6th International Conference on
Combinatorics, Cryptography, Computer Science and Computing
November. 17-18, 2021

Chemical significance of Zagreb coindices.

Melaku Berhe Belay[1]
Addis Ababa Science and Technology University, Addis Ababa, 16417, Ethiopia

Chunxiang Wang
Central China Normal University, Wuhan, 430079, P.R. China.

Abstract

Molecular structure descriptors(or topological indices) are numerical parameters associated with chemical constitution for correlation of chemical structures with their various physical properties, chemical reactivity or biological activity. So far the degree (valency) based topological indices are the oldest and most successful class of descriptors. A molecular graph is a representation of the structural formula of a chemical compound whose vertices correspond to the atoms and edges correspond to the chemical bonds. In this paper, the chemical applicability of the Zagreb coindices is studied and their predictive power is compared with the well known degree based indices so-called the Zagreb indices.

Keywords: Topological index, Zagreb indices and coindices, octane isomers, chemical graphs.
Mathematics Subject Classification [2010]: 05C90, 05C76, 05C07.

1 Introduction

In the fields of chemical graph theory, molecular topology, and mathematical chemistry, a topological index also known as a connectivity index is a type of a molecular descriptor that is calculated based on the molecular graph of a chemical compound. Topological indices are numerical parameters of a graph which characterize its topology and are usually graph invariant. Topological indices are used for example in the development of quantitative structure-activity relationships (QSARs) in which the properties of molecules are correlated with their chemical structure[6, 23].These properties related by topological indices with the molecular structures could be physical properties, biological activity or chemical reactivity.

In recent years, topological indices have been used to study the molecular complexity, drug design, chemical documentation, isomer discrimination, chirality, lead optimization database selection, in deriving multilinear regression models and rational combinatorial library design. Especially in the field of medicine the notions of topological indices help pharmacists by providing several information based upon the structure of materials which reduce their workload. Computing topological indices of a compound may help in approximating its medicinal behavior. The idea of understanding compounds through different topological indices has gained significant importance in the field of medicine because it requires only molecular graph computation with no chemical-related apparatus to study. It is believed that absence of degeneracy and high discriminating power are two properties of an ideal topological index. Among all topological indices the descriptors based on valencies (degrees of end vertices of molecular graphs) of atoms in molecule have been used for several chemical resources. The Randić index [20], one of the best-known topological indices in chemical graph theory, belongs to this class of indices. There is no doubt that Randić index is the most studied, most often

[1]speaker

applied and most popular among all topological indices. In fact many papers and few books are devoted to this molecular descriptor (see, e.g.[17, 15, 16, 22] and the references cited therein).
In mathematical chemistry, the molecular structures are represented by undirected graphs where each vertex represents an atoms and each edge represents the chemical bond between these atoms. So in this paper, all the graphs considered are simple, connected and undirected graphs. Let G be a finite simple graph on n vertices and m edges. The vertex set and edge set of G are denoted respectively by $V(G)$ and $E(G)$. The complement of a graph G which is denoted by $\overline{G}$ is the simple graph with the same vertex set $V(G)$ and any two vertices $uv \in E(\overline{G})$ if and only if $uv \notin E(G)$. Since $E(\overline{G}) \cup E(G) = E(K_n)$, we have that $\overline{m} = n(n-1)/2 - m$, where $\overline{m} = |E(\overline{G})|$. The degree of a vertex v in G is denoted by $d_G(v)$ and the degree of the same vertex v in $\overline{G}$ is then given by $d_{\overline{G}}(v) = n - 1 - d_G(v)$.
In [12] Gutman and Trinajstić studied the structure-dependency of total π -electron energy and the following approximate formula were obtained

$$M_1(G) = \sum_{u \in V(G)} d_G(u)^2 \tag{1}$$

and

$$M_2(G) = \sum_{uv \in E(G)} [d_G(u) d_G(v)]. \tag{2}$$

Immediately, it was recognized that they have applications in studying the extent of branching of the carbon-atom skeleton i.e they provide quantitative measures of molecular branching. Because of their applicability Balaban et al.[3] included these quantities in topological indices and named them as "Zagreb group indices". Nowadays M_1 and M_2 are known as the first and second Zagreb index respectively. Alternative expression for the first Zagreb index which is defined on the edges of G, is given in [18] as

