Three New/Old Vertex–Degree–Based Topological Indices

Ivan Gutman¹, Boris Furtula¹, Clive Elphick²

¹Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia
gutman@kg.ac.rs, boris.furtula@gmail.com
²clive.elphick@gmail.com

(Received January 25, 2014)

Abstract

Three vertex–degree-based graph invariants are presented, that earlier have been considered in the chemical and/or mathematical literature, but that evaded the attention of most mathematical chemists. These are the reciprocal Randić index (RR), the reduced second Zagreb index $RM_2$, and the reduced reciprocal Randić index (RRR). If $d_1, d_2, \ldots, d_n$ are the degrees of the vertices of the graph $G = (V, E)$, then

$$RR = \sum_{ij \in E} \sqrt{d_i d_j}, \quad RM_2 = \sum_{ij \in E} (d_i - 1)(d_j - 1), \quad RRR = \sum_{ij \in E} \sqrt{(d_i - 1)(d_j - 1)}.$$

We outline the literature sources of these topological indices, their main mathematical properties, and establish their correlating abilities w.r.t. characteristic physico-chemical properties of alkanes.

1 Introduction

Let $G$ be a simple graph with $n$ vertices and $m$ edges, with vertex set $V(G) = \{v_1, v_2, \ldots, v_n\}$ and edge set $E(G)$. The edge connecting the vertices $v_i$ and $v_j$ will be denoted by $ij$.

The degree of the vertex $v_i$, denoted by $d_i$, is the number of first neighbors of $v_i$ in the underlying graph.
In the chemical literature, several dozens of vertex–degree–based topological indices have been and are currently considered and applied in QSPR/QSAR studies. Three oldest and most thoroughly investigated are the first Zagreb index $M_1$, the second Zagreb index $M_2$, and the Randić (or connectivity) index $R$. These are defined as

\[
M_1 = M_1(G) = \sum_{v_i \in V(G)} (d_i)^2 \quad (1)
\]

\[
M_2 = M_2(G) = \sum_{ij \in E(G)} d_i d_j \quad (2)
\]

\[
R = R(G) = \sum_{ij \in E(G)} \frac{1}{\sqrt{d_i d_j}} \quad . \quad (3)
\]

In addition, we mention here the geometric–arithmetic index $GA$

\[
GA = GA(G) = \sum_{ij \in E(G)} \frac{\sqrt{d_i d_j}}{(d_i + d_j)/2} \quad . \quad (4)
\]

Details of these and other degree–based topological indices can be found in the books [20,21,28], surveys [5–7,19,34–36], recent review [18], and the references cited therein. Comparative studies of the chemical applicability of these indices were recently reported [16,23]. Three degree–based indices that we consider in the present paper, although not truly new, have so far not been studied in mathematical chemistry. In particular, these are not mentioned at all either in the works [5–7,16,19–21,23,28,34–36] or in the handbooks [37,38]. The aim of the present article is to contribute towards filling this gap.

The new/old topological indices studied in the present paper are the following.

• The **reciprocal Randić index** is defined as

\[
RR = RR(G) = \sum_{ij \in E(G)} \sqrt{d_i d_j} \quad . \quad (5)
\]

It is, of course, a special case of the earlier much examined “general Randić index” $\sum_{ij \in E(G)} (d_i d_j)^\alpha$, where $\alpha$ is a variable parameter (see [20,28]). Yet, in the theory of general Randić index, the choice $\alpha = 1/2$ did not attract any particular attention.

The invariant $RR$ seems to be first encountered in a paper by Favaron, Mahéo, and Saclé [15], but only in connection with a marginal result stated in their
Proposition 2.8 (see Theorem 2 below). The usage of $RR$ for measuring graph irregularity [11] is outlined in the subsequent section.

- The **reduced second Zagreb index** is defined as

$$\text{RM}_2 = \text{RM}_2(G) = \sum_{ij \in E(G)} (d_i - 1)(d_j - 1). \quad (6)$$

Because of the identity

$$\text{RM}_2(G) = M_2(G) - M_1(G) + m \quad (7)$$

this graph invariant is necessarily encountered within studies of the difference between the two Zagreb indices [3,17,32]. In addition, if the graph $G$ is a tree, then $\text{RM}_2(G)$ is equal to the number of pairs of vertices at distance 3 [9,27,29,40], which in mathematical chemistry is often referred to as the “Wiener polarity index”.

