CROATICA CHEMICA ACTA
CCACAA **76** (2) 113–124 (2003)

ISSN-0011-1643
CCA-2861
Authors' Review

The Zagreb Indices 30 Years After

Sonja Nikolić,* Goran Kovačević, Ante Miličević, and Nenad Trinajstić

The Rugjer Bošković Institute, P. O. Box 180, HR-10002 Zagreb, Croatia

RECEIVED JANUARY 30, 2003; REVISED MAY 12, 2003; ACCEPTED MAY 13, 2003

The original formulation of the Zagreb indices is presented and their relationship to topological indices such as self-returning walks, Platt, Gordon-Scantlebury and connectivity indices is discussed. Their properties are also listed. Modified Zagreb indices are introduced and the Zagreb complexity indices reviewed. Their use in QSPR is illustrated by modeling the structure-boiling point relationship of C_3–C_8 alkanes using the CROMRsel procedure. The obtained models are in fair agreement with experimental data and are better than many models in the literature. However, in general, the Zagreb indices do not contribute to the best structure-boiling point models of alkanes. Nevertheless, it is interesting to note that the best five-descriptor model that we found in the literature contains the Zagreb M_2 index.

Key words
connectivity indices
molecular graphs
QSPR
Zagreb indices

INTRODUCTION

A pair of topological indices, denoted by symbols M_1 and M_2, was introduced 30 years ago.[1] They were given different names in the literature, such as the Zagreb Group indices,[2] the Zagreb group parameters[3] and, most often, the Zagreb indices.[4] In this review, we report what was happening with these two topological indices since their inception because when they appeared only the Wiener index[5] and the Hosoya *Z*-index[6] were known and used. Randić introduced in 1975 a bond-additive topological index as a descriptor for characterizing molecular branching,[7] which he called the branching index. This index was soon renamed the connectivity index[8] and generalized to connectivity indices of various orders, starting with the zeroth-order connectivity index.[9] The first-order connectivity index is grounded on the Zagreb M_2 index. The connectivity index and its variants are used more frequently than any other topological index in QSPR and QSAR.[2,4,10–22]

Recently, the Zagreb indices and their variants have been used to study molecular complexity,[23–27] chirality,[28] *ZE*-isomerism[29] and heterosystems[30] whilst the overall Zagreb indices[31] exhibited a potential applicability for deriving multilinear regression models. Zagreb indices are also used by various researchers in their QSPR and QSAR studies.[2,4,12,13,32–43] Mathematical properties of the Zagreb indices have also been studied.[44–46]

Zagreb indices are referred to in most books reporting topological indices and their uses in QSPR and QSAR.[2,4,12,13] They are also included in a number of programs used for the routine computation of topological indices, such as POLLY,[47] OASIS,[48] DRAGON,[49] Cerius,[50] TAM,[51] DISSIM,[52] *etc*.

Since many authors who use the Zagreb indices as well as other topological indices employ the concepts and terminology of chemical graph theory[53] in their research, we will do the same. Graphs are generated from molecules by replacing atoms with vertices and bonds with edges. We will consider only simple molecular graphs, that is, graphs without loops and multiple edges.[54] Be-

* Author to whom correspondence should be addressed. (E-mail: sonja@rudjer.irb.hr)

sides, graphs that we will use will represent only bare molecular skeletons, that is, molecular skeletons without hydrogen atoms. In Figure 1, we give as an illustrative example a tree T (tree is a connected acyclic graph)[54] corresponding to 2-methylbutane.

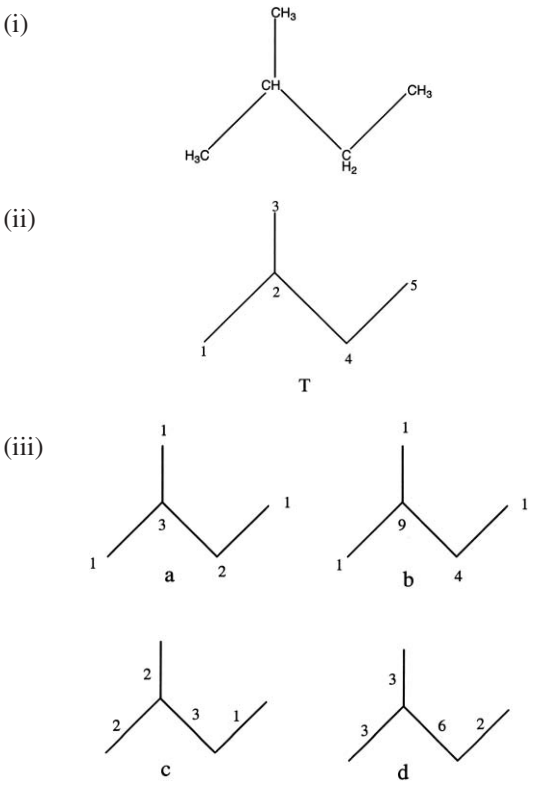

Figure 1. Labeled hydrogen-suppressed tree T representing 2-methylbutane and its vertex-degrees, squared vertex-degrees, edge-degrees and edge-weights. (i) 2-Methylbutane; (ii) Hydrogen suppressed labeled tree T corresponding to 2-methylbutane; (iii) Vertex-degrees (a), squared vertex-degrees (b), edge-degrees (c) and edge-weights (d) in T.

The development and uses of the Zagreb indices, including some new results, are reviewed below. We first give definitions of these indices, list some of their properties and discuss their connection with several topological indices. Then, we present the modified and Zagreb complexity indices. Finally, we describe their uses in QSPR modeling. We end the article with our concluding remarks.

DEFINITIONS OF ZAGREB INDICES AND THEIR PROPERTIES

Definitions

In the early work of the Zagreb Mathematical Chemistry Group on the topological basis of the π-electron energy, two terms appeared in the topological formula for the total π-energy of conjugated molecules,[1] which were first used as branching indices[55] and later as topological indices in QSPR and QSAR studies.[4]

The original Zagreb indices are defined as follows:

$$M_1 = \sum_{\text{vertices}} d(i)\, d(i) \tag{1}$$

$$M_2 = \sum_{\text{edges}} d(i)\, d(j) \tag{2}$$

where $d(i)$ is the degree of vertex i and $d(i)\ d(j)$ is the weight of edge i-j. For example, the values of M_1 and M_2 indices for the tree T representing 2-methylbutane are $M_1 = 1 + 9 + 1 + 4 + 1 = 16$ and $M_2 = 3 + 3 + 6 + 2 = 14$.

The total adjacency index A[56–59] may be considered as the precursor to the M_1 index:

$$A = \sum_{\text{vertices}} d(i)\,. \tag{3}$$

The total adjacency index, of course, *via* the handshaking lemma,[54] equals twice the number of edges E in a graph.

$$\sum_{i} d(i) = 2E\,. \tag{4}$$

The connectivity index, also called the vertex-connectivity index,[60,61] to differentiate it from the edge-connectivity index, denoted by χ, is given by:[7]

$$\chi = \sum_{\text{edges}} \left[d(i)\, d(j)\right]^{-0.5}\,. \tag{5}$$

This index is also called the first-order (vertex-)connectivity index and is sometimes denoted by ${}^1\chi$.[9] The zero-order (vertex-)connectivity index is given by:

$${}^0\chi = \sum_{\text{vertices}} \left[d(i)\right]^{-0.5}\,. \tag{6}$$

The Zagreb indices (and the connectivity index) can be also obtained directly from the adjacency matrix $\boldsymbol{A}$ of a graph since the row-sums of this matrix are equal to vertex-degrees. For example, the adjacency matrix corresponding to the labeled tree T, given in Figure 1, is:

					$d(i)$
0	1	0	0	0	1
1	0	1	1	0	3
0	1	0	0	0	1
0	1	0	0	1	2
0	0	0	1	0	1

Zagreb indices can be also given in terms of the diagonal elements of the squared adjacency matrix:

$$M_1 = \sum_{\text{vertices}} (\boldsymbol{A}^2)_{ii}\, (\boldsymbol{A}^2)_{ii} \tag{7}$$

$$M_2 = \sum_{\text{edges}} (\boldsymbol{A}^2)_{ii}\, (\boldsymbol{A}^2)_{jj}\,. \tag{8}$$

Eqs. (7) and (8) are directly connected with Eqs. (1) and (2) by the equality:

$$d(i) = (\boldsymbol{A}^2)_{ii} . \quad (9)$$

To illustrate this, we give below the squared adjacency matrix $\boldsymbol{A}^2$ of the labeled tree T, given in Figure 1:

$$\begin{matrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{matrix}$$

Zagreb indices and the zero- and first-order connectivity indices for the lowest alkanes are given in Table I.