$$M_1(G) = \sum_{uv \in E(G)} [d_G(u) + d_G(v)]. \tag{3}$$

These indices were also proposed to measure molecular complexity [19]. Zagreb indices are also among thoroughly studied valency (degree) based topological indices. Some results regarding Zagreb indices can be referred to the reviews[18, 5, 25]. Taking the contributions of non adjacent pairs of vertices into consideration, the first ($\overline{M}_1$) and second ($\overline{M}_2$) Zagreb coindices were introduced originally by Došlic [8] while computing weighted Wiener polynomial of some composite graphs. These invariants were introduced for second time in [1] where the author of [8] is participated. In both papers, the first and second Zagreb coindices are defined respectively as

$$\overline{M}_1(G) = \sum_{uv \notin E(G)} [d_G(u) + d_G(v)] \tag{4}$$

and

$$\overline{M}_2(G) = \sum_{uv \notin E(G)} [d_G(u) d_G(v)]. \tag{5}$$

Some mathematical results on the Zagreb coindices can be referred to [1, 2, 4, 14, 10, 13]. Note that the Zagreb indices are defined over pairs of adjacent vertices whereas Zagreb coindices are defined over pairs of non-adjacent vertices.
Gutman [11, 9] have studied comparative testing on how well some degree based topological indices are correlated with two physico-chemical properties of octane isomers. These two properties were chosen to be the normal boiling points and standard heats of formation. The normal boiling point were considered as representative for van-dar-Waals-type interactions and standard heats of formation as representative for thermochemical properties. Even though the Zagreb indices and Randić index are among the studied once, Gutman did not include Zagreb coindices in his studies. So far, to the best of our knowledge, in the literature of chemical graph theory the applicability and chemical significance of Zagreb coindices have not been studied. Thus in our study we consider this for the Zagreb coindices and their predictive power are studied and compared with their respective Zagreb indices.
For any concepts and terms not defined here we recommend the reader to refer any standard books such as [24, 7].

2 Chemical significance of the Zagreb coindices.

Octane isomers are an important set of organic molecules to test the applicability of various topological parameters in quantitative structure activity/property relationships. These organic compounds are structurally diverse enough to yield considerable variation in shape, branching and non polarity. The number of their structural isomers is also large enough to make statistical analysis and the availability of experimental data for all the isomers makes octanes exceptionally convenient for such kind of testings. The octane isomers data set consists of 16 important physico-chemical properties: the boiling point (BP), melting point (MP), heat capacity at P constant (CP), heat capacity at T constant (CT), Entropy (S), density ($DENS$), standard enthalpy of vaporization ($DHVAP$), enthalpy of vaporization ($HVAP$), enthalpy of formation ($HFORM$), standard enthalpy of formation ($DHFORM$), motor octane number (MON), molar refraction (MR), acentric factor ($AcenFac$), total surface area (TSA), octanol-water partition coefficient ($LogP$), and molar volume (MV). The experimental data of these properties are found at http://www.moleculardescriptors.eu/dataset/dataset.htm. In a comprehensive study of numerous properties of octane isomers, Randić et al[21] have used single molecular descriptors and concluded that different physicochemical properties depend on different descriptors. Sofar in the literature of chemical graph theory, the chemical applicability of the Zagreb coindices have not been studied whether they can predict physicochemical properties of compounds or not.

Table 1: Topological Indices and some experimental values of octane isomers.

