molecules

On Use of the Variable Zagreb $^{v}M_2$ Index in QSPR: Boiling Points of Benzenoid Hydrocarbons

Sonja Nikolić [1],*, Ante Miličević [1], Nenad Trinajstić [1] and Albin Jurić [2]

[1] The Rugjer Bošković Institute, Bijenička 54, P.O.B. 180, HR-10002 Zagreb, Croatia
[2] Department of Chemistry, Agricultural Institute, HR-43260 Križevci, Croatia

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

Abstract: The variable Zagreb $^{v}M_2$ index is introduced and applied to the structure-boiling point modeling of benzenoid hydrocarbons. The linear model obtained (the standard error of estimate for the fit model S_{fit}=6.8 ^{o}C) is much better than the corresponding model based on the original Zagreb M_2 index (S_{fit}=16.4 ^{o}C). Surprisingly, the model based on the variable vertex-connectivity index (S_{fit}=6.8 ^{o}C) is comparable to the model based on $^{v}M_2$ index. A comparative study with models based on the vertex-connectivity index, edge-connectivity index and several distance indices favours models based on the variable Zagreb $^{v}M_2$ index and variable vertex-connectivity index. However, the multivariate regression with two-, three- and four-descriptors gives improved models, the best being the model with four-descriptors (but $^{v}M_2$ index is not among them) with S_{fit}=5 ^{o}C, though the four-descriptor model contaning $^{v}M_2$ index is only slightly inferior (S_{fit}=5.3 ^{o}C).

Keywords: Zagreb M_2 index, variable Zagreb $^{v}M_2$ index, QSPR modeling, benzenoid hydrocarbons, variable connectivity index, distance indices

Introduction

The concept of the variable molecular descriptors was proposed as an alternative way of characterizing heteroatoms in molecules [1,2], but also to assess the structural differences, such as, for example, the relative role of carbon atoms of acyclic and cyclic parts in alkylcycloalkanes [3]. The

idea behind the variable molecular descriptors is that the variables are determined during the regression so that the standard error of estimate for a studied property is as small as possible.

Several molecular descriptors have already been tested in their variable forms in *QSPR* and *QSAR* [4-12]. Here we report the use of the variable Zagreb $^vM_2$ index in the structure-boiling point modeling of benzenoid hydrocarbons. We selected benzenoid hydrocarbons because there are several structure-boiling point models of these compounds already published [13,14]. Due to this fact, we were also able to carry out a comparative study of the model based on $^vM_2$ index against the models based on the standard vertex-connectivity index, variable vertex-connectivity index, edge-connectivity index and several distance indices. Since the Zagreb index in its original form was derived using graph-theoretical concepts and terminology [15], we will use these in the present report. Graphs will be generated from molecules in the usual way by replacing atoms with vertices and bonds with edges [16]. Besides, graphs that we will use will represent only carbon skeletons of benzenoid hydrocarbons. Therefore, benzenoid hydrocarbons in this report will be presented as various arrangements of hexagons in the plane.

The Zagreb M_2 index and Its Variable Form $^vM_2$

Originally, the Zagreb M_2 index together with the Zagreb M_1 index appeared in the topological formula for the total π-electron energy of conjugated molecules [17]:

$$M_2 = \sum_{\text{edges}} d(i)\, d(j) \tag{1}$$

where $d(i)$ is the degree of vertex i and $d(i)\, d(j)$ is the weight of edge i-j. This index was first used as a branching index [18] and later as a useful molecular descriptor in various forms in QSPR and QSAR studies [19-23].

The easiest way to introduce the variable Zagreb $^vM_2$ index is by means of an example. For this purpose we will use a graph *G* representing the carbon skeleton of naphthalene (see Figure 1).