Properties

Zagreb indices possess many interesting properties. Here we will describe a few of them. There is a relationship between M_1 and the number of vertices having degree 1 (corresponding to the number of primary carbon atoms, P), degree 2 (corresponding to the number of secondary carbon atoms, S), degree 3 (corresponding to the number of tertiary carbon atoms, T) or degree 4 (corresponding to the number of quaternary carbon atoms, Q) for trees representing alkanes:

$$M_1 = \mathrm{P} + 4\ \mathrm{S} + 9\ \mathrm{T} + 16\ \mathrm{Q} . \quad (10)$$

Eq. (10) transforms into:

$$M_1 = 4\ \mathrm{V} + 2\ \mathrm{T} + 6\ \mathrm{Q} - 6 \quad (11)$$

by means of the following equations:

$$E = V - 1 \quad (12)$$

$$V = \mathrm{P} + \mathrm{S} + \mathrm{T} + \mathrm{Q} \quad (13)$$

$$E = (\mathrm{P} + 2\ \mathrm{S} + 3\ \mathrm{T} + 4\ \mathrm{Q})/2 \quad (14)$$

where V stands for the number of vertices in a tree.

If we transform a graph G into the corresponding line graph L(G), then the M_1 index of graph G equals twice the number of vertices and edges in the corresponding line graph L(G),[62] as shown below. The line graph L(G) of a simple graph G is the graph derived from G in such a way that the edges in G are replaced by vertices in L(G).[54] Two vertices in L(G) are connected whenever the corresponding edges in G are adjacent.

The line graph L(T) corresponding to the tree T from Figure 1 is given in Figure 2.

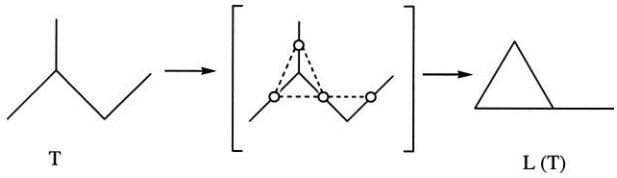

Figure 2. Construction of the line graph L(T) from tree T given in Figure 1.

The number of vertices and edges of graph G and the corresponding line graph L(G) are related by:

$$V[\mathrm{L(G)}] = E(\mathrm{G}) \quad (15)$$

$$E[\mathrm{L(G)}] = (1/2) \sum_{\text{vertices}} d(i)\, d(i)\ (\mathrm{G}) - \mathrm{E(G)} . \quad (16)$$

Using Eq. (1) we obtain:

$$E[\mathrm{L(G)}] = (1/2)\, M_1 - E(\mathrm{G}) . \quad (17)$$

Eq. (17) transforms into:

$$M_1 = 2\ \{\mathrm{E[L(G)]} + E(\mathrm{G})\} \quad (18)$$

and, utilizing Eq. (15), into:

$$M_1 = 2\ \{[V[\mathrm{L(G)}] + E[\mathrm{L(G)}]\} . \quad (19)$$

M_1 is also equal to the number of walks of length 2 ($\mathrm{mwc_2}$) in a graph:[63]

$$M_1 = \mathrm{mwc_2} \quad (20)$$

and is related to the number of self-returning walks of length 4 ($\mathrm{srw_4}$):[64,65]

$$M_1 = (\mathrm{srw_4} + 2\ E)/2 . \quad (21)$$

Since $\mathrm{mwc_2}$ is equal to the sum of all elements in the squared adjacency matrix, Eq. (20) can be given as:

$$M_1 = \sum_{i=1}^{V} \sum_{j=1}^{V} (\boldsymbol{A}^2)_{ij} . \quad (22)$$

Eq. (21) can also be transformed into:

$$M_1 = (\sum_{\text{vertices}} (\boldsymbol{A}^4)_{ii} + 2\ E)/2 \quad (23)$$

by means of:

$$\mathrm{srw_4} = \sum_{\text{vertices}} (\boldsymbol{A}^4)_{ii} . \quad (24)$$

While Eq. (20) is generally valid, Eq. (21) and therefore Eq. (23) are limited to graphs without 4-membered rings. Note that a walk in a graph is any sequence of consecutive edges. The length of the walk is the number

TABLE I. C_3–C_8 alkanes, their Zagreb indices (M_1, M_2), modified Zagreb indices ($^mM_1$, $^mM_2$), the number of paths of length 3 (p_3), the zero-order and the first-order connectivity indices ($^0\chi$, $^1\chi$), overall Zagreb complexity indices (TM_1, TM_1*, TM_2, TM_2*), total walk counts (*twc*) and their experimental boiling points (bp)

	Alkane	M_1	M_2	$^mM_1$	$^mM_2$	p_3	$^0\chi$	$^1\chi$	TM_1	TM_1*	TM_2	TM_2*	*twc*	bp
1	Propane	6	4	2.25	1	0	2.7071	1.4142	22	10	8	6	5	–42.1
2	Butane	10	8	2.5	1.25	1	3.4142	1.9142	56	28	28	19	16	–0.5
3	2-Methylpropane	12	9	3.1111	1	0	3.5774	1.7321	87	36	36	24	18	–11.7
4	Pentane	14	12	2.75	1.5	2	4.1213	2.4142	110	60	64	44	44	36.0
5	2-Methylbutane	16	14	3.3611	1.3332	2	4.2845	2.2701	168	80	94	59	53	27.8
6	2,2-Dimethylpropane	20	16	4.0625	1	0	4.5	2.0	292	112	128	80	70	9.5
7	Hexane	18	16	3	1.75	3	4.8284	2.9142	188	110	120	85	111	68.7
8	2-Methylpentane	20	18	3.6111	1.5832	3	4.9916	2.7701	277	146	172	114	134	60.3
9	3-Methylpentane	20	19	3.6111	1.6665	4	4.9916	2.8081	300	158	199	125	142	63.3
10	2,2-Dimethylbutane	24	22	4.3125	1.375	3	5.2071	2.5607	505	222	290	173	185	49.7
11	2,3-Dimethylbutane	22	21	4.2222	1.4444	4	5.1547	2.6427	404	196	264	156	165	58.0
12	Heptane	22	20	3.25	2	4	5.5355	3.4142	294	182	200	146	268	98.5
13	2-Methylhexane	24	22	3.8611	1.8332	4	5.6987	3.2701	418	238	278	193	329	90.0
14	3-Methylhexane	24	23	3.8611	1.9166	5	5.6987	3.3081	468	266	330	219	354	92.0
15	3-Ethylpentane	24	24	3.8611	2	6	5.6987	3.3461	516	294	384	246	378	93.5
16	2,2-Dimethylpentane	28	26	4.5625	1.625	4	5.9142	3.0607	762	370	476	302	489	79.2
17	2,3-Dimethylpentane	26	26	4.4722	1.7777	6	5.8618	3.1807	668	354	487	297	436	89.8
18	2,4-Dimethylpentane	26	24	4.4722	1.6666	4	5.8618	3.1259	584	312	384	256	399	80.5
19	3,3-Dimethylpentane	28	28	4.5625	1.75	6	5.9142	3.1213	850	414	580	344	526	86.1
20	2,2,3-Trimethylbutane	30	30	5.1736	1.4999	6	6.0774	2.9434	1065	494	732	414	588	80.9
21	Octane	26	24	3.5	2.25	5	6.2426	3.9142	432	280	308	231	627	125.7
22	2-Methylheptane	28	26	4.1111	2.0832	5	6.4058	3.7701	595	360	416	300	764	117.6
23	3-Methylheptane	28	27	4.1111	2.1666	6	6.4058	3.8081	676	408	497	345	838	118.9
24	4-Methylheptane	28	27	4.1111	2.1666	6	6.4058	3.8081	703	424	520	360	856	117.7
25	3-Ethylhexane	28	28	4.1111	2.25	7	6.4058	3.8461	780	472	604	407	928	118.5
26	2,2-Dimethylhexane	32	30	4.8125	1.875	5	6.6213	3.5607	1067	560	706	471	1142	106.8
27	2,3-Dimethylhexane	30	30	4.7222	2.0277	7	6.5689	3.6807	984	558	748	482	1068	115.6
28	2,4-Dimethylhexane	30	29	4.7222	2	6	6.5689	3.6639	923	528	668	451	997	109.4
29	2,5-Dimethylhexane	30	28	4.7222	1.9166	5	6.5689	3.6259	804	462	556	389	911	109.1
30	3,3-Dimethylhexane	32	32	4.8125	2	7	6.6213	3.6213	1255	660	902	567	1301	112.0
31	3,4-Dimethylhexane	30	31	4.7222	2.1111	8	6.5689	3.7187	1072	608	846	531	1136	117.7
32	3-Ethyl-2-methylpentane	30	31	4.7222	2.1111	8	6.5689	3.7187	1097	624	868	548	1152	115.6
33	3-Ethyl-3-methylpentane	32	34	4.8125	2.125	9	6.6213	3.6820	1408	744	1086	651	1441	118.2
34	2,2,3-Trimethylpentane	34	35	5.4236	1.8333	8	6.7845	3.4814	1669	840	1247	735	1536	109.8
35	2,2,4-Trimethylpentane	34	32	5.4236	1.7082	5	6.7845	3.4165	1425	724	944	616	1317	99.2
36	2,3,3-Trimethylpentane	34	36	5.4236	1.8749	9	6.7845	3.5040	1734	874	1334	769	1609	114.8
37	2,3,4-Trimethylpentane	32	33	5.3333	1.8888	8	6.7321	3.5534	1360	734	1059	645	1296	113.5
38	2,2,3,3-Tetramethylbutane	38	40	6.125	1.5625	9	7	3.2500	2526	1160	1888	1024	2047	106.5

