

Scientific paper

# On Relationships of Eigenvalue–Based Topological Molecular Descriptors

Izudin Redžepović and Boris Furtula\*

Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia

\* Corresponding author: E-mail: boris.furtula@pmf.kg.ac.rs

Received: 08-23-2019

## Abstract

Three eigenvalue-based topological molecular descriptors are compared using several datasets of alkanes. Two of them are well-known and frequently employed in various QSPR/QSAR investigations, and third-one is a newly derived whose predictive potential is yet to be proven. The relations among them are found and discussed. Structural parameters that govern these relations are identified and the corresponding formulas based on multiple linear regression have been obtained. It has been shown that all three investigated indices are encoding almost the same structural information of a molecule. They differ only by the extent of the sensitivity on a structural branching of a molecule and on the number of non-bonding molecular orbitals.

**Keywords:** Graph energy, Estrada index, Resolvent energy of a graph, eigenvalues, adjacency matrix.

## 1. Introduction

Molecular descriptors are the fundamental tools in QSPR/QSAR modeling, which are frequently employed in diverse fields of chemistry.<sup>1–3</sup> Among them, topological indices are the usual choice, because of their low computational complexity and fairly simple identification of structure–property relationships.<sup>4–7</sup> There are hundreds of topological descriptors.<sup>2</sup> A natural way for their classification is by the origin of parameters that are used in their definitions. Thus, one differentiates degree–, distance–, and eigenvalue–based topological molecular descriptors, although there is a couple of them that cannot be strictly designated as members of any of the above-mentioned classes.

Interest for the eigenvalue-based topological molecular descriptors had been aroused after the explanation of the physical meaning of eigenvalues in HMO theory.<sup>8</sup> This happened in the seventies of the last century. Probably the first eigenvalue-based topological descriptor that had been introduced is the graph energy. This index is defined using the eigenvalues of an “ordinary” adjacency matrix in the following way:

$$E(G) = \sum_{i=1}^n |\lambda_i| \quad (1)$$

where  $\lambda_i$  is the  $i$ -th eigenvalue of a graph  $G$ .

The graph energy is tightly connected to the total  $\pi$ -electron energy of alternate conjugated molecules. It is a popular research topic both in chemically, and in mathematically oriented investigations. Several books and numerous papers are devoted to this particular topological invariant.<sup>9</sup> Nowadays, there are numerous eigenvalue-based topological indices, but just a couple of them are based on the eigenvalues that come from the adjacency matrix. These indices have been used as molds for defining almost all other topological invariants belonging to this class. Thus, beside graph energy, one could find indices like Laplacian energy, distance energy, Randić energy, etc. (e.g. see<sup>10–12</sup>).

Next to graph energy, the second most investigated topological molecular descriptor based on eigenvalues of an adjacency matrix is Estrada index. It was designed to model the folding in some biomolecules.<sup>13</sup> Estrada index is defined as follows:

$$EE(G) = \sum_{i=1}^n e^{\lambda_i} \quad (2)$$

where  $\lambda_i$  is the  $i$ -th eigenvalue of a graph  $G$ .

Its undeniable success led to a vigorous research of this quantity (see<sup>14–16</sup> and references cited therein). This invoked the introduction of many other Estrada-like invariants.<sup>10,17–19</sup>

Recently, another topological invariant based on “ordinary” eigenvalues has been introduced.<sup>20</sup> This descriptor is named resolvent energy after the resolvent matrix which eigenvalues are used in its definition.

$$ER(G) = \sum_{i=1}^n \frac{1}{n - \lambda_i} \quad (3)$$

where  $\frac{1}{n - \lambda_i}$  is the  $i$ -th eigenvalue of the resolvent matrix of the adjacency matrix of a graph.

Despite its juvenility, this quantity attracted much attention, which led to a couple of its descendants. A question that arises here is, whether there are some relationships among these topological invariants. Since the beginnings of the Estrada index, a connection between it and the energy of a graph has been investigating. Many inequalities, connecting these two descriptors, have been derived. However, the correlation between these two topological molecular descriptors seems to be never investigated. Also, the connections between the resolvent energy and the other two indices have not been tested yet.

This paper is devoted to the relationships among the recently introduced resolvent energy and other two eigenvalue-based topological descriptors. These relationships will shed a light onto relation between graph energy and the Estrada index as well.