Figure 1. Graph G representing the carbon skeleton of naphthalene

G

Since it is known from the chemistry of benzenoid hydrocarbons [24,25] that the carbon atoms with two adjacent carbon atoms possess different characteristics then the carbon atoms with three adjacent atoms, we assess their relative roles by differentiating these two groups with variable parameters. Using the graph-theoretical approach, the difference between the two groups of carbon atoms is expressed by means of the degrees of the corresponding vertices plus the variables. Hence, the vertex-degree of the carbon atom adjacent to two other carbon atoms is taken to be:

$$d = 2 + x \qquad\qquad (2)$$

and likewise the vertex-degree of the carbon atom adjacent to three other carbon atoms is given by:

$$d = 3 + y \qquad\qquad (3)$$

Putting (2) and (3) into (1), one obtains the following Zagreb $^vM_2$ index for naphthalene as a function of the variables x and y:

$$^vM_2 = 6\,(2 + x)^2 + 4\,(2 + x)\,(3 + y) + (3 + y)^2 \qquad\qquad (4)$$

In general, the variable Zagreb $^vM_{2i}$ index of a benzenoid hydrocarbon i can be given as:

$$^vM_{2i} = c_{1i}\,(2 + x)^2 + c_{2i}\,(2 + x)\,(3 + y) + c_{3i}\,(3 + y)^2 \qquad\qquad (5)$$

We denote the boiling point of a benzenoid hydrocarbon i ($i=1,\dots,21$) by bp_i. Thus, for naphthalene ($i=1$), $bp_1=218°C$ and c_{11}, c_{21} and c_{31} are respectively 6, 4 and 1. In Table 1 we give bp_i, c_{1i}, c_{2i} and c_{3i} values for 21 benzenoid hydrocarbons whose graphs are given in Figure 2.

Figure 2. Graphs corresponding to studied benzenoid hydrocarbons

Table 1. The values of experimental boiling points (bp_i in $^\circ$C, i=1,…,21) and
coefficients (c_{1i}, c_{2i} and c_{3i}) of the variable Zagreb $^vM_2$ indices of 21
benzenoid hydrocarbons.

Benzenoid hydrocarbon	bp_i	$^vM_2$ index		
		c_{1i}	c_{2i}	c_{3i}
1	218	6	4	1
2	338	7	6	3
3	340	6	8	2
4	431	8	8	5
5	425	7	10	4
6	429	9	6	6
7	440	6	12	3
8	496	7	10	7
9	493	8	8	8
10	497	8	8	8
11	547	6	12	9
12	542	7	10	10
13	535	9	10	7
14	536	8	12	6
15	531	8	12	6
16	519	9	10	7
17	590	6	12	12
18	592	9	10	10
19	596	8	12	9
20	594	8	12	9
21	595	9	10	10

Experimental values of boiling points of considered benzenoid hydrocarbons are taken from
Randić [13].

Expression (5) for x=0 and y=0 reduces to:

$$M_{2i} = 2^2\, c_{1i} + 2\cdot3\, c_{2i} + 3^2\, c_{3i} \qquad\qquad (6)$$

which is the formula for computing the original Zagreb M_2 index of a given benzenoid i. It should also
be noted that the variable connectivity index $^v\chi_i$ is related to the Zagreb $^vM_{2i}$ index with the same set of
coefficients:

$$^v\chi_i = c_{1i}\,(2+x)^{-1} + c_{2i}\,[(2+x)\,(3+y)]^{-1/2} + c_{3i}\,(3+y)^{-1} \qquad (7)$$

For $x=0$ and $y=0$, eq. (7) reduces to one for computing the vertex-connectivity index of a benzenoid hydrocarbon i:

$$^{v}\chi_i = 2^{-1} c_{1i} + (2 \cdot 3)^{-1/2} c_{2i} + 3^{-1} c_{3i} \tag{8}$$