of edges in it. A self-returning walk is a walk starting and ending at the same vertex.

The Zagreb M_1 index is related to the Gordon-Scantlebury index[66] and the Platt index (or the Platt number).[67] The Gordon-Scantlebury index S of a graph G is equal to the number of paths of length two p_2 in G:

$$S = p_2 . \tag{25}$$

A path in a graph G is a sequence of adjacent edges, which do not pass through the same vertex more than once and the length of the path is the number of edges in it.[68]

The Gordon-Scantlebury index can also be defined by means of the squared adjacency matrix of G:[69]

$$S = (1/2) \sum_i (\boldsymbol{A}^2)_{ii} (\boldsymbol{A}^2)_{ii} - E . \tag{26}$$

Using Eq. (9), the above expression converts into:

$$S = (1/2) \sum_i d(i) d(i) - E \tag{27}$$

or, using Eq. (1), to:

$$M_1 = 2 (S + E) . \tag{28}$$

Comparison between this equation and Eq. (17) tells us that the Gordon-Scantlebury index of G enumerates edges in the corresponding line graph.

The Platt index F of a graph G is equal to the total sum of edge-degrees $\varepsilon(i)$ in G:

$$F = \sum_{\text{edges}} \varepsilon(i) . \tag{29}$$

The degree $\varepsilon(i)$ of an edge i is equal to the number of its adjacent edges (see Figure 1). The Platt index can be also defined by means of the squared adjacency matrix:[69]

$$F = \sum_{\text{edges}} \left[(\boldsymbol{A}^2)_{ii} + (\boldsymbol{A}^2)_{jj} - 2 \right] . \tag{30}$$

Eq. (30) can be reformulated as:

$$F = \sum_i \left[(\boldsymbol{A}^2)_{ii} (\boldsymbol{A}^2)_{ii} - (\boldsymbol{A}^2)_{ii} \right] . \tag{31}$$

Using the handshaking lemma (4) and Eq. (9), Eq. (31) converts into:

$$F = \sum_i \left[d(i) d(i) \right] - 2E . \tag{32}$$

Combination of Eqs. (27) and (32) produces a simple relationship between F and S:

$$F = 2 S . \tag{33}$$

Similarly, the relationship between M_1 and F is also simple:

$$M_1 = F + 2 E . \tag{34}$$

Analytical Formulas for Some Homologous Structures

The M_1 and M_2 indices can be given in a closed form for homologous structures. Here we give analytical formulas for several classes of regular structures:

(i) n-alkanes (n stands for normal, that is, unbranched alkanes):

$$M_1 = 4 (V - 2) + 2 \tag{35}$$

$$M_2 = 4 (V - 2) \text{ for } V > 2 \tag{36}$$

(ii) V-cycloalkanes (V stands for the size of cycloalkanes in terms of the number of vertices):

$$M_1 = M_2 = 4 V . \tag{37}$$

(iii) polyacenes

$$M_1 = 26 R - 2 \tag{38}$$

$$M_2 = 33 R - 9 . \tag{39}$$

(iv) polyphenanthrenes (zig-zag benzenoids)

$$M_1 \text{ (polyphenanthrene)} = M_1 \text{ (polyacene)} \tag{40}$$

$$M_2 = 34 R - 11 \tag{41}$$

where R is the number of hexagons in a polyacene or polyphenanthrene.

The equality $M_1 = M_2$ for V-cycloalkanes is a consequence of the fact that the number of carbon atoms in these molecules is equal to the number of carbon-carbon bonds, that is, $V = E$. The equality M_1(polyphenanthrene) = M_1(polyacene) is a consequence of the definition of the M_1 index. Index M_1 depends only on the valencies of atoms (degrees) in a molecule (graph) and these are the same in isomeric benzenoids.

MODIFIED ZAGREB INDICES

A problem with the Zagreb indices, as well as with many other topological indices, but not, for example, with the vertex-connectivity indices,[70,71] is that their contributing parts give greater weights to inner (interior) vertices and edges and smaller weights to outer (terminal) vertices and edges of a graph, as can be seen from Figure 1 (diagrams (b) and (d)). This opposes intuitive reasoning that outer atoms and bonds should have greater weights than inner vertices and bonds because outer vertices and

bonds are associated with a larger part of the molecular surface and are consequently expected to make a greater contribution to physical, chemical and biological properties. Chemical intuition should not be disregarded even in theoretical research, as some tend to do,[72] because many crucial discoveries in chemistry, such as the periodic law and the benzene structural formula, were achieved relying on intuitive rules.[73,74] Researcher's intuition is a very important guidance in many areas of modern chemistry and especially in drug design.[75]

One way to amend the Zagreb indices is to insert inverse values of the vertex-degrees into Eqs. (1) and (2). We call these indices the modified Zagreb indices and denote them ${}^{m}M_1$ and ${}^{m}M_2$. They are given below:

$$ {}^{m}M_1 = \sum_{\text{vertices}} [1/d(i)][1/d(i)] = \sum_{\text{vertices}} [d(i)d(i)]^{-1} \quad (42) $$

$$ {}^{m}M_2 = \sum_{\text{edges}} 1/[d(i)d(j)] = \sum_{\text{edges}} [d(i)d(j)]^{-1} . \quad (43) $$

It should be pointed out that ${}^{m}M_2$ is identical to the first-order overall index, denoted by ${}^{1}ON$ and introduced by Bonchev.[58]

A modified total adjacency index is:

$$ {}^{m}A = \sum_{\text{vertices}} 1/d(i) = \sum_{\text{vertices}} [d(i)]^{-1} . \quad (44) $$

The values of ${}^{m}A$, ${}^{m}M_1$, ${}^{m}M_2$, ${}^{0}\chi$ and ${}^{1}\chi$ indices for the 2-methylbutane graph are given in Figure 3.