- In the same manner as the reduced second Zagreb index, Eq. (6), is related with the ordinary second Zagreb index, Eq. (2), the **reduced reciprocal Randić index**

$$\text{RRR} = \text{RRR}(G) = \sum_{ij \in E(G)} \sqrt{(d_i - 1)(d_j - 1)} \quad (8)$$

might be viewed as the reduced analogue of the reciprocal Randić index, Eq. (5). However, short time ago, the invariant $\text{RRR}$ has appeared in the physico-chemical literature in a fully unrelated context [31]: within a novel approach for modeling normal boiling points of hydrocarbons.

In the later parts of this paper we refer to three standard special graphs: the star $S_n$, the path $P_n$, and the complete graph $K_n$. The star $S_n$ is the $n$-vertex tree in which one vertex has degree $n - 1$. The path $P_n$ is the $n$-vertex tree in which exactly two vertices have degree one. The complete graph $K_n$ is the $n$-vertex graph in which the degree of all vertices is $n - 1$.

## 2 Measuring Irregularity of Graphs

If $G$ is a graph of order $n$ with vertex degrees $d_1, d_2, \ldots, d_n$ and if

$$d_1 = d_2 = \cdots = d_n \quad (9)$$
then $G$ is said to be regular. If, in addition, $d_1 = d_2 = \cdots = d_n = r$, then $G$ is regular of degree $r$. If condition (9) is not satisfied, then $G$ is said to be irregular. In order to establish “how irregular” a given graph is, several irregularity measures, denoted here by $irr(G)$, have been proposed. Each of such measures is required to have the property $irr(G) \geq 0$ and $irr(G) = 0$ if the graph $G$ is regular.

In chemical applications, quantifying the irregularity of molecular graphs and of biomolecular networks seems to be of marginal importance. Yet, some results along these lines have been achieved [12,16,22].

The oldest measure of graph irregularity is that of Collatz and Sinogowitz [4], namely $\mu - d$, where $\mu$ is the largest eigenvalue of the $(n,m)$-graph $G$, and $d = 2m/n$ is its average vertex degree. It is based on the relation [4]

$$\mu \geq \frac{2m}{n}$$

in which equality holds if and only if the graph $G$ is regular. Therefore, it is convenient to write the Collatz–Sinogowitz measure as

$$irr_{CS}(G) = \frac{n\mu}{2m} - 1.$$ (10)

Other proposed irregularity measures are in terms of vertex degrees, which indeed is the natural option. Edwards [10] considered the measure $irr_E$ defined via

$$\nu = \nu(G) := 1 + irr_E(G)^2 = \frac{n}{4m^2} \sum_{v_i \in V(G)} (d_i)^2.$$ (11)

We immediately recognize its close relation with the first Zagrebel index, Eq. (1):

$$irr_E(G) = \sqrt{\frac{n M_1(G)}{4m^2} - 1}.$$ Condition $irr_E(G) \geq 0$ is equivalent to the well-known inequality $M_1(G) \geq 4m^2/n$ in which equality holds if and only if $G$ is regular.

A mathematically equivalent approach to Edwards’ irregularity is that of Bell [2], whose measure is the variance of the vertex degrees:

$$irr_B(G) = \frac{1}{n} \sum_{v_i \in V(G)} (d_i)^2 - \left(\frac{1}{n} \sum_{v_i \in V(G)} d_i\right)^2.$$ Consequently,

$$irr_B(G) = \frac{M_1(G)}{n} - \left(\frac{2m}{n}\right)^2.$$
Albertson’s measure \([1]\) is defined as
\[
irr_A(G) = \sum_{ij \in E(G)} |d_i - d_j| .
\]
(12)

Various classes of graphs whose \(irr_A\)-value is extremal have been studied \([14,30,41]\). It is worth noting that, not being aware of Albertson’s work \([1]\), Fath-Tabar proposed \([13]\) that the quantity on the right-hand side of Eq. (12) be called the “third Zagreb index”.

The papers \([22,24]\) report comparative studies of the irregularity measures \(irr_B\) and \(irr_A\).