## 2. Results

The results are separated into three parts. In the first subsection, the relation between graph energy and the resolvent energy of a graph will be elaborated. The second subsection is devoted to the relation between the Estrada index and the resolvent energy of a graph, and the last one is reserved for the relation between graph energy and the Estrada index. All relations are investigated in the case of alkanes.

### 2. 1. Graph Energy Versus Resolvent Energy

The resolvent energy of a graph has emerged in 2016 as a modification of the resolvent Estrada index, which was introduced few years before.<sup>17</sup> Several papers have appeared dealing with the mathematical properties of this quantity.<sup>21–29</sup>

Due to similar definitions of the resolvent energy (3) and the graph energy (1), the question about their relation naturally occurs. Fig. 1 shows the relation between the resolvent energy and the graph energy.

It is obvious from the Fig. 1 that the values of the graph energy are clustered into three distinct groups. Also, the values of the resolvent energy of graphs lies onto several nearly parallel lines.

We determined, by direct checking, that the values of the energy of acyclic connected graphs are classified into three distinct groups by the number of zeros ( $n_0(T)$ ) in their spectra (number of non-bonding orbitals in a molecule). On the other hand, the values of resolvent energy are separated onto nearly parallel lines by the values of the first Zagreb index ( $Zg_1(T)$ ) (a rough measure of a structural branching in a molecule). The alkanes that are lying on the same line have the same  $Zg_1(T)$ . Therefore, the correlation between the graph energy and the resolvent energy of a graph should involve these two parameters as well.

$$ER(T) \approx A \cdot E(T) + B \cdot Zg_1(T) + C \cdot n_0(T) + D \quad (4)$$

We made an in-house Python program for testing multiple linear relation shown in (4) using scikit-learn module.<sup>30</sup> Results are given in the Fig. 2 and Table 1.

The data presented in the Table 1, as well as the example shown in Fig. 2, demonstrate the remarkably good correlation between the values of the resolvent energy and the values obtained by the model given in (4). The first Zagreb index and the number of zeros in the spectra almost completely explain the dependence between the energy of graph and the resolvent energy in the case of trees.

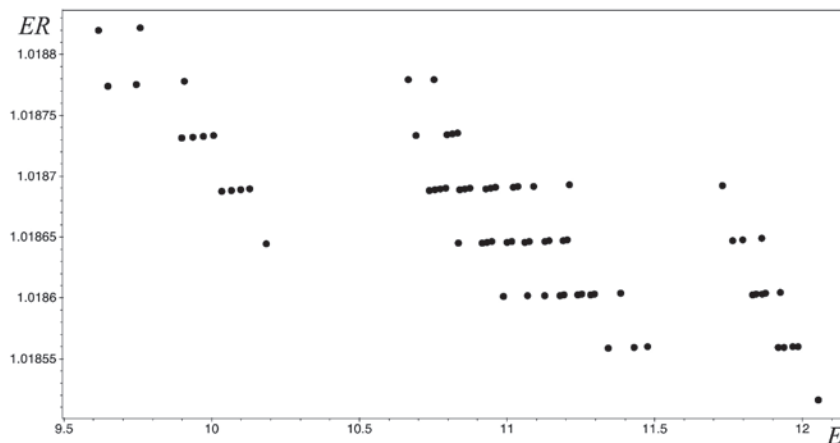


Figure 1. The relation between the resolvent energy of graph (ER) and the graph energy (E) in the case of 75 decanes.

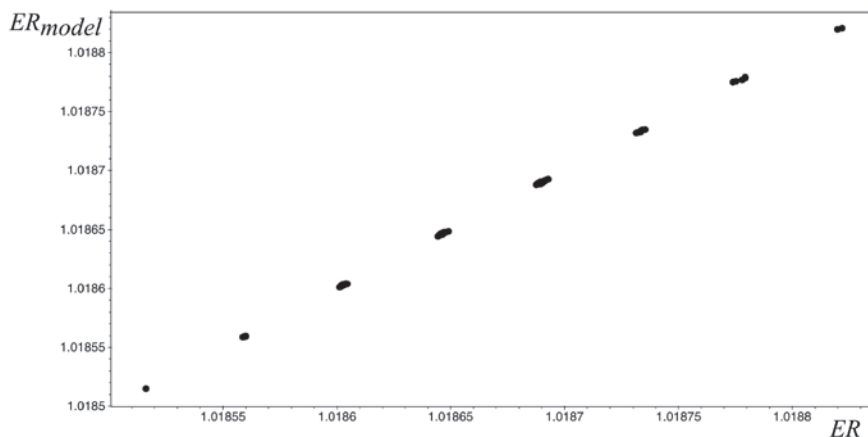


Figure 2. The correlation between the ER-values of chemical trees with 10 vertices and the values of ER calculated using eq. (4).