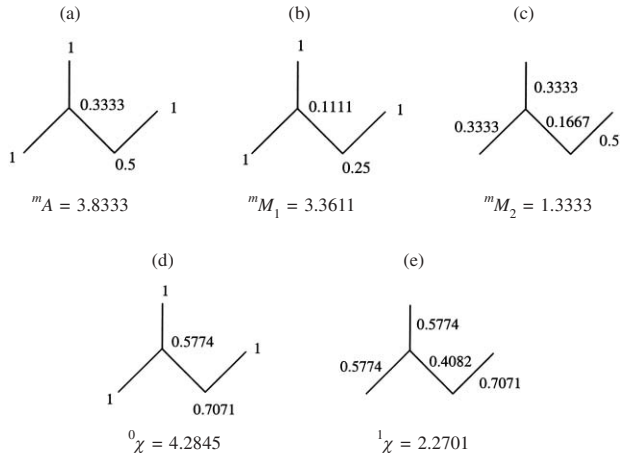

Figure 3. Vertex-weights and edge-weights contributing to (a) modified total adjacency index ${}^{m}A$, (b) modified Zagreb indices ${}^{m}M_1$, (c) ${}^{m}M_2$ and to the first two vertex-connectivity indices (d) ${}^{0}\chi$ and (e) ${}^{1}\chi$.

All three modified indices follow the zero-order and first-order vertex-connectivity indices by giving greater weights to outer vertices and edges than to inner vertices and edges in (molecular) graphs.

The modified Zagreb indices ${}^{m}M_1$ and ${}^{m}M_2$ for the lowest alkanes are also given in Table I. The reader should note that the values of ${}^{m}M_2$ for C_3–C_8 alkanes given in Table I are identical with those for ${}^{1}ON$, given in Table III of Ref. 58.

Formal similarity between formulas (5) and (6), and (43) and (44) is obvious. Hence, one may say that the road to the connectivity indices consisted of the following stops: $A \rightarrow {}^{m}A \rightarrow {}^{0}\chi$ and $M_2 \rightarrow {}^{m}M_2 \rightarrow {}^{1}\chi$. However, this road was not taken — the vertex-connectivity indices ${}^{0}\chi$ and ${}^{1}\chi$ were obtained in quite a different manner than indices ${}^{m}A$ and ${}^{m}M_2$.

ZAGREB COMPLEXITY INDICES

The Zagreb complexity indices, denoted by TM_1, TM_1*, TM_2 and TM_2*, are defined as:[27]

$$ TM_1 = \sum_{\mathbf{S}} \sum_{\text{vertices}} d_i^{\,2}(\mathbf{S}) \quad (45) $$

$$ TM_2 = \sum_{\mathbf{S}} \sum_{\text{edges}} d_i d_j (\mathbf{S}) . \quad (46) $$

The first sum in the above formulas represents the sum of the M_1 (M_2) indices for all connected subgraphs **S** of a graph G. It should be noted that in TM_1, the M_1 index is calculated for each subgraph **S** using the vertex-degrees as they are in graph G. If the M_1 index is calculated for each subgraph **S** using the vertex-degrees as they are in isolated subgraphs, then the corresponding index is TM_1*. In the case of the TM_2 index, the M_2 index is calculated for each subgraph **S** using the edge-weights as they are in graph G. If the M_2 index is calculated for each subgraph **S** using the edge-weights as they are in isolated subgraphs, then the corresponding index is TM_2*. Indices TM_1, TM_1*, TM_2 and TM_2* were derived by combining the ideas of Bonchev[76] about calculating indices for all connected subgraphs of a graph and summing them up and using the original Zagreb indices as indices.

The overall first Zagreb index OM_1 is equal to TM_1, but the overall second Zagreb index OM_2 differs from TM_2 because in the formula for computing OM_2 the second sum in (46) is replaced by the product of the edge-weights for each connected subgraph **S** and it also contains the value of the total adjacency index (see Eq. (3)):

$$ OM_2 = \sum_{\mathbf{S}} \prod_{\text{edges}} d_i d_j (\mathbf{S}) + A . \quad (47) $$

We illustrate the application of this formula by using the 2-methylbutane tree and its subgraphs, as given in Figure 4: (iii.2) $1 \times 3 + 1 \times 3 + 3 \times 2 + 2 \times 1 = 14$; (iii.3) $1 \times 3 \times 1 + 1 \times 3 \times 2 + 1 \times 3 \times 2 + 3 \times 2 \times 1 = 21$; (iv.4) $1 \times 3 \times 2 \times 1 + 1 \times 3 \times 2 \times 1 = 12$; (iii.5) $1 \times 3 \times 1 \times 2 = 6$; (iii.6) $1 \times 3 \times 1 \times 2 \times 1 = 6$; $A = 8$ and OM_2 (2-methylbutane) = 67. Bonchev and Trinajstić[31] said that the reason why they used the total adjacency index in formula (47) was to compute both OM_1 and OM_2 for a complete series of connected subgraphs.

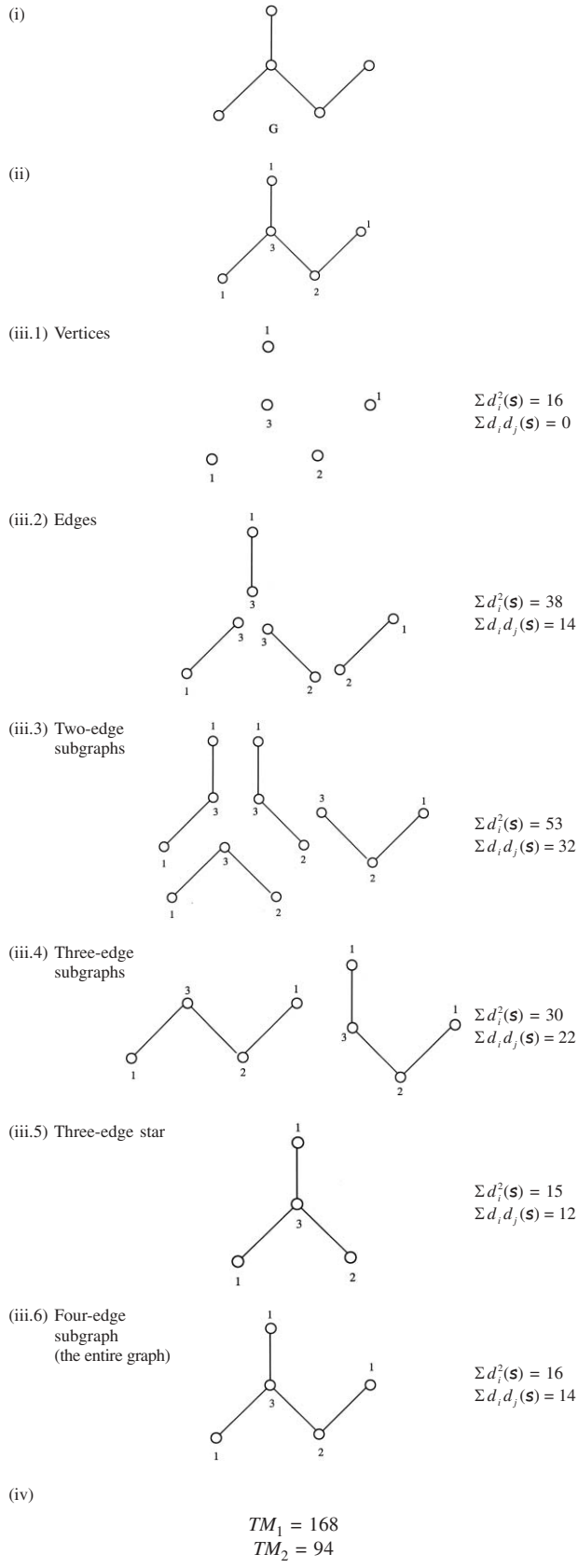

Figure 4. Calculation of the Zagreb complexity indices TM_1 and TM_2 for the 2-methylbutane tree. (i) 2-Methylbutane tree; (ii) Vertex-degrees in G; (iii) All connected subgraphs of G. Vertices in subgraphs retain degrees they have in G; (iv) The TM_1 and TM_2 indices of G.

Computation of the Zagreb complexity indices TM_1 and TM_2 for the 2-methylbutane tree is shown in Figure 4 whilst the computation of the TM_1* and TM_2* indices for the same tree is given in Figure 5.