Alkane Isomers	$\overline{M}_1$	M_1	$\overline{M}_2$	M_2	S	$HVAP$	$DHVAP$	$AcentFac$	MON
n-octane	72	26	61	24	111.67	73.19	9.915	0.397898	$--$
2-methyl-heptane	70	28	58	26	109.84	70.30	9.484	0.377916	23.1
3-methyl-heptane	70	28	57	27	111.26	71.30	9.521	0.371002	35
4-methyl-heptane	70	28	57	27	109.32	70.91	9.483	0.371504	39
3-ethyl-hexane	70	28	56	28	109.43	71.70	9.476	0.362472	52.4
2,2-dimethyl-hexane	66	32	52	30	103.42	67.70	8.915	0.339426	77.4
2,3-dimethyl-hexane	68	30	53	30	108.02	70.20	9.272	0.348247	78.9
2,4-dimethyl-hexane	68	30	54	29	106.98	68.50	9.029	0.344223	69.9
2,5-dimethyl-hexane	68	30	55	28	105.72	68.60	9.051	0.35683	55.7
3,3-dimethyl-hexane	66	32	50	32	104.74	68.50	8.973	0.322596	83.4
3,4-dimethyl-hexane	68	30	52	31	106.59	70.20	9.316	0.340345	81.7
2-methyl-3-ethyl-pentane	68	30	52	31	106.06	69.70	9.209	0.332433	88.1
3-methyl-3-ethyl-pentane	66	32	48	34	101.48	69.30	9.081	0.306899	88.7
2,2,3-trimethyl-pentane	64	34	46	35	101.31	67.30	8.826	0.300816	99.9
2,2,4-trimethyl-pentane	64	34	49	32	104.09	64.87	8.402	0.30537	100
2,3,3-trimethyl-pentane	64	34	45	36	102.06	68.10	8.897	0.293177	99.4
2,3,4-trimethyl-pentane	66	32	49	33	102.39	68.37	9.014	0.317422	95.9
2,2,3,3-tetramethylbutane	60	38	39	40	93.06	66.20	8.410	0.255294	$--$

Thus by using the suggestion of International Academy of Mathematical Chemistry, the predictive ability of these descriptors is testified here for the first time. To carry out this, we first considered the experimental data of all the above 16 physico-chemical properties of octane isomers. We then calculate the Zagreb indices and coindices values of those 18 octane isomers. To calculate this we use graph structural analysis along with the definitions of M_1, M_2, $\overline{M}_1$ and $\overline{M}_2$ on the molecular graphs (carbon-atom skeleton) of octane isomers. The calculated Zagreb indices and coindices values as well as some important physcio-chemical properties of those octanes are presented in Table 1. Having those calculated and experimental data, curvilinear fittings were testified in all possible cases. However non of the cases could be established with such correlation. For this reason we used linear correlation coefficients as one means of comparisons. The calculated quantities are then correlated with each of these physico-chemical properties and regression analysis is carried out by standard statistical software package. In the following two sections we will study how well the Zagreb

coindices are correlated with physico-chemical properties of octane isomers and results will also be compared with those obtained using the well known Zagreb indices.

3 Main results

3.1 Comparative testings between the first Zagreb index and coindex

To testify the chemical applicability of the first Zagreb coindex five physico-chemical properties have been modeled: Entropy (S), Enthalpy of vaporization ($HVAP$), Standard enthalpy of vaporization ($DHVAP$) ,Acentric factor ($AcenFac$) and Motor octane number (MON) and results are compared with those obtained using the first Zagreb index. From a practical point of view, topological indices for which the absolute value of the correlation coefficients is less than 0.8 can be characterized as useless [9]. Thus, the reason behind choosing only these five physico-chemical properties is that both the first Zagreb index and coindex gives relatively good linear correlation i.e (the respective absolute value of their correlation coefficients are greater than 0.8) with only the four mentioned physico-chemical properties.
Scatter plots that shows correlations between M_1 and $\overline{M}_1$ with each of the four physico-chemical properties were used as means of simple comparisons. But it was not easy to identify which descriptor gives better results. Hence, we conducted statistical analysis to compare them and results are presented in Table 2.

Table 2: The correlation coefficients for the first Zagreb coindex and the first Zagreb index with some physco-chemical properties of octane isomers. RQR_1 is the ratio of quadratic mean of residuals of $\overline{M}_1$ to M_1.

Property	$\overline{M}_1$	M_1	$1 - RQR_1(\%)$
Entropy (S)	0.954	−0.954	0.000
Enthalpy of vaporization (HVAP)	0.886	−0.886	0.000
Standard enthalpy of vaporization (DHVAP)	0.936	−0.936	0.000
Acentric factor (AcenFac)	0.973	−0.973	0.000
Motor octane number MON (n=16)	−0.881	0.881	0.000

As we can see in the correlation out put data in Table 2, in all the properties both the first Zagreb index and its coindex gives the same correlation coefficients except that $\overline{M}_1$ is positively correlated and M_1 is negatively correlated with the properties. This unexpected result leads us to check the bivariate correlation between the descriptors. We found that there is perfectly negative relationship between $\overline{M}_1$ and M_1 i.e the correlation coefficient between them is -1. To check whether both descriptors really have same predictive ability we further check by using the ratio of quadratic means of residuals (RQR_1) :
Where:

$$RQR_1 = \sqrt{\frac{\frac{\sum_{i=1}^{n}[\acute{a}\cdot\overline{M}_{1i}+\acute{b}-Exp_i]^2}{n}}{\frac{\sum_{i=1}^{n}[a\cdot M_{1i}+b-Exp_i]^2}{n}}} = \sqrt{\frac{\sum_{i=1}^{n}[\acute{a}\cdot\overline{M}_{1i}+\acute{b}-Exp_i]^2}{\sum_{i=1}^{n}[a\cdot\ M_{1i}+b-Exp_i]^2}}$$

