 4. Statistical parameters for cubic regresstion model: QSPR analysis of HA of a graph.

Physical Properties	r	r^2	F	SE	MSE	$RMSE$	P	Indicator
BP	0.657	0.432	5.693	81.425	4956.376	70.402	0.014	Sig
E	0.619	0.383	4.658	10.723	89.943	9.484	0.027	Sig
FP	0.652	0.425	5.553	49.944	1926.464	43.892	0.016	Sig
MR	0.960	0.921	87.943	5.044	18.920	4.350	0.000	Sig
P	0.960	0.921	87.905	2.000	2.981	1.727	0.000	Sig
MV	0.858	0.735	20.835	28.537	646.435	25.425	0.000	Sig

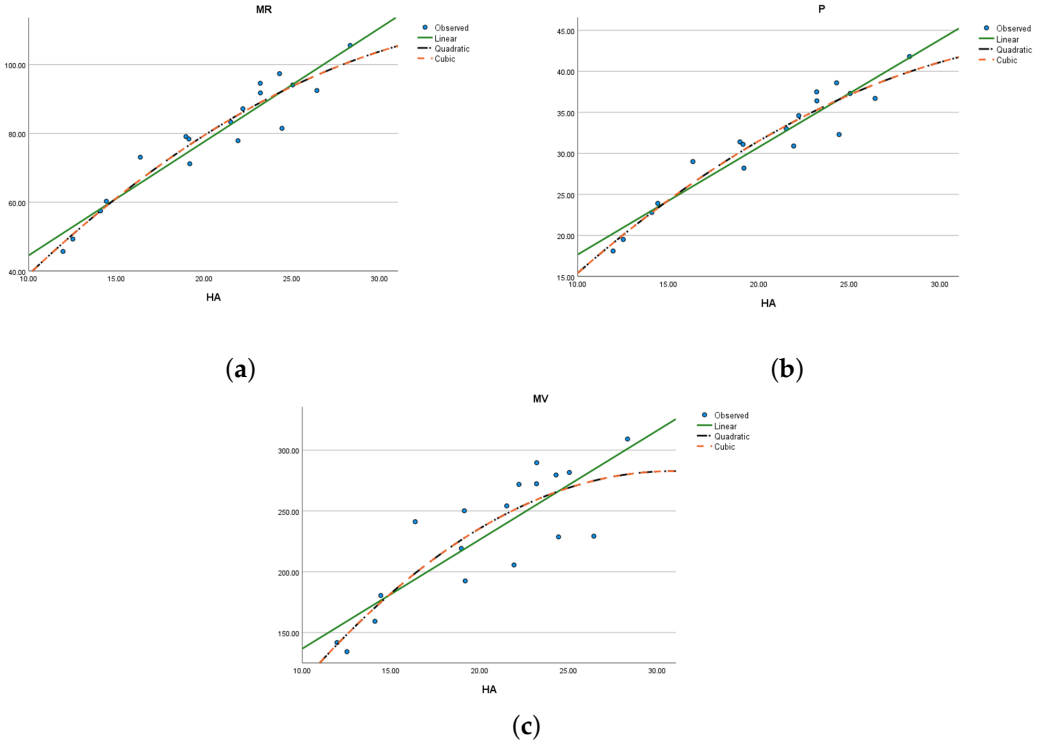


Figure 6. The curvilinear regression equation graphs between (**a**) MR and HA (**b**), between P and HA and (**c**) between MV and HA.

In [10], a statistical analysis of Parkinson's disease was examined using linear regression models, and then the authors calculated the values that showed strong positive correlations between various topological indices and the physical properties of anti-Parkinson's drugs. Among the 14 topological indices analyzed, the symmetric division degree index (SDD) and the Harmonic index (H) exhibit highly significant positive correlations. Upon comparing the H index in [10] with the HA index, we observed that MR, P and MV have the most similar correlation coefficient values. This significance is based on choosing the maximum correlation coefficient (r) and minimum RMSE value in the linear regression model (refer to Figure 7).

Observation:

- Similar to the linear regression model, the quadratic and cubic regression models yield the highest correlation coefficient r for the H index and HA index in relation to physical properties (PP), with nearly identical values.
- The correlation coefficient values for the quadratic and cubic regression models are follows (see Figure 8):
  The correlation coefficients for the H index with MR and the HA index with MR are as follows: $r = 0.968$ and $r = 0.96$, respectively.
  The correlation coefficients for the H index with P and the HA index with P are as follows: $r = 0.968$ and $r = 0.96$, respectively.

Also, the correlation coefficients for the H index with MV and the HA index with MV are as follows: $r = 0.87$ and $r = 0.858$, respectively.