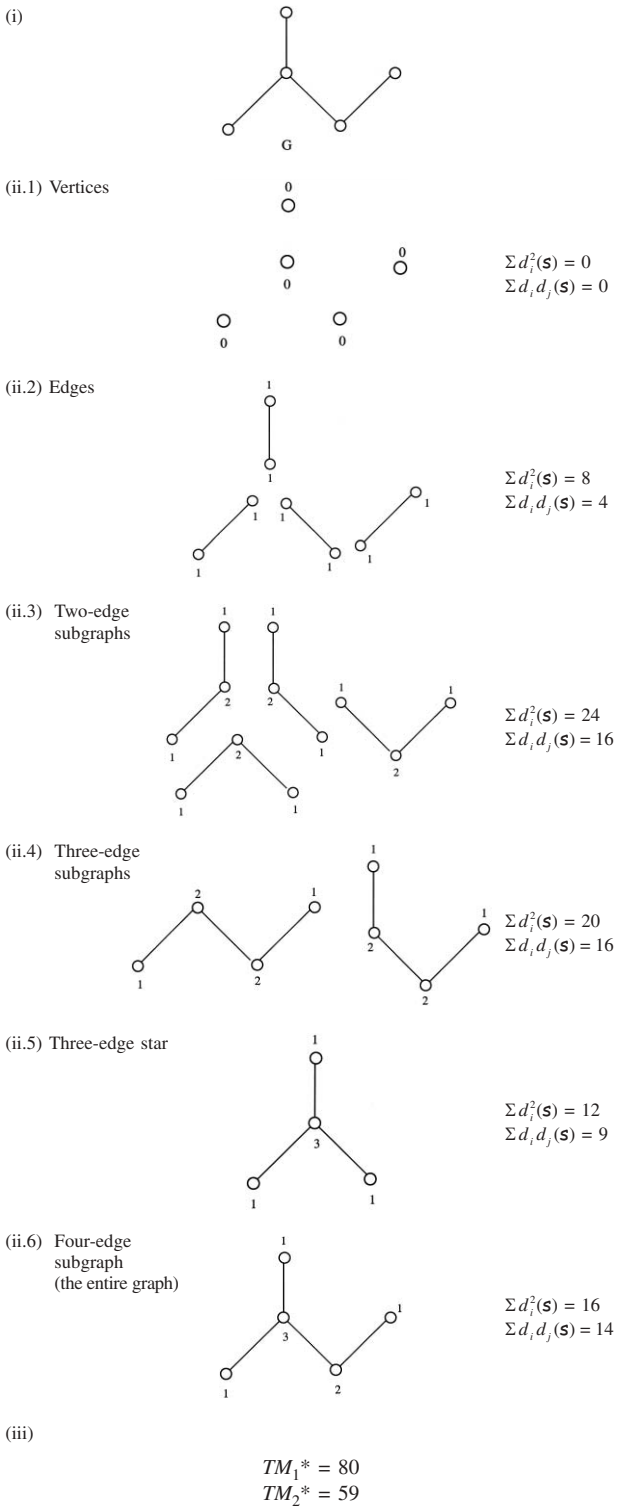

Figure 5. Calculation of the Zagreb complexity indices TM_1* and TM_2* for the 2-methylbutane tree. (i) 2-Methylbutane tree (ii) All connected subgraphs of G. Vertex-degrees are taken as they are in the isolated subgraphs; (iii) The TM_1* and TM_2* indices of G.

USE OF THE ZAGREB INDICES IN QSPR

We tested the use of the following Zagreb indices: M_1, M_2, ${}^mM_1$, ${}^mM_2$, TM_1, TM_1*, TM_2, and TM_2* in modeling boiling points of 38 C_3–C_8 alkanes. These compounds and their boiling points were selected because several QSPR studies that can be used for comparison already exist in the literature.[76,77] In Table I we give, besides the Zagreb indices and connectivity indices ${}^0\chi$ and ${}^1\chi$, the total walk count (*twc*) index[77–79] and the polarity number (p_3).[4,5,80] The latter two indices were included in our modeling for the following reasons: p_3 is the number of paths of length 3 and encodes the steric aspects of an alkane,[81] whilst the *twc* index has been shown to be very useful in producing multilinear structure-boiling point correlations.[77]

Our structure-boiling point modeling is based on the CROMRsel procedure.[38,82–84] This is a multivariate procedure that picks up the best possible model among the set of models obtained for a given number of parameters; the criterion of the model goodness being the standard error of estimate. In building our models we considered all possible combinations between topological indices given in Table I, starting from a single index up to five indices. The quality of models is expressed by fitted (descriptive) statistical parameters: the correlation coefficient r_{fit}, the standard error of estimate s_{fit} and F is the result of Fisher's test. In addition, the models were cross(internally)-validated by a leave-one-out procedure. Statistical parameters for the cross-validated models are denoted as r_{cv} and s_{cv}, where cv stands for the cross-validation procedure.

The standard errors of estimate, s_{fit} and s_{cv}, were computed using the following expressions:

$$s_{\text{fit}} = \left\{ \left[\sum_{i=1}^{N} (\mathrm{P}_i - \mathrm{P(fit)}_i^{\text{est}})^2 \right] / (N - I - 1) \right\}^{1/2} \quad (48)$$

$$s_{\text{cv}} = \left\{ \left[\sum_{i=1}^{N} (\mathrm{P}_i - \mathrm{P(cv)}_i^{\text{est}})^2 \right] / (N - I - 1) \right\}^{1/2} \quad (49)$$

where P_i denotes experimental properties, $\mathrm{P(fit)}_i^{\text{est}}$ estimated properties based on the fit statistical procedure, $\mathrm{P(cv)}_i^{\text{est}}$ estimated properties based on the leave-one-out cross-validation procedure, N is the number of data points used in the model building and I is the number of descriptors contained in the model.

Below we give the best models with $I = 1, 2, ..., 5$ descriptors:

(I) $I = 1$

$$\mathrm{bp} = -118.3\ (\pm 4.9) + 63.9\ (\pm 1.5)\ {}^1\chi \quad (50)$$

$N = 38$ $r_{\text{fit}} = 0.990$ $r_{\text{cv}} = 0.988$
$s_{\text{fit}} = 5.91$ $s_{\text{cv}} = 6.42$ $F = 1775$

(II) $I = 2$

$$\mathrm{bp} = -110.7\ (\pm 3.5) + 54.2\ (\pm 1.8)\ {}^1\chi + 0.9\ (\pm 0.1)\ M_2 \quad (51)$$

$N = 38$ $r_{\text{fit}} = 0.996$ $r_{\text{cv}} = 0.994$
$s_{\text{fit}} = 4.00$ $s_{\text{cv}} = 4.57$ $F = 1963$

(III) $I = 3$

$$\mathrm{bp} = -128.7\ (\pm 3.1) + 65.8\ (\pm 1.1)\ {}^1\chi + 0.0317\ (\pm 0.0030)\ TM_1 - 0.0288\ (\pm 0.0035)\ twc \quad (52)$$

$N = 38$ $r_{\text{fit}} = 0.998$ $r_{\text{cv}} = 0.997$
$s_{\text{fit}} = 2.69$ $s_{\text{cv}} = 3.20$ $F = 2895$

(IV) $I = 4$

$$\mathrm{bp} = -148.6\ (\pm 3.8) + 67.9\ (\pm 1.1)\ {}^1\chi + 6.8\ (\pm 0.9)\ {}^mM_1 + 0.12\ (\pm 0.01)\ TM_2 - 0.22\ (\pm 0.02)\ TM_2^* \quad (53)$$

$N = 38$ $r_{\text{fit}} = 0.999$ $r_{\text{cv}} = 0.998$
$s_{\text{fit}} = 2.10$ $s_{\text{cv}} = 2.82$ $F = 3582$

(V) $I = 5$

$$\mathrm{bp} = -107.1\ (\pm 2.4) + 41.9\ (\pm 1.3)\ M_1 - 45.3\ (\pm 1.6)\ M_2 + 49.3\ (\pm 1.5)\ p_3 + 0.0750\ (\pm 0.0087)\ TM_1 - 0.170\ (\pm 0.017)\ TM_1^* \quad (54)$$

$N = 38$ $r_{\text{fit}} = 0.999$ $r_{\text{cv}} = 0.998$
$s_{\text{fit}} = 1.94$ $s_{\text{cv}} = 3.10$ $F = 3355$.