Table 1. Coefficients A, B, C, and D, in (4), computed to achieve the best correlation coefficient for chemical trees from 6 to 20 vertices. Last two columns contain obtained correlation coefficients and the average relative errors for all data sets used.

n	A	B	C	D	R	ARE
6	-2.21E-05	3.19E-04	-2.13E-05	1.044	1.00000	8.43E-08
7	6.90E-05	1.47E-04	2.11E-05	1.033	0.99999	1.15E-06
8	2.80E-05	7.24E-05	7.48E-06	1.026	0.99996	1.29E-06
9	1.55E-05	3.91E-05	4.56E-06	1.021	0.99998	6.39E-07
10	7.46E-06	2.25E-05	2.05E-06	1.018	0.99997	4.12E-07
11	4.01E-06	1.37E-05	1.15E-06	1.015	0.99998	2.29E-07
12	2.12E-06	8.74E-06	5.71E-07	1.013	0.99998	1.42E-07
13	1.25E-06	5.79E-06	3.48E-07	1.011	0.99998	8.71E-08
14	7.29E-07	3.96E-06	1.94E-07	1.009	0.99999	5.62E-08
15	4.61E-07	2.78E-06	1.25E-07	1.008	0.99999	3.63E-08
16	2.89E-07	2.00E-06	7.65E-08	1.007	0.99999	2.45E-08
17	1.92E-07	1.47E-06	5.11E-08	1.006	0.99999	1.67E-08
18	1.28E-07	1.10E-06	3.35E-08	1.006	0.99999	1.17E-08
19	8.83E-08	8.36E-07	2.32E-08	1.005	1.00000	8.35E-09
20	6.14E-08	6.45E-07	1.60E-08	1.005	1.00000	6.07E-09

## 2. 2. Estrada Index Versus Resolvent Energy

An illustrative example of correlation between the Estrada index and the resolvent energy of a graph is shown in Fig. 3.

Although the correlation in this example is quite well, it is evident that the points in Fig. 3 are clustered into several nearly parallel lines. It was empirically determined

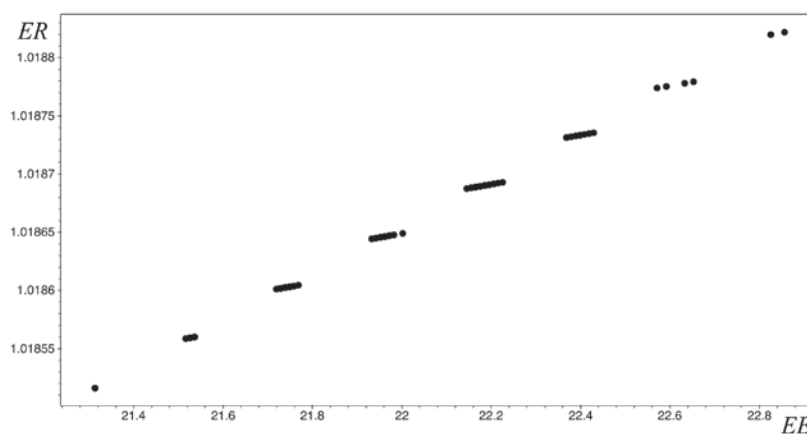


Figure 3. Correlation between the Estrada index and the resolvent energy of graph for all chemical trees with 10 vertices.