Results and Discussion

In order to find optimal variable Zagreb $^{v}M_2$ index, the values of x were varied in the range between –2 and 2 and values of y were varied in the range between –3 and 3, both in steps of 0.1. This range of x and y values was imposed by the degrees of valences in benzenoid graphs. In non-optimized Zagreb index (M_2), the values of variables x and y are equal to 0.0. We want to see are there optimal values of x and y near their non-optimized values (0.0, 0.0) for which the standard error of estimate of the structure-boiling point model reaches minimum. For each pair (x, y) in the given range, coefficients a_0 and a_1 in the linear regression model:

$$bp = a_0 + a_1 \,^{v}M_2 \tag{9}$$

were computed using the least square fitting procedure as implemented in the CROMRsel program [26-28]. 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, the Fisher´s values. S_{fit} was computed with N and N-I-1 in the denominator, where N is the number of considered benzenoid hydrocarbons and I is the number of descriptors used in the model. In addition, the models were cross(internally)-validated using the leave-one-out method. Statistical parameters for the cross-validated models are denoted by R_{cv} and S_{cv}, where cv denotes the cross-validation.

The linear structure-boiling point model based on the Zagreb $^{v}M_2$ index with the optimum values of x (0.0) and y (-1.2) is as follows:

$$bp = -11.0 \,(\pm 8.2) + 5.767 \,(\pm 0.093) \,^{v}M_2 \tag{10}$$

$N=21$ $R_{fit}=0.998$ $S_{fit}(N)=6.8$ $S_{fit}(N\text{-}I\text{-}1)=7.2$ $R_{cv}=0.997$ $S_{cv}(N)=8.0$ $S_{cv}(N\text{-}I\text{-}1)=8.4$ $F=3866$

Figure 3. Scatter plot between the fit standard error of estimate ($S_{fit}(N)$) and x values in the range from –0.4 to +1.1 for the optimum value of y (-1.2)

Figure 4. Scatter plot between the fit standard error of estimate ($S_{fit}(N)$) and *y* values in the range from –3 to +1.1 for the optimum value of x (0.0)

Figure 5. Scatter plot between the cross validated standard errors of estimate ($S_{cv}(N)$) and *x* values in the range from –0.4 to +1.1 for the optimum value of *y* (-1.2)

Figure 6. Scatter plot between the cross validated standard errors of estimate ($S_{cv}(N)$) and *y* values in the range from −3 to +1.1 for the optimum value of x (0.0)

Figure 7. Scatter plots between the experimental (bp_{exp}) and calculated values of the fit (bp_{fit}) and cross-validated (bp_{cv}) models, respectively

Figure 8. Scatter plots of fit residuals versus fitted values (bp_{fit}) and of cross validated residuals versus cross validated values (bp_{cv}), respectively

In Figure 3, we give the scatter plot between $S_{fit}(N)$ and x values for the optimum value of y (-1.2) and in Figure 4, we give the scatter plot between $S_{fit}(N)$ and y values for the optimum value of x (0.0). The results for the fitted models were supported by the results for the cross-validated models (see Figures 5 and 6).

In Figure 7, we give plots between the experimental and calculated boiling points for fit and cross-validated models and in Figure 8 we give the scatter plots of fit residuals against the fitted values and cross-validated residuals against cross-validated boiling points, respectively.

Model (10) is much superior to the linear model based on the original M_2 index:

$$bp = 98 \ (\pm 16) + 2.63 \ (\pm 0.10) \ M_2 \tag{11}$$
$N=21 \ R_{fit}=0.986 \ S_{fit}(N)=16.4 \ S_{fit}(N\text{-}I\text{-}1)=17.2 \ R_{cv}=0.980 \ S_{cv}(N)=19.6 \ S_{cv}(N\text{-}I\text{-}1)=20.6 \ F=656$