Judging by the statistical parameters, model (53) appears to be the best model; model (54) seems to be less stable, since the values of s_{fit} (1.94) and s_{cv} (3.10) are very different.

In Figure 6, we give the plot between the experimental and calculated values of C_3–C_8 alkane boiling points for the fit and cross-validated model (53).

In Table II, we give fitted standard errors of estimate s_{fit} for some structure-boiling point models for C_3–C_8 alkanes taken from Table VII of Ref. 77, which incorporates Bonchev's models based on topological complexity indices.[76]

There are no single-descriptor models for C_3–C_8 alkanes given in Ref. 77. However, it is known[10,11] that the connectivity index gives good structure-boiling point models. Recently, Randić's very detailed study[90] on regressions based on a single descriptor has shown that the *Z*-index of Hosoya[6] is the only index that surpasses the

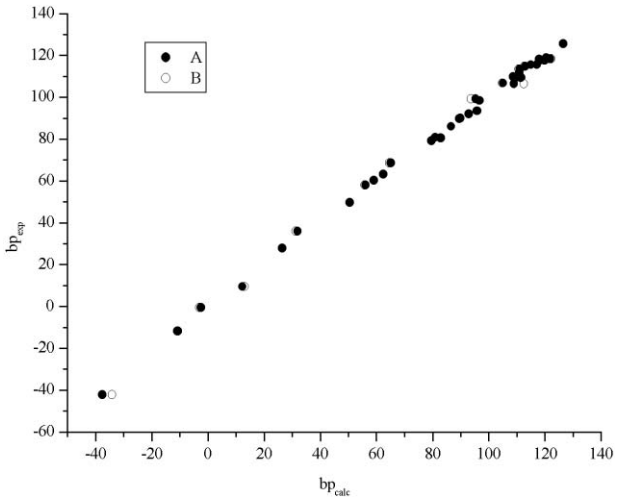

Figure 6. Scatter plots between the experimental and calculated alkane C_3–C_8 boiling points for the fit (A) and cross-validated (B) model (53).

TABLE II. Fitted standard errors of estimate s_{fit} for some structure-boiling point models for C_3–C_8 alkanes taken from Table VII of Ref. 77, which includes Bonchev's models based on topological complexity indices (Ref. 76)

Descriptors	I	s_{fit}
${}^0TC$,[a] ${}^1TC$[a]	2	4.92
$\chi^{0.33}$,[b] mwc_5[c]	2	2.21
${}^0TC$,[a] ${}^3TC$,[a] ${}^1TC1$[a]	3	3.75
p_3,[d] $W^{0.25}$,[e] V[f]	3	1.38
${}^0K$,[g] TC,[a] ${}^1TC$,[a] $TC1$[a]	4	2.37
I_{WG},[h] $D^{(2)}$,[i] ${}^2TC$,[a] ${}^3K$[j]	4	0.91
p_3,[d] $W^{0.25}$,[e] dia[k], V^2[l]	4	0.79
${}^1TC$,[a] TC,[a] ${}^1TC1$,[a] ${}^3TC1$,[a] $TC1$[a]	5	1.75
p_3[d], W^2,[m] $W^{0.25}$,[e] dia,[k] V^2[l]	5	0.72
I_{WG},[h] $D^{(2)}$,[i] M_2,[n] ${}^2TC1$,[a] ${}^2TC$[a]	5	0.64

[a] Bonchev's topological complexity indices.[76] [b] $\chi^{0.33}$ = the cubic root of Randić's connectivity index.[7,85] [c] mwc_5 = molecular walk count of order 5.[63,77–79,86] [d] p_3 = number of paths of length 3.[4,5,80] [e] W^{025} = the fourth root of the Wiener index.[5] [f] V = number of carbon atoms in an alkane. [g] ${}^0K = V$.[58,76] [h] I_{WG} = information-theoretic analog of the 3D Wiener index.[87,88] [i] $D^{(2)}$ = Balaban's distance-based topological index called the mean square distance.[89] [j] ${}^3K$ = total number of connected subgraphs of a graph with three edges.[58,76] [k] *dia* = graph-theoretical diameter, *i.e.*, the longest distance in a graph. [l] V^2 = number of carbon atoms in an alkane squared. [m] W^2 = Wiener number, squared. [n] M_2 = Second Zagreb index, defined in the text by Eq.(2).

connectivity indices in modeling structure-boiling point relationships for isomeric alkanes.

Among the two-descriptors models, the best is the model based on $\chi^{0.33}$ and mwc_5, obtained by Rücker and Rücker[77] (s_{fit} = 2.21 °C); however, our model (51) (s_{fit} = 4.00 °C) is better than two-parameter models we found in the literature, including Bonchev's model (s_{fit} = 4.92 °C)[76] and the model based on Randić's indices ${}^0\chi$ and ${}^2\chi$ (s_{fit} = 5.13 °C).[76]

In the case of three-descriptor models, again the model of Rücker and Rücker[77] (s_{fit} = 1.38 °C) is better than other models and our three-descriptor model (52) comes second (s_{fit} = 2.69 °C), surpassing all three-parameter models from the literature known to us, including Bonchev's model (s_{fit} = 3.75 °C)[76] and the model based on Randić's indices ${}^0\chi$, ${}^2\chi$ and ${}^5\chi$ (s_{fit} = 4.58 °C).[76] The same is true of four-descriptor models; the best model being the one reported by Rücker and Rücker[77] (s_{fit} = 0.79 °C). The second best models, close to the Rücker-Rücker model (s_{fit} = 0.91 °C), is the model by Bonchev.[76] Our four-descriptor model (54) (s_{fit} = 2.10 °C) is poorer than these two models, but it is a slightly better model than the other four-descriptor model reported by Bonchev (s_{fit} = 2.37 °C),[76] whilst the model based on four Randić's indices, ${}^0\chi$, ${}^2\chi$, ${}^3\chi$ and ${}^4\chi$, is poorer than our model (s_{fit} = 4.01 °C).[76]

The best five-descriptor model is that of Bonchev[76] (s_{fit} = 0.64 °C), based on the combination of the Zagreb M_2 index with the mean square distance,[89] two complexity indices and the information-theoretic analog of the 3D Wiener number.[87,88] Close to it is the Rücker-Rücker model[77] (s_{fit} = 0.72 °C). Our model (53) with s_{fit} = 1.71 °C is similar to another Bonchev's model[76] (s_{fit} = 1.75 °C), based on the set of complexity indices. Bonchev[58] also obtained several other five-parameter models based on the overall connectivity indices; the best being the model (s_{fit} = 1.60 °C) based on four complexity indices: the first-order overall connectivity index (${}^1TC$; 1 stands for edges in a graph), the third-order overall connectivity index for cluster subgraphs (${}^3TC_c$; c stands for cluster and 3 denotes clusters of three edges), the fourth-order overall connectivity index for path subgraphs (${}^4TC_p$; p stands for path subgraphs and 4 denotes the length of a path) and the fifth-order overall connectivity index for path subgraphs (${}^5TC_p$) and the total number of subgraphs with a single vertex (${}^0K$) is simply the number of vertices V. A five-descriptor model based on Randić's indices: ${}^0\chi$, ${}^2\chi$, ${}^3\chi$, ${}^4\chi$ and ${}^5\chi$ is poorer (s_{fit} = 3.32 °C) than any of the above models.[76]

INTERCORRELATION BETWEEN THE ZAGREB INDICES

We investigated the intercorrelation between the indices given in Table I. The intercorrelation matrix reflecting the pairwise linear correlation between the twelve indices computed for C_3–C_8 alkanes is given in Table III.