that the first Zagreb index is a parameter which governs this classification. Namely, the chemical trees belonging to a cluster have the same first Zagreb index. The relevance of this parameter for the correlation between the resolvent energy of a graph and the Estrada index is mathematically corroborated below. In order to do this, some well-known facts from spectral graph theory need to be outlined. The  $k$ -th spectral moment of a graph  $G$  is defined in the following manner:<sup>31</sup>

$$M_k(G) = \sum_{i=1}^n \lambda_i^k \quad (5)$$

where  $k$  is an integer greater than or equal to 0, and  $\lambda_i$  is the  $i$ -th eigenvalue of a graph  $G$ . The Estrada index and resolvent energy can be expressed in terms of the spectral moments using Taylor series (e.g.<sup>15,20</sup>):

$$EE(G) = \sum_{k \geq 0} \frac{M_k(G)}{k!} \quad (6)$$

$$ER(G) = \sum_{k \geq 0} \frac{M_k(G)}{n^{k+1}} \quad (7)$$

Then, using formulas (6) and (7) the following equality can be established:

$$ER(G) = \frac{EE(G)}{n} + \sum_{k \geq 0} \frac{k! - n^k}{k! \cdot n^{k+1}} M_k(G) \quad (8)$$

The odd  $k$ -th spectral moments are equal to 0 in the case of bipartite graphs (the chemical trees are bipartite). Using (8), the  $ER(T)$  can be approximated in terms of  $EE(T)$  and a few of the first spectral moments:

$$ER(T) \approx \frac{\alpha}{n} EE(T) + \frac{1-\alpha}{n} M_0(T) + \frac{2-\alpha \cdot n^2}{2n^3} M_2(T) + \frac{24-\alpha \cdot n^4}{24n^5} M_4(T) \quad (9)$$

where  $\alpha$  is a fitting parameter.

Since the  $k$ -th spectral moment of a graph is equal to the number of self-returning walks of the length  $k$ ,<sup>31</sup> the relations between spectral moments for  $k = 0, 2$ , and  $4$  and some other, easily divisible graph invariants, are derived years ago (e.g. see<sup>32</sup>):

$$M_0(T) = n$$

$$M_2(T) = 2(n-1) \quad (10)$$

$$M_4(T) = 2(Zg_1(T) - n + 1)$$

where  $n$ , and  $Zg_1(T)$  are the number of vertices and the first Zagreb index of a tree  $T$ .

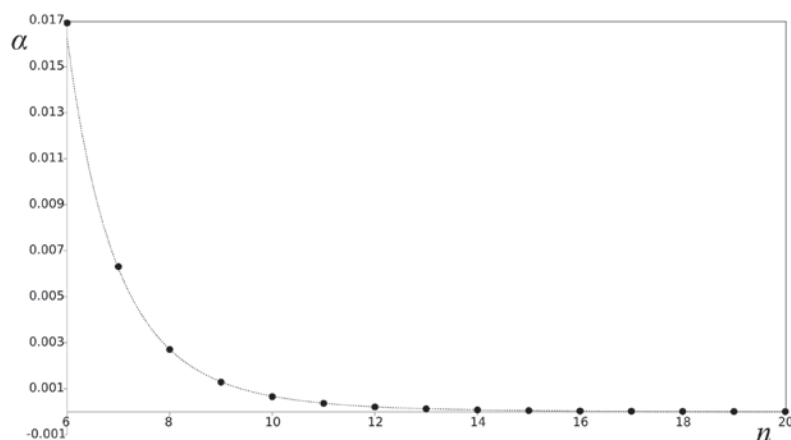
Incorporating equations shown in (10) into (9), the formula, relating the resolvent energy of a graph, Estrada index and the first Zagreb index, is obtained:

$$ER(T) \approx \frac{\alpha}{n} EE(T) - \frac{\alpha \cdot n^4 - 24}{12n^5} Zg_1(T) - \frac{11 \cdot \alpha \cdot n^4 - 24n^2 + 24}{12n^5} (n-1) - \alpha + 1 \quad (11)$$

In order to get the fitting parameter  $\alpha$  that appears in (11), we made an in-house computer program. This program is written in Python and the  $\alpha$  values are obtained for all chemical trees from 6 to 20 vertices. The results are shown in Table 2. The correlation coefficients are so high, and they are equal to 1 rounded to 7 decimals.

**Table 2.** The values of the fitting parameter  $\alpha$  from formula (11) for which the best correlation coefficients are obtained.