Wocjan and one of the present authors \([11]\) considered the edge-variant of Edwards’ measure, Eq. (11), and introduced
\[
\epsilon = \epsilon(G) := 1 + irr_{ew}(G)^2 = \frac{n}{2m^2} \sum_{ij \in E(G)} \sqrt{d_i d_j} .
\]
(13)
which in view of Eq. (5) is tantamount to
\[
irr_{ew}(G) = \sqrt{\frac{n \text{RR}(G)}{2m^2}} - 1 .
\]
(14)

The Elphick–Wocjan measure of irregularity is based on Theorem 5, proven in Subsection 3.1.

We see that some of the measures of graph irregularity are intimately connected with degree-based topological indices. This observation could be further strengthened by pointing out that the Randić index, Eq. (3), and the geometric-arithmetic index, Eq. (4), could serve for constructing irregularity measures.

For a connected graph \(G\) with \(n\) vertices, \(R(G) \leq n/2\), with equality if and only if \(G\) is regular \([15,28]\). For a connected graph with \(m\) edges, \(GA(G) \leq m\), with equality if and only if \(G\) is regular \([39]\). Therefore,
\[
irr_R(G) = 1 - \frac{2R(G)}{n}
\]
\[
irr_{GA}(G) = 1 - \frac{GA(G)}{m}
\]
could be viewed as measures of graph irregularity. Analogous irregularity measures could be designed by means of any other degree-based topological index.
3 Mathematical Properties of Topological Indices $RR$, $RRR$, and $RM_2$

3.1 Bounds

As before, by $\mu$ we denote the largest eigenvalue of the adjacency matrix of the graph $G$. The following result seems to have been first reported by Hofmeister in 1988 [25].

**Theorem 1.** [25] Let the first Zagreb index $M_1(G)$ be defined via Eq. (1), and let $\mu$ be the largest eigenvalue of the $(n,m)$-graph $G$. Then,

$$\mu \geq \sqrt{\frac{M_1(G)}{n}}$$

with equality if and only if $G$ is regular.

The next two theorems were obtained by Favaron, Mahéo, Saclé [15] as early as in 1993.

**Theorem 2.** [15] Let the reciprocal Randić index $RR(G)$ be defined via Eq. (5), and let $\mu$ be the largest eigenvalue of the $(n,m)$-graph $G$. Then,

$$\mu \geq \frac{RR(G)}{m}.$$  

Equality is attained in the case of regular graphs, but the necessary and sufficient conditions for equality are much more complicated, see Proposition 2.8 in [15].

**Theorem 3.** [15] Let the Randić index $R(G)$ and the reciprocal Randić index $RR(G)$ of the $(n,m)$-graph $G$ be defined via Eqs. (3) and (5). Then,

$$R(G) \cdot RR(G) \geq m^2$$

with equality if and only if $G$ is regular.

**Proof.** Replace $R(G) \cdot RR(G)$ by the right–hand sides of Eqs. (3) and (5), use the Cauchy–Schwartz inequality, and recall that $|E(G)| = m$. □

We now consider three measures of irregularity, $irr_{cs}$, $\nu$, and $\epsilon$, as defined by Eqs. (10), (11), and (13), respectively.
Theorem 4. Let $G$ be a graph of order $n$, with no isolated vertices. Then,

$$
\left( \text{irr}_{cs}(G) + 1 \right)^2 \leq \frac{n^2}{4(n-1)} \quad (15)
$$

$$
\nu(G) \leq \frac{n^2}{4(n-1)} \quad (16)
$$

$$
\epsilon(G)^2 \leq \frac{n^2}{4(n-1)} \quad (17)
$$

with equality in all three cases if $G \cong S_n$.

Proof.

Upper bound (15): Hong [26] proved that $\mu^2 \leq 2m - n + 1$ for connected graphs. Therefore,

$$
\left( \text{irr}_{cs} + 1 \right)^2 = \frac{n^2 \mu^2}{4m^2} \leq \frac{n^2 (2m - n + 1)}{4m^2} \leq \frac{n^2}{4(n-1)}
$$

because of

$$
0 \leq m^2 - 2m(n-1) + (n-1)^2 = (m - (n-1))^2
$$

which is true for all $m$ and $n$.

Upper bound (16): In [11] it has been shown that $\nu \leq [n \mu/(2m)]^2$, so (16) immediately follows from the bound (15).