The degree of the intercorrelation was appraised by the correlation coefficient r. Pairs of indices with $r \geq 0.97$ are considered highly intercorrelated, those with $0.90 \leq r \leq 0.97$ appreciably correlated, those with $0.50 \leq r \leq 0.89$ weakly correlated and finally the pairs of indices with low r-values (< 0.50) not intercorrelated.[91] It appears that, according to the above classification, all

TABLE III. Intercorrelation matrix

	M_1	M_2	${}^mM_1$	${}^mM_2$	p_3	${}^0\chi$	${}^1\chi$	TM_1	TM_1*	TM_2	TM_2*	*twc*
M_1	1.000	0.990	0.936	0.646	0.879	0.977	0.827	0.876	0.906	0.863	0.896	0.879
M_2		1.000	0.935	0.653	0.931	0.965	0.818	0.906	0.936	0.902	0.929	0.904
${}^mM_1$			1.000	0.364	0.787	0.852	0.595	0.914	0.909	0.893	0.901	0.842
${}^mM_2$				1.000	0.752	0.786	0.957	0.373	0.484	0.405	0.481	0.552
p_3					1.000	0.892	0.830	0.835	0.885	0.863	0.885	0.868
${}^0\chi$						1.000	0.926	0.789	0.845	0.785	0.836	0.842
${}^1\chi$							1.000	0.561	0.654	0.576	0.647	0.697
TM_1								1.000	0.990	0.996	0.989	0.952
TM_1*									1.000	0.992	1.000	0.980
TM_2										1.000	0.993	0.960
TM_2*											1.000	0.981
twc												1.000

the indices considered are intercorrelated to a certain extent, except for five pairs (${}^mM_1$ and ${}^mM_2$, ${}^mM_2$ and TM_1, ${}^mM_2$ and TM_1*, ${}^mM_2$ and TM_2, ${}^mM_2$ and TM_2*). There are ten pairs of highly intercorrelated indices, but there is only one pair (TM_1* and TM_2*) with perfect intercorrelation (r = 1.00). All other indices (51 pairs) are either appreciably intercorrelated (16 pairs) or weakly intercorrelated (35).

It is noteworthy that many good models listed above contain combinations of indices regardless of their intercorrelation status. This teaches us, as Randić stressed several times,[16,92,93] that the intercorrelation criterion should not be always used for filtering descriptors to be used in building QSPR models.

CONCLUDING REMARKS

A number of known results about the Zagreb M_1 and M_2 indices are reviewed. A novel form of Zagreb indices, named the modified Zagreb indices and denoted by ${}^mM_1$ and ${}^mM_2$, is introduced. The modified Zagreb ${}^mM_1$ index gives a greater contribution to outer atoms than to inner atoms in a molecule. Similarly, the modified Zagreb ${}^mM_2$ index gives a greater contribution to outer bonds than to inner bonds in a molecule. This is opposite to the behavior of the original Zagreb indices and in agreement with the chemists' understanding that the most important contributions to the interactions between molecules that are essential for many of their physical, chemical, biological and even technological properties arise from the more exposed atoms and bonds. Zagreb M_1 and M_2 indices, their modified forms ${}^mM_1$ and ${}^mM_2$ and Zagreb complexity indices TM_1, TM_1*, TM_2 and TM_2* together with the connectivity indices ${}^0\chi$ and ${}^1\chi$, the total walk count *twc* index and the polarity number p_3 have been used in the structure-boiling point modeling of C_3–C_8 alkanes, using the CROMRsel procedure. Models obtained are not as good as those produced by Rücker and Rücker[77] and some of Bonchev,[76] but they are comparable, if not better, to other models of Bonchev and other authors.[76]

Acknowledgments. – This work was supported in part by the Ministry of Science and Technology of the Republic of Croatia – Grant No. 0098034. We are grateful to Dr. Bono Lučić for helping with the CROMRsel program. We thank the referees for reviewing twice our manuscript and for many useful suggestions to improve the presentation.

REFERENCES

1. I. Gutman and N. Trinajstić, *Chem. Phys. Lett*. **17** (1972) 535–538.
2. J. Devillers and A. T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon & Breach, Amsterdam, 1999.
3. S. C. Basak, B. D. Gute, and G. D. Grunwald, *Croat. Chem. Acta* **69** (1996) 1159–1173.
4. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000.
5. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
6. H. Hosoya, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.
7. M. Randić, *J. Am. Chem. Soc*. **97** (1975) 6609–6615.
8. L. B. Kier, L. H. Hall, W. J. Murray, and M. Randić, *J. Pharm. Sci.* **64** (1975) 1971–1974.
9. L. B. Kier, L. H. Hall, W. J. Murray, and M. Randić, *J. Pharm. Sci.* **65** (1976) 1226–1230.
10. L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic, New York, 1976.
11. L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Wiley, New York, 1986.
12. A. T. Balaban (Ed.), *From Chemical Topology to Three-Dimensional Geometry*, Plenum, New York, 1997.
13. M. Karelson, *Molecular Descriptors in QSAR/QSPR*, Wiley-Interscience, New York, 2000.
14. M. V. Diudea (Ed.), *QSPR/QSAR Studies by Molecular Descriptors*, Nova, Huntington, N. Y., 2001.