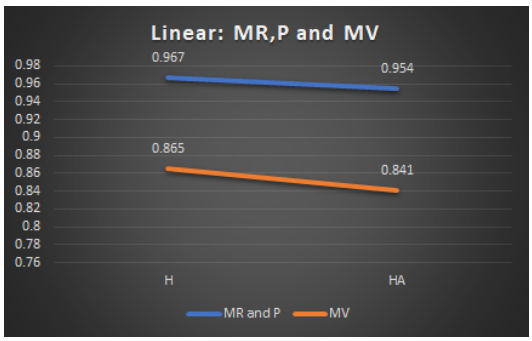
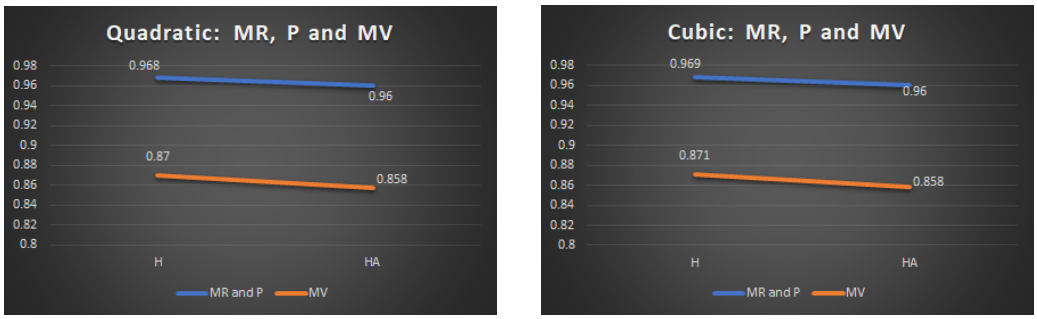


Figure 7. Comparition of the correlation coefficient of topological indices with properties: MR, P and MV (linear regression model).



(**a**) (**b**)

Figure 8. Comparison of the correlation coefficients of topological indices with properties: (**a**) quadratic MR, P and MV; (**b**) cubic MR, P and MV.

5. Discussion

The HA index is distinctive due to its unique properties and the way it quantifies molecular and graph structures, setting it apart from other topological indices. In Section 2, this paper highlighted the unique application of the HA index in comparing different tree structures $T \in \mathfrak{T}_{n,\Delta}$. While previous research established that these trees have a minimum first irregularity Sombor index [12], this study demonstrates that they can exhibit both maximum and minimum HA indices, showcasing the versatility and unique application of the HA index in characterizing graph structures.

In Section 3, the QSPR analysis of the Harmonic–Arithmetic (HA) topological index on various drugs reveals its potential in identifying physical and chemical properties strongly correlated with the topological structure of anti-Parkinson's drugs. The comparison of correlation coefficients between the Harmonic index in [10] and the HA index demonstrates that both indices are effective in correlating with the drugs' physical properties. The quadratic and cubic regression models, like the linear model, yield similar high correlation coefficients r for the HA index and H index, indicating their strong relevance to property prediction. This highlights the application of the HA index as a valuable tool in QSPR analyses for evaluating and predicting the physicochemical attributes of drug compounds.

6. Conclusions

In this paper, we computed the bounds for the Harmonic–Arithmetic (HA) index of starlike trees with a maximum node degree Δ and also conducted a comparative study between the H index and the HA index using the QSPR model to support chemists in discovering novel drugs because it applies regression techniques to predict molecular prop-

erties, typically in fields like drug design, material science and environmental chemistry. Future research could explore the HA index's application in a broader range of molecular structures and its integration with machine learning algorithms to enhance predictive modeling. Additionally, investigating the HA index's performance across different classes of drugs and comparing it with other emerging indices could provide further insights into its reliability and adaptability. Finally, extending this study to real-world pharmacokinetic and pharmacodynamic data would deepen our understanding of its practical implications in drug development.

Author Contributions: K.R.: Writing—review and editing; S.L.: supervision and validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data are included within the article.

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

References

1. Trinajstic, N. *Chemical Graph Theory*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 1992. [CrossRef]
2. Sadybekov, A.V.; Katritch, V. Computational approaches streamlining drug discovery. *Nature* **2023**, *616*, 673–685. [CrossRef] [PubMed]
3. Gutman, I. Degree-based topological indices. *Croatica Chem. Acta* **2013**, *86*, 351–361. [CrossRef]
4. Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20. [CrossRef] [PubMed]
5. Randic, M. Characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. [CrossRef]
6. Albalahi, A.M.; Ali, A.; Alanazi, A.M.; Bhatti, A.A.; Hamza, A.E. Harmonic-arithmetic index of (molecular) trees. *arXiv* **2023**, arXiv:2302.11099. [CrossRef]
7. Albalahi, A.M.; Ali, A.; Alanazi, A.M.; Bhatti, A.A.; Hamza, A.E. Inequalities involving the harmonic-arithmetic index. *Afr. Mat.* **2024**, *35*, 46. [CrossRef]
8. Dehgardi, N.; Liu, J.B. Lanzhou index of trees with fixed maximum degree. *MATCH Commun. Math. Comput. Chem.* **2021**, *86*, 3–10.
9. Kirana, B.; Shanmukha, M.; Usha, A. A QSPR analysis and curvilinear regression models for various degree-based topological indices: Quinolone antibiotics. *Heliyon* **2024**, *10*, e32397. [CrossRef] [PubMed]
10. Ramesh, K.; Loganathan, S.; Cangul, I.N. QSPR analysis of antiparkinson's drugs with degree based topological indices. *Montes Taurus J. Pure Appl. Math.* **2024**, *6*, 239–254.
11. Goetz, C.G. The history of Parkinson's disease: Early clinical descriptions and neurological therapies. *Cold Spring Harb. Perspect. Med.* **2011**, *1*, 008862. [CrossRef] [PubMed]
12. Dehgardi, N.; Shang, Y. First irregularity Sombor index of trees with fixed maximum degree. *Res. Math.* **2024**, *11*, 2291933. [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.