$n$	$\alpha$	$n$	$\alpha$
6	0.0169	13	0.00013
7	0.00631	14	8.20E-05
8	0.00271	15	5.40E-05
9	0.00129	16	3.60E-05
10	0.00066	17	2.60E-05
11	0.00037	18	1.80E-05
12	0.00021	19	1.30E-05
		20	9.00E-06



**Figure 4.** The relation between the fitting parameter  $\alpha$  and the number of vertices in chemical trees.

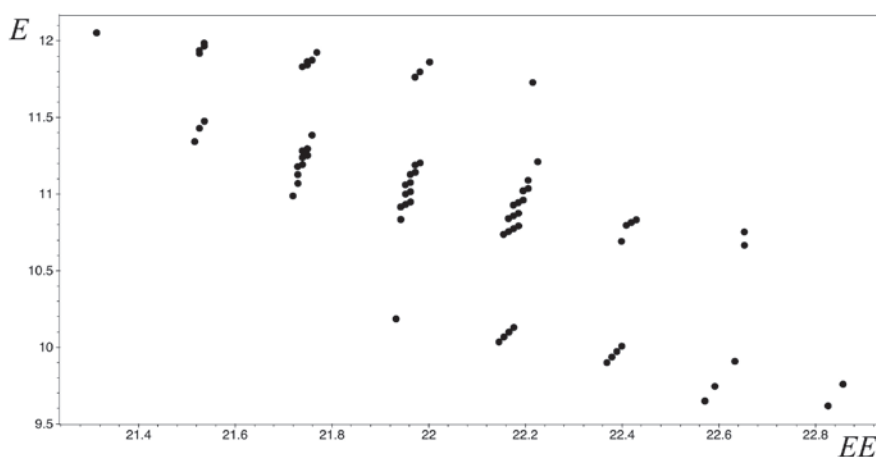


Figure 5. The relation between the Estrada index and the energy of a graph in the case of decanes.

The relation between the  $\alpha$  and the number of vertices is given in the Fig. 4. Although this relation is rather complex, the value of  $\alpha$  is completely determined by the  $n$ . Such finding suggests that the parameter, which solely influences the relation between the resolvent energy and Estrada index of isomeric chemical trees, is the first Zagreb index.

### 2. 3. Energy of Graph Versus Estrada Index

An introduction of the Estrada index also initiated the investigations of its connection with the energy of a graph. There are several papers presenting various bounds for the Estrada index in terms of the energy of a graph.<sup>14–16</sup>

However, the relation between these two indices is complex and has never been investigated thoroughly. Fig. 5 shows an illustrative example of the relation between the Estrada index and the energy of a graph. The approximate relations shown in (4) and (11) suggest that the

first Zagreb index and the number of zeros in the spectra of a graph are the parameters who largely influence the relationship between the Estrada index and the graph energy in the case of trees. Thence, we conjectured that the energy of a graph could be modeled by the following formula:

$$E(T) \approx A \cdot EE(T) + B \cdot Zg_1(T) + C \cdot n_0(T) + D \quad (12)$$

We tested the conjecture given in (12) using an in-house built Python program and the results are summarized in the Table 3 and the Fig. 6.

The statistics given in the Table 3 indicate that the model (12) explains more than 98% of the data variations. ARE-values are also considerably small. However, it is evident from Figs. 5 and 6 that beside the first Zagreb index and the number of zeros in the spectra of a graph, some other parameter(s) has an influence on this relation.

Table 3. The coefficients A, B, C, and D, the correlation coefficients and the average relative errors for the model given in (12).

$n$	A	B	C	D	R	ARE
6	-215.6	22.85	-1.384	2224.4	0.99995	0.07%
7	12.08	-1.388	-0.314	-135.9	0.99936	0.23%
8	7.78	-0.933	-0.31	-96.5	0.99671	0.47%
9	7.48	-0.903	-0.318	-104.3	0.99746	0.36%
10	6.26	-0.771	-0.307	-94.9	0.99563	0.45%
11	5.72	-0.713	-0.322	-94.3	0.99543	0.40%
12	5.53	-0.695	-0.309	-99.1	0.99427	0.42%
13	5.21	-0.659	-0.319	-100.1	0.99423	0.41%
14	5.11	-0.65	-0.31	-105.6	0.99339	0.41%
15	4.93	-0.63	-0.315	-108.5	0.99337	0.39%
16	4.86	-0.623	-0.31	-114	0.99278	0.39%
17	4.74	-0.609	-0.312	-117.6	0.99270	0.38%
18	4.69	-0.604	-0.309	-122.9	0.99229	0.37%
19	4.6	-0.595	-0.31	-127	0.99219	0.36%
20	4.56	-0.59	-0.309	-132.2	0.99189	0.36%