Upper bound (17): In [11] it also has been shown that $\epsilon \leq n \mu/(2m)$, so (17) immediately follows from the bound (15).

It is straightforward to show that all three bounds (15)–(17) are exact for the star graph.

For a graph with clique number $\omega$, Nikiforov has proved that [33]

$$
\mu^2 \leq 2m(\omega - 1)/\omega .
$$

Bearing this result in mind, we get

$$
\left( \text{irr}_{cs} + 1 \right)^2 = \frac{n^2 \mu^2}{4m^2} \leq \frac{2n^2 m(\omega - 1)}{4m^2 \omega} = \frac{n^2 (\omega - 1)}{2m \omega} \leq \frac{n^2}{2m} .
$$

Therefore for triangle–free graphs, $\text{irr}_{cs} \leq \sqrt{n^2/4m} - 1$. 

Theorem 5. Let $G$ be a graph with $n$ vertices and $m$ edges. Let $\epsilon(G)$ and $\text{irr}_{EW}(G)$ be the quantities defined via Eqs. (13) and (14). Then $\epsilon(G) \geq 1$ and, equivalently, $\text{irr}_{EW}(G) \geq 0$ holds, with equalities if and only if the graph $G$ is regular.

Proof. Bearing in mind Theorem 3, we get

$$\epsilon = \frac{n \cdot RR}{2m^2} = \frac{n \cdot R \cdot RR}{2R \cdot m^2} = \frac{nm^2}{2R \cdot m^2} = \frac{n}{2R}.$$ 

For every edge $ij \in E(G)$,

$$\left( \frac{1}{\sqrt{d_i}} - \frac{1}{\sqrt{d_j}} \right)^2 \geq 0$$

from which it follows

$$\frac{2}{\sqrt{d_i} d_j} \leq \frac{1}{d_i} + \frac{1}{d_j}.$$ 

Therefore, by Eq. (3),

$$\frac{n}{2R} = \frac{n}{2} \left[ \sum_{ij \in E(G)} \frac{1}{\sqrt{d_i} d_j} \right]^{-1} \geq n \left[ \sum_{ij \in E(G)} \left( \frac{1}{d_i} + \frac{1}{d_j} \right) \right]^{-1} = 1$$

because of

$$\sum_{ij \in E(G)} \left( \frac{1}{d_i} + \frac{1}{d_j} \right) = n.$$  \hspace{1cm} (18)

Recall that (18) is just a special case of a more general identity, recently reported in [8].

The fact that $\epsilon(G) = 1$ and $\text{irr}_{EW}(G) = 0$ for regular graphs is directly verified from Eqs. (13) and (14). \hfill \Box

It should be noted that the result of Theorem 5 is equivalent to the relation

$$RR(G) \geq \frac{2m^2}{n}$$

in which, again, equality holds if and only if $G$ is regular.
3.2 An identity for reduced second Zagreb index

The reduced second Zagreb index $RM_2$ is defined via Eq. (6). By taking into account the definitions of the first and second Zagreb indices, Eqs. (1) and (2), one straightforwardly arrives at the equality (7).

Several authors [9,27,29,32,40] have recognized that if the edge $ij$ does not belong to a triangle, then the term $(d_i - 1)(d_j - 1)$ is equal to the number of pairs of vertices of which one is adjacent to $v_i$ and the other to $v_j$. In other words, $(d_i - 1)(d_j - 1)$ is equal to the number of $P_4$-subgraphs of the graph $G$, whose central edge is $ij$. Consequently, $RM_2$ is equal to the total number of 4-vertex path–subgraphs of $G$. This implies:

**Theorem 6.** If $G$ is a triangle–free graph (which includes the case of trees), then $RM_2(G) = p_4$, where $p_4$ is the number of subgraphs of $G$ isomorphic to $P_4$.

The above theorem is stated in several papers [9,27,29,40]. If, in addition, the graph $G$ does not possess cycles whose size is smaller than 7 (which, again, includes the case of trees), then $p_4$ coincides with the number of vertex pairs at distance 3 (the “Wiener polarity number”) [29].