From the last column in Table 2, one can see that $1 - RQR_1 = 0.000\%$ for all the properties. This clearly guarantee that both $\overline{M}_1$ and M_1 have the same predicting power of physicochemical properties. To the best of our knowledge this fact was never mentioned in past studies[11, 9].
Note that we have checked and compared the correlation coefficients of all the 16 properties of octanes with both descriptors. We have found that in all the properties both have the same correlation coefficients (though it is weak) except that in all but in melting point, heat capacity at P constant, density and standard enthalpy of formation, $\overline{M}_1$ is positively correlated and M_1 is negatively correlated with the properties.

3.2 Comparative testings between the second Zagreb index and coindex

To testify the chemical applicability of the second Zagreb coindex four physico-chemical properties have been modeled: Entropy (S), Standard enthalpy of vaporization ($DHVAP$), Acentric factor ($AcenFac$), and

Motor octane number (MON) and results are compared with those obtained using the second Zagreb index. As in section 2.1, the reason behind choosing only these three physico-chemical properties is that both the second Zagreb index and coindex gives relatively good linear correlation i.e (the respective absolute value of correlation coefficients are greater than 0.8) with the mentioned three physico-chemical properties. The respective correlation coefficients are reported in Table 3. Looking at a glance on the correlation coefficients

Table 3: The correlation coefficients and ratio of quadratic mean of residuals of the second Zagreb coindex ($\overline{M}_2$), and the second Zagreb index (M_2) with some physco-chemical properties of octane isomers.

Property	$\overline{M}_2$	M_2	$1 - RQR_2(\%)$
Entropy (S)	0.955	−0.942	11.886
Standard enthalpy of vaporization (DHVAP)	0.854	−0.812	10.987
Acentric factor (AcenFac)	0.993	−0.986	25.696
Motor octane number MON (n=16)	−0.926	0.916	5.909

(in Table 3) it does not seem both have significant difference even though $\overline{M}_2$ shows better correlation coefficients than M_2 in all the properties. Thus as a means of comparing the predicting power of the second Zagreb coindex ($\overline{M}_2$) over the second Zagreb index (M_2), we use the ratio of quadratic mean of residuals RQR_2 :
where,

$$RQR_2 = \sqrt{\frac{\sum_{i=1}^{n}[\acute{a} \cdot \overline{M}_{2i} + \acute{b} - Exp_i]^2}{\sum_{i=1}^{n}[a \cdot M_{2i} + b - Exp_i]^2}}$$

The last column of Table 3 elaborates this. Accordingly, in all of the properties the prediction power of the second Zagreb coindex is at least 10% better than the second Zagreb index. In case of Entropy (resp. Standard enthalpy of vaporization) the predicting power of the second Zagreb coindex is at least 10% (resp. 11%) better than the predicting power of the second Zagreb index. The greatest improvement in prediction power of the second Zagreb coindex comparing to the second Zagreb index is obtained in the case of Acentric Factor which is more than 25%. Lastly the correlation between the descriptors was checked and it gives strongly negative relationship with $R = -0.996$. Thus from the above statistical analysis the second Zageb coindex should also be considered in the designining of quantitative structure-property relationships.

4 Conclusion

By using experimental data from standard data bases and computed data, the chemical applicability and predictive powers of the Zagreb coindices are studied for the first time and comparisons are held with their respective indices. Some of these results were surprising. For example, it has been found that the first Zagreb index and its coindex have the same predicting power for all 16 physico-chemical properties of octanes. Where as the second Zagreb coindex has better predicting power than second Zagreb index. Therefore, for researchers dealing with applications of chemical graph theory we recommend them to treat the first Zagreb index and the first Zagreb coindex equally. In general Zagreb coindices predict certain physico-chemical properties such as the Entropy and Acentric Factor with high accuracy (the correlation coefficient is greater than 0.954.) even better than Randić index.