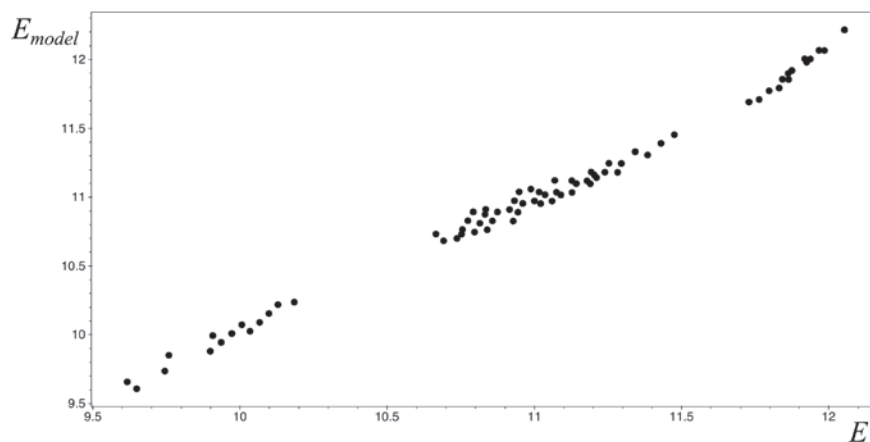


Figure 6. The model given in (12) versus the graph energy in the case of decanes.

### 3. Conclusion

The approximate relations among three eigenvalue-based topological indices whose definitions are based on the eigenvalues of the adjacency matrix are presented. It is shown that the first Zagreb index, as a measure of structural branching in a molecule, and the number of non-bonding orbitals, are the parameters that significantly influence these relations. In the (4) and (11) these graph invariants almost completely explain the relations between the  $ER(T)$  and  $E(T)$ , and  $ER(T)$  and  $EE(T)$ . The formulas (4) and (11) suggest that the relation between the  $E(T)$  and  $EE(T)$  can be modeled in terms of the first Zagreb index and the number of zeros in a graph. This model (12) has been tested and it is shown that it explains more than 98% of the data variation in the case of alkanes. However, for the complete description of a relation between the graph energy and the Estrada index, some other parameter(s), beside  $n_0(T)$  and  $Zg_1(T)$ , needs to be involved.

### Acknowledgments

This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through the grant no. 174033.

### 4. References

1. A. Dudek, T. Arodz, J. Galvez, *Comb. Chem. High Throughput Screen.* **2006**, *9*, 213–228. DOI:10.2174/138620706776055539
2. R. Todeschini, V. Consonni, *Molecular Descriptors for Chemoinformatics*, Wiley-VCH, Weinheim, Germany, **2009**. DOI:10.1002/9783527628766
3. A. Mauri, V. Consonni, R. Todeschini, in: J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, M. K. Shukla (Ed.): *Handbook of Computational Chemistry*, Springer, Cham, Switzerland, **2017**, pp. 2065–2093.
4. J. Devillers, A. T. Balaban (Ed.): *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, The Netherlands, **1999**.
5. A. Talevi, C. L. Bellera, M. Di Ianni, P. R. Duchowicz, L. E. Bruno-Blanch, E. A. Castro, *Curr. Comput. Aided Drug Des.* **2012**, *8*, 172–181. DOI:10.2174/157340912801619076
6. R. Zanni, M. Galvez-Llompart, R. García-Domenech, J. Galvez, *Expert Opin. Drug Discov.* **2015**, *10*, 945–957. DOI:10.1517/17460441.2015.1062751
7. J. C. Dearden, in: K. Roy (Ed.): *Advances in QSAR Modeling – Applications in Pharmaceutical, Chemical, Food, Agricultural and Environmental Sciences*, Springer, Cham, Switzerland, **2017**, pp 57–88. DOI:10.1007/978-3-319-56850-8\_2
8. D. Cvetković, I. Gutman, N. Trinajstić, *Croat. Chem. Acta* **1972**, *44*, 365–374.
9. X. Li, Y. Shi, I. Gutman, *Graph Energy*, Springer, New York, USA, **2012**.
10. G. P. Clemente, A. Cornaro, *MATCH Commun. Math. Comput. Chem.* **2017**, *77*, 673–690.
11. I. Gutman, B. Furtula, *Croat. Chem. Acta* **2017**, *90*, 359–368. DOI:10.5562/cca3189
12. N. J. Rad, A. Jahanbani, I. Gutman, *MATCH Commun. Math. Comput. Chem.* **2018**, *79*, 371–386.
13. E. Estrada, *Chem. Phys. Lett.* **2000**, *319*, 713–718. DOI:10.1016/S0009-2614(00)00158-5
14. A. Ilić, D. Stevanović, *J. Math. Chem.* **2009**, *47*, 305–314. DOI:10.1007/s10910-009-9570-0
15. B. Li, *MATCH Commun. Math. Comput. Chem.* **2017**, *77*, 701–706.
16. W. Wang, Y. Xue, *Appl. Math. Comput.* **2019**, *354*, 32–41. DOI:10.1016/j.amc.2019.02.019
17. X. Chen, J. Qian, *MATCH Commun. Math. Comput. Chem.* **2015**, *73*, 163–174.
18. J. Li, L. Qiao, N. Gao, *Appl. Math. Comput.* **2018**, *317*, 143–149. DOI:10.1016/j.amc.2017.09.015
19. Y. Shang, *Bull. Aust. Math. Soc.* **2013**, *88*, 106–112. DOI:10.1017/S0004972712000676
20. I. Gutman, B. Furtula, E. Zogić, E. Glogić, *MATCH Commun.*