15. L. H. Hall and L. B. Kier, *J. Mol. Graphics Modell.* **20** (2001) 4–18.
16. M. Randić, *J. Mol. Graphics Modelling* **20** (2001) 19–35.
17. J. Rada, O. Araujo, and I. Gutman, *Croat. Chem. Acta* **74** (2001) 225–235.
18. S. Nikolić and N. Raos, *Croat. Chem. Acta* **74** (2001) 621–631.
19. I. Gutman, *Croat. Chem. Acta* **75** (2002) 357–369.
20. L. B. Kier and L. H. Hall, *Croat. Chem. Acta* **75** (2002) 371–382.
21. (a) L. Pogliani, *Chem. Rev.* **100** (2000) 3827–3858. (b) L. Pogliani, *Croat. Chem. Acta* **75** (2002) 409–432.
22. N. Kezele, L. Klasinc, J. von Knop, S. Ivaniš, and S. Nikolić, *Croat. Chem. Acta* **75** (2002) 651–661.
23. S. H. Bertz and W. F. Wright, *Graph Theory Notes of New York* **35** (1998) 32–48.
24. S. Nikolić, I. M. Tolić, and N. Trinajstić, *MATCH – Comm. Math. Comput. Chem.* **40** (1999)187–201.
25. S. Nikolić, N. Trinajstić, and I. M. Tolić, *J. Chem. Inf. Comput. Sci.* **40** (2000) 920–926.
26. S. Nikolić, I. M. Tolić, N. Trinajstić, and I. Baučić, *Croat. Chem. Acta* **73** (2000) 909–921.
27. S. Nikolić, N. Trinajstić, I. M. Tolić, G. Rücker, and C. Rücker, in: D. Bonchev and D. H. Rouvray (Eds.), *Complexity – Introduction and Fundamentals*, Francis & Taylor, London, 2003, pp. 29–89.
28. A. Golbraikh, D. Bonchev, and A. Tropsha, *J. Chem. Inf. Comput. Sci.* **41** (2001) 147–158.
29. A. Golbraikh, D. Bonchev, and A. Tropsha, *J. Chem. Inf. Comput. Sci.* **42** (2002) 769–787.
30. A. Miličević and S. Nikolić, *Croat. Chem. Acta,* in press.
31. D. Bonchev and N. Trinajstić, *SAR QSAR Environ. Res*. **12** (2001) 213–236.
32. A. T. Balaban, *Theoret. Chim. Acta* **53** (1979) 355–375.
33. A. T. Balaban, *Chem. Phys. Lett*. **89** (1982) 399–404.
34. S. C. Basak and G. D. Grunwald, *SAR QSAR Environ. Res.* **3** (1995) 265–277.
35. S. C. Basak, G. D. Grunwald, and G. J. Niemi, in: A. T. Balaban (Ed.), *From Chemical Topology to Three-Dimensional Geometry*, Plenum Press, New York, 1997, pp. 73–116.
36. B. D. Gute and S. C. Basak, *SAR QSAR Environ. Res.* **3** (1997) 117–131.
37. S. C. Basak, B. D. Gute, and S. Ghatak, *J. Chem. Inf. Comput. Sci*. **39** (1999) 255–260.
38. S. C. Basak, B. D. Gute, B. Lučić, S. Nikolić, and N. Trinajstić, *Comput. Chem.* **24** (2000) 181–191.
39. S. C. Basak, A. T. Balaban, G. D. Grunwald, and B. D. Gute, *J. Chem. Inf. Comput. Sci.* **40** (2000) 891–898.
40. X. Yao, X. Zhang, R. Zhang, M. Liu, Z. Hu, and B. Fan, *Comput. Chem.* **25** (2001) 475–482.
41. P. V. Khadikar, S. Karmarkar, and V. K. Agrawal, *J. Chem. Inf. Comput. Sci.* **41** (2001) 934–949.
42. S. Sardana and A. K. Madan, *MATCH – Comm. Math. Comput. Chem.* **45** (2002) 35–53.
43. R. Natarajan, I. Nirdosh, S. C. Basak, and D. R. Mills, *J. Chem. Inf. Comput. Sci.* **42** (2002) 1425–1430.
44. E. Estrada, L. Rodriguez, and A. Gutierrez, *MATCH – Comm. Math. Comput. Chem.* **35** (1997) 145–156.
45. E. Estrada, *Chem. Phys. Lett.* **336** (2001) 248–252.
46. D. Vukičević and N. Trinajstić, *Croat. Chem. Acta* **76** (2003) 183–187.
47. S. C. Basak, D. K. Harriss, and V. R. Magnuson, POLLY 2.3, Copyright of the University of Minnesota, 1988.
48. O. Mekenyan, S. Karaburnaliev, and D. Bonchev, *Comput. Chem.* **14** (1990) 193–200.
49. R. Todeschini, DRAGON, WebSite – http://www.disat.unimib.it/chm/
50. Cerius^2 3.0 QSAR+, Molecular Simulation Inc., 9685 Scranton Road, San Diego, CA 92121, USA.
51. M. Vedrina, S. Marković, S. Medić-Šarić, and N. Trinajstić, *Comput. Chem*. **21** (1997) 355–361.
52. D. R. Flower, *J. Mol. Graphics Modelling* **16** (1998) 239–253.
53. N. Trinajstić, *Chemical Graph Theory*, 2nd revised edition, CRC Press, Boca Raton, FL, 1992.
54. F. Harary, *Graph Theory*, second printing, Addison-Wesley, Reading, MA, 1972.
55. I. Gutman, B. Ruščić, N. Trinajstić, and C. F. Wilcox, Jr., *J. Chem. Phys.* **62** (1975) 3399–3405.
56. D. Bonchev, O. Mekenyan, and A. T. Balaban, *J. Chem. Inf. Comput. Sci.* **17** (1980) 106–113.
57. D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures*, Wiley, Chichester, 1983.
58. D. Bonchev, *J. Mol. Graphics Modelling* **20** (2001) 65–75.
59. J. M. Quigley and S. M. Naughton, *J. Chem. Inf. Comput. Sci*. **42** (2002) 976–982.
60. E. Estrada, *J. Chem. Inf. Comput. Sci*. **35** (1996) 1001–1009.
61. S. Nikolić, N. Trinajstić, and I. Baučić, *J. Chem. Inf. Comput. Sci.* **38** (1998) 42–46.
62. S. C. Basak, S. Nikolić, N. Trinajstić, D. Amić, and D. Bešlo, *J. Chem. Inf. Comput. Sci*. **40** (2000) 927–933.
63. I. Gutman, C. Rücker, and G. Rücker, *J. Chem. Inf. Comput. Sci* **41** (2001) 739–745.
64. D. Bonchev, *J. Mol. Struct. (Theochem)* **336** (1995) 137–156.
65. D. Bonchev, X. Liu, and D. J. Klein, *Croat. Chem. Acta* **66** (1993) 1421–150.
66. M. Gordon and G. R. Scantlebury, *Trans. Faraday Soc*. **60** (1964) 604–621.
67. J. R. Platt, *J. Chem. Phys.* **15** (1947) 419–420.
68. W. T. Tutte, *Connectivity in Graphs*, University of Toronto Press & University of Oxford Press, London, 1966.
69. M. Barysz, D. Plavšić, and N. Trinajstić, *MATCH – Comm. Math. Chem.* **19** (1986) 89–116.
70. M. Randić and J. Zupan, *J. Chem. Inf. Comput. Sci.* **41** (2001) 550–560.
71. S. Nikolić, N. Trinajstić, and M. Randić, *Chem. Phys. Lett*. **333** (2001) 319–321.
72. S. H. Bertz, in: R. B. King (Ed.), *Chemical Applications of Topology and Graph Theory*, Elsevier, Amsterdam, 1983, pp. 206–221.
73. W. H. Brock, *The Fontana History of Chemistry*, Fontana Press, London, 1992.
74. D. Grdenić, *History of Chemistry*, Novi Liber/Školska knjiga, Zagreb, 2001.
75. *e.g.*, I. Muegge, *Chem. Eur. J.* **8** (2002) 1976–1981.
76. D. Bonchev, *SAR QSAR Environ. Res*. **7** (1997) 23–43.
77. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **39** (1999) 788–802.
78. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **33** (1993) 683–695.
79. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **40** (2000) 99–106.

80. I. Lukovits and W. Linert, *J. Chem. Inf. Comput. Sci.* **38** (1998) 715–719.
81. J. R. Platt, *J. Phys. Chem*. **15** (1947) 328–336.
82. D. Amić, D. Bešlo, B. Lučić, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **38** (1998) 819–822.
83. B. Lučić and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **39** (1999) 121–132.
84. B. Lučić, A. Miličević, S. Nikolić, and N. Trinajstić, *Croat. Chem. Acta* **75** (2002) 847–868.
85. M. Randić, P. J. Hansen, and P. C. Jurs, *J. Chem. Inf. Comput. Sci*. **38** (1988) 60–68.
86. G. Rücker and C. Rücker, *J. Chem. Inf. Comput. Sci*. **34** (1994) 534–538.
87. O. Mekenyan, D. Peitchev, D. Bonchev, N. Trinajstić, and I. Bangov, *Drug Res.* **36** (1986) 176–183.
88. B. Bogdanov, S. Nikolić, and N. Trinajstić, *J. Math. Chem.* **3** (1989) 299–309.
89. A. T. Balaban, *Pure Appl. Chem*. **55** (1983) 199–206.
90. M. Randić, *Croat. Chem. Acta* **66** (1993) 289–312.
91. N. Trinajstić, S. Nikolić, S. C. Basak, and I. Lukovits, *SAR QSAR Environ. Res*. **12** (2001) 31–54.
92. M. Randić and N. Trinajstić, *J. Mol. Struct*. **284** (1993) 209–221.
93. M. Randić and N. Trinajstić, *J. Mol. Struct*. **300** (1993) 551–571.

SAŽETAK

Zagrebački indeksi 30 godina poslije

Sonja Nikolić, Goran Kovačević, Ante Miličević i Nenad Trinajstić

Prikazana je izvorna formulacija Zagrebačkih indeksa i razmatrana je njihova veza s nekoliko topologijskih indeksa kao što su povratne šetnje, Plattov indeks, Gordon-Scantleburyjev indeks i indeksi povezanosti. Također su navedena i neka svojstva izvornih Zagrebačkih indeksa. Uvedeni su modificirani Zagrebački indeksi, a ukratko su prikazani Zagrebački indeksi kompleksnosti koji su nedavno predloženi. Uporaba Zagrebačkih indeksa u modeliranju odnosa strukture i svojstava molekula ilustrirana je na primjeru predviđanja vrelišta nižih alkana. Modeliranje je provedeno pomoću našega izvornoga postupka, nazvanoga CROMRsel. Dobiveni modeli su bolji od mnogih objavljenih u literaturi, ali je primijećeno da Zagrebački indeksi obično ne sudjeluju u gradnji najboljih modela za predviđanja vrelišta alkana. Međutim, najbolji model u literaturi za predviđanja vrelišta nižih alkana sagrađen je od pet deskriptora od kojih je jedan Zagrebački indeks označen s M_2.