Acknowledgment

The first author would like to thank to all the staff of the department of mathematics at Addis Ababa Science and Technology University for their love and encouragement.

References

[1] A. R. Ashrafi, T. Došlić, and A. Hamzeh. The zagreb coindices of graph operations. *Discrete applied mathematics*, 158(15):1571–1578, 2010.

[2] A. R. Ashrafi, T. Došlic, and A. Hamzeh. Extremal graphs with respect to the zagreb coindices. *MATCH Commun. Math. Comput. Chem*, 65(1):85–92, 2011.

[3] A. T. Balaban, I. Motoc, D. Bonchev, and O. Mekenyan. Topological indices for structure-activity correlations. In *Steric effects in drug design*, pages 21–55. Springer, 1983.

[4] K. Das, I. Gutman, and B. Horoldagva. Comparison between zagreb indices and zagreb coindices. *MATCH Commun. Math. Comput. Chem*, 68(1):189–198, 2012.

[5] K. C. Das and I. Gutman. Some properties of the second zagreb index. *MATCH Commun. Math. Comput. Chem*, 52(1):3–1, 2004.

[6] M. V. Diudea. *QSPR/QSAR studies by molecular descriptors*. Nova Science Publishers, 2001.

[7] M. V. Diudea, I. Gutman, and L. Jantschi. *Molecular topology*. Nova Science Publishers Huntington, NY, 2001.

[8] T. Došlić. Vertex-weighted wiener polynomials for composite graphs. *Ars Mathematica Contemporanea*, 1(1):66–80, 2008.

[9] I. Gutman. Degree-based topological indices. *Croatica Chemica Acta*, 86(4):351–361, 2013.

[10] I. Gutman, B. Furtula, Z. K. Vukicevic, and G. Popivoda. On zagreb indices and coindices. *MATCH Commun. Math. Comput. Chem*, 74(1):5–16, 2015.

[11] I. Gutman and J. Tošović. Testing the quality of molecular structure descriptors. vertex–degree-based topological indices. *Journal of the Serbian Chemical Society*, 78(6):805, 2013.

[12] I. Gutman and N. Trinajstić. Graph theory and molecular orbitals. total π -electron energy of alternant hydrocarbons. *Chem. Phys. Lett.*, 17:535–538, 1972.

[13] H. Hua, A. R. Ashrafi, and L. Zhang. More on zagreb coindices of graphs. *Filomat*, 26(6):1215–1225, 2012.

[14] H. Hua and S. Zhang. Relations between zagreb coindices and some distance-based topological indices. *Match-Communications in Mathematical and Computer Chemistry*, 68(1):199, 2012.

[15] X. Li and I. Gutman. Mathematical aspects of randic–type molecular structure descriptors, univ. *Kragujevac, Kragujevac*, 2006.

[16] X. Li and Y. Shi. A survey on the randic index. *MATCH Commun. Math. Comput. Chem*, 59(1):127–156, 2008.

[17] B. Liu and I. Gutman. On a conjecture on randic indices. *Match*, 62(1):143, 2009.

[18] S. Nikolić, G. Kovačević, A. Miličević, and N. Trinajstić. The zagreb indices 30 years after. *Croatica chemica acta*, 76(2):113–124, 2003.

[19] S. Nikolić, I. M. Tolić, N. Trinajstić, and I. Baučić. On the zagreb indices as complexity indices. *Croatica chemica acta*, 73(4):909–921, 2000.

[20] M. Randic. Characterization of molecular branching. *Journal of the American Chemical Society*, 97(23):6609–6615, 1975.

[21] M. Randic. Comparative structure-property studies: Regressions using a single descriptor. *Croat. Chem. Acta*, 66:289–312, 1993.

[22] J. Rodríguez and J. Sigarreta. On the randi\'{c} index and conditional parameters of a graph. *arXiv preprint arXiv:1311.7316*, 2013.

[23] R. Todeschini and V. Consonni. *Handbook of molecular descriptors*, volume 11. John Wiley & Sons, 2008.

[24] N. Trinajstic. *Chemical graph theory*. Routledge, 2018.

[25] B. Zhou and I. Gutman. Further properties of zagreb indices. *MATCH Commun. Math. Comput. Chem*, 54:233–239, 2005.

Email: melaku.berhe@aastu.edu.et
Email: wcxiang@mail.ccnu.edu.cn