In the general case, when the graph possesses triangles, Theorem 6 has to be modified as [32]:

**Theorem 7.** If $G$ is a graph with $t$ triangles, then $RM_2(G) = p_4 + 3t$.

3.3 Graphs and trees with extremal RR, RRR, and RM$_2$ indices

**Theorem 8.** Let $G$ be a connected graph of order $n$. Then

$$\begin{align*}
RR(P_n) & \leq RR(G) \leq RR(K_n) \\
RM_2(S_n) & \leq RM_2(G) \leq RM_2(K_n) \\
RRR(S_n) & \leq RRR(G) \leq RRR(K_n).
\end{align*}$$

Equalities in (19) are attained if and only if $G \cong P_n$ and $G \cong K_n$, respectively. Analogous conditions apply also for equalities in (20) and (21).
Proof. All inequalities in (19)–(21), except the left one in (19), are immediate and obvious consequences of the definitions (5), (6), and (8).

From Eq. (5), it is evident that by deleting an edge from the graph \( G \), its RR-value will diminish. By deleting from \( G \) as many edges as possible, so that the subgraph thus obtained remains connected, we arrive at a tree. In Theorem 3.3.6 in the book \[28\] it is proven that in the case of \( n \)-vertex trees, \( n \geq 5 \), and for \( \alpha > 0 \), the quantity \( \sum_{ij \in E(G)} (d_i d_j)\alpha \) attains its minimal value for \( G \cong P_n \). The left–hand side bound in (19) is the special case of this result for \( \alpha = 1/2 \). \qed

Theorem 9. If \( T \) is a tree of order \( n \), then
\[
RR(P_n) \leq RR(T) \leq RR(S_n).
\]
Equality is attained if and only if, \( T \cong P_n \) and \( T \cong S_n \), respectively.

Proof. The left inequality is a special case of (19). The right inequality is a special case of Theorem 3.3.2 in the book \[28\], according to which the star is the \( n \)-vertex tree with maximal \( \sum_{ij \in E(G)} (d_i d_j)\alpha \) for all \( 0 < \alpha < 1 \). \qed

If \( T \) is a tree of order \( n \), then, evidently, the star \( S_n \) has minimal \( RM_2 \)- and \( RRR \)-values, both equal to zero. The case of trees with maximal \( RM_2 \)- and \( RRR \)-values is somewhat more complicated.

Let \( n \) and \( \Delta \) be fixed integers, \( n \geq 4 \), \( 2 \leq \Delta \leq n - 2 \). Construct the set \( T(n, \Delta) \) of \( n \)-vertex trees by attaching (in any possible way) \( n - \Delta - 1 \) pendent vertices to the pendent vertices of the star \( S_{\Delta+1} \).

Theorem 10. Let \( n \) be a fixed integer, \( n \geq 4 \).

(a) If \( n \) is even, then the \( n \)-vertex trees with greatest \( RM_2 \)-value are the elements of \( T(n, n/2) \). For all \( T \in T(n, n/2) \), this maximal value is \( RM_2(T) = \frac{1}{4} (n - 2)^2 \).

(b) If \( n \) is odd, then the \( n \)-vertex trees with greatest \( RM_2 \)-value are the elements \( T(n, \lfloor n/2 \rfloor) \cup T(n, \lceil n/2 \rceil) \). For all \( T \in T(n, \lfloor n/2 \rfloor) \cup T(n, \lceil n/2 \rceil) \), this maximal value is \( RM_2(T) = \frac{1}{4} (n - 1)(n - 3) \).

The proof of Theorem 10 is lengthy and has been communicated elsewhere \[17\]. An illustrative example is given in Fig. 1.
Fig. 1. Trees of order \( n = 10 \) and \( n = 11 \) with maximal reduced second Zagreb indices: \( T(10, 5) = \{T_1, T_2, T_3, T_4, T_5\} \); \( T(11, 5) = \{T_6, T_7, T_8, T_9, T_{10}, T_{11}, T_{12}\} \); \( T(11, 6) = \{T_{13}, T_{14}, T_{15}, T_{16}, T_{17}\} \). Note that \( T_6 \) and \( T_{13} \) are isomorphic; therefore there is a total of 11 mutually non-isomorphic trees of order 11, with maximal \( RM_2 \).