- Math. Comput. Chem.* **2016**, *75*, 279–290.
21. L. E. Allem, J. Capaverde, V. Trevisan, I. Gutman, E. Zogić, E. Glogić, *MATCH Commun. Math. Comput. Chem.* **2017**, *77*, 95–104.
22. M. Bianchi, A. Cornaro, J. L. Palacios, A. Torriero, *MATCH Commun. Math. Comput. Chem.* **2018**, *80*, 459–465.
23. K. C. Das, *MATCH Commun. Math. Comput. Chem.* **2019**, *81*, 453–464.
24. Z. Du, *MATCH Commun. Math. Comput. Chem.* **2017**, *77*, 85–94.
25. A. Farrugia, *Appl. Math. Comput.* **2018**, *321*, 25–36.  
DOI:10.1016/j.amc.2017.10.020
26. M. Ghebleh, A. Kanso, D. Stevanović, *MATCH Commun. Math. Comput. Chem.* **2017**, *77*, 635–654.  
DOI:10.1007/s11042-017-4634-9
27. I. Gutman, B. Furtula, E. Zogić, E. Glogić, E. in: I. Gutman, X. Li (Ed.): *Energies of Graphs – Theory and Applications*, Univ. Kragujevac, Kragujevac, Serbia, **2016**, pp. 277–290.
28. Z. Zhu, *J. Math. Anal. Appl.* **2017**, *447*, 957–970.  
DOI:10.1016/j.jmaa.2016.10.043
29. E. H. Zogić, E. R. Glogić, *Sci. Publ. State Univ. Novi Pazar* **2017**, *9A*, 187–191. DOI:10.5937/SPSUNP1702187Z
30. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
31. D. Cvetković, M. Doob, H. Sachs, *H. Spectra of Graphs – Theory and Application*, Johann Ambrosius Barth Verlag: Heidelberg, Leipzig, Germany, **1995**.
32. I. Gutman, K. C. Das, *MATCH Commun. Math. Comput. Chem.* **2004**, *50*, 83–92.

## Povzetek

Tri topološke molekularne deskriptorje smo primerjali s podatkovnimi bazami alkanov. Dva od teh deskriptorjev sta dobro znana in pogosto uporabljena v različnih QSPR/QSAR preiskavah, tretji pa je na novo izpeljan in je treba njegove napovedovalne možnosti še dokazati. Našli smo povezave med temi deskriptorji in o njih razpravljali. Z uporabo večkratne linearne regresije smo opredelili strukturne parametre in ustrezne enačbe, ki določajo te povezave. Pokazali smo, da vsi trije preiskovani indeksi kodirajo skoraj iste strukturne informacije o molekuli. Razlikujejo se le po obsegu občutljivosti na strukturno razvejanje molekule in po številu neveznih molekularnih orbital.



Except when otherwise noted, articles in this journal are published under the terms and conditions of the Creative Commons Attribution 4.0 International License