Let \( n \) be a fixed integer, \( n \geq 3 \). If \( n = 2k \), construct the tree \( T_{RRR}(n) \) by attaching one pendent vertex to \( k - 1 \) pendent vertices of \( S_{k+1} \), not attaching anything to a single pendent vertex of \( S_{k+1} \), see Fig. 2. If \( n = 2k + 1 \), construct the tree \( T_{RRR}(n) \) by attaching one pendent vertex to each of the \( k \) pendent vertices of the star \( S_{k+1} \), see Fig. 2.

Fig. 2. The tree \( T_{RRR}(n) \) which by Theorem 11 has greatest reduced reciprocal Randić index.

The following result was obtained by an extensive computer search (embracing trees up to \( n = 20 \)). At the present moment its exact proof is not known. We,
nevertheless, state it in the form of a theorem.

**Theorem 11.** The tree of order 6 with maximal \( RRR \)-value is \( P_6 \). The tree of order 8 with maximal \( RRR \)-value is the tree obtained by attaching a pendent vertex to a pendent vertex of \( T_{RRR}(7) \). For \( n \geq 3, n \neq 6,8 \), the tree of order \( n \) with maximal \( RRR \)-value is \( T_{RRR}(n) \).

4 Correlation of Topological Indices \( RR \), \( RRR \), and \( RM_2 \) with Physico–Chemical Parameters of Alkanes

In order to get some preliminary information on whether the three new/old topological indices possess any potential applicability in chemistry (especially in QSPR/QSAR studies), we have correlated them with two characteristic physico–chemical properties. The standard enthalpy of formation (\( \Delta H_f^\circ \)) was chosen to represent thermodynamic (energetic) molecular properties. The normal boiling point (\( b.p. \)) was chosen as a feature reflecting intermolecular (van der Waals type) interactions. As in our earlier studies [16,23], the set consisting of all isomeric octanes was used, for which molecular–size–based effects, as well as effects caused by polar groups can be disregarded.

In Table 1 are given the correlation coefficients (\( r \)) of the studied correlations. In addition to \( RR \), \( RM_2 \) and \( RRR \), we report \( r \)-values also for the Randić index (\( R \)), geometric–arithmetic index (\( GA \)), atom–bond connectivity index (\( ABC \)), and augmented Zagreb index (\( AZI \)). The latter two topological indices were included into the test because according to earlier studies [16,23], these possess the best correlating abilities. (More details on \( ABC \) and \( AZI \) can be found in the recent review [18].)

<table>
<thead>
<tr>
<th>index</th>
<th>( r(\Delta H_f^\circ) )</th>
<th>( r(b.p.) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( RR )</td>
<td>−0.640</td>
<td>−0.609</td>
</tr>
<tr>
<td>( RM_2 )</td>
<td>0.059</td>
<td>0.086</td>
</tr>
<tr>
<td>( RRR )</td>
<td>0.898</td>
<td>0.877</td>
</tr>
<tr>
<td>( R )</td>
<td>0.850</td>
<td>0.821</td>
</tr>
<tr>
<td>( GA )</td>
<td>0.858</td>
<td>0.823</td>
</tr>
<tr>
<td>( ABC )</td>
<td>−0.894</td>
<td>−0.863</td>
</tr>
<tr>
<td>( AZI )</td>
<td>0.922</td>
<td>0.923</td>
</tr>
</tbody>
</table>
Table 1. Correlation coefficients ($r$) of correlations between selected degree–based topological indices and two physico–chemical properties of octane isomers. For details see text.

As expected [16,23], also in this case, $AZI$ was found to possess the best correlating ability among the examined degree–based topological indices. As a sort of promising surprise, the new $RRR$-index performed slightly better than the previously second–ranked $ABC$-index. Thus, the reduced reciprocal Randić index qualifies to be included among structure–descriptors deserving attention of colleagues performing QSPR/QSAR studies. The present results also corroborate the work of Manso et al. [31], whose correlations pertain to sets of non-isomers, in which case the effect of molecular size usually screens any other topology–based influence.

In contrast to this, the reciprocal Randić index and, especially, the reduced second Zagreb index did not pass our tests. These were found to be inferior to the standard and most frequently employed molecular structure descriptors (such as the Randić index), and their further utilization for QSPR/QSAR purposes cannot be recommended.

Acknowledgement. B. F. thanks the Ministry of Science and Education of Serbia for financial support through the project No. 174033.

References