We also derived the structure-boiling point model using the variable vertex-connectivity index ${}^{v}\chi$ and the CROMRsel procedure. The following linear model is obtained for the optimum parameters x (0.0) and y (0.5):

$$bp = -11.5 \ (\pm 8.2) + 46.27 \ (\pm 0.74) \ {}^{v}\chi \tag{12}$$
$N=21 \ R_{fit}=0.998 \ S_{fit}(N)=6.8 \ S_{fit}(N\text{-}I\text{-}1)=7.2 \ R_{cv}=0.997 \ S_{cv}(N)=8.0 \ S_{cv}(N\text{-}I\text{-}1)=8.4 \ F=3866$

Both models, (10) and (12), possess practically identical statistical parameters. This is unexpected results, since the vertex-connectivity index [29] is superior to the Zagreb M_2 index in building QSPR models, though there are indications [22] that the various variable forms of these two indices lead to the models of the same quality regarding the statistical parameters. However, that is so because both eqs. (5) and (7) are based on the same coefficients c_{1i}, c_{2i} and c_{3i} at constant factors.

We also considered the following linear model:

$$bp_i = b_0 + b_1 \, c_{1i} + b_2 \, c_{2i} + b_3 \, c_{3i} \tag{13}$$

where c_{1i}, c_{2i} and c_{3i} are taken from Table 1.

The resulting model:

$$bp = -11.2 \ (\pm 14.2) + 22.8 \ (\pm 1.6) \ c_1 + 21.2 \ (\pm 0.9) \ c_2 + 18.4 \ (\pm 0.7) \ c_3 \qquad (14)$$

$N=21 \quad R_{fit}=0.998 \quad S_{fit}(N)=6.8 \quad S_{fit}(N\text{-}I\text{-}1)=7.5 \quad R_{cv}=0.996 \quad S_{cv}(N)=9.2 \quad S_{cv}(N\text{-}I\text{-}1)=10.2 \quad F=1176$

possesses fit statistical parameters R_{fit} and $S_{fit}(N)$ almost identical to those in the models (10) and (12). These parameters would be exactly the same if x and y were obtained more accurately. Because of the greater number of descriptors used in model (14) its $S_{fit}(N\text{-}I\text{-}1)$, $S_{cv}(N)$, $S_{cv}(N\text{-}I\text{-}1)$, R_{cv} and F values are somewhat worse.

Models (10) and (14) are related by:

$$b_0 = a_0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (15)$$
$$b_1 = a_1 \ (2+x)^2 \qquad\qquad\qquad\qquad\qquad\qquad\quad (16)$$
$$b_2 = a_1 \ (2+x)\ (3+y) \qquad\qquad\qquad\qquad\qquad (17)$$
$$b_3 = a_1 \ (3+y)^2 \qquad\qquad\qquad\qquad\qquad\qquad\quad (18)$$

Similar relationships exist also between models (12) and (14).

Models (10) and (12) possess better statistical properties than the models we found in the literature. Below we give the models based on the vertex-connectivity index [13,14] and the edge-connectivity index [14], respectively:

$$bp = -34.4 \ (\pm 9.7) + 52.58 \ (\pm 0.95) \ \chi \qquad\qquad\qquad (19)$$

$N=21 \quad R_{fit}=0.997 \quad S_{fit}(N\text{-}I\text{-}1)=8.1 \quad F=3045$

$$bp = 13.9 \ (\pm 8.6) + 40.07 \ (\pm 0.71) \ \varepsilon \qquad\qquad\qquad (20)$$

$N=21 \quad R_{fit}=0.997 \quad S_{fit}(N\text{-}I\text{-}1)=7.9 \quad F=3179$

where χ denotes the vertex-connectivity index [29], while ε denotes the edge-connectivity index [30]. Both models (19) and (20) are fitted models − they were not cross-validated.

The quadratic models with χ and ε do not represent an improvement and they also have not been cross-validated [13,14]:

$$bp = -50 \ (\pm 35) + 56.3 \ (\pm 8.1) \ \chi - 0.21 \ (\pm 0.45) \ \chi^2 \qquad (21)$$

$N=21 \quad R_{fit}=0.997 \quad S_{fit}(N\text{-}I\text{-}1)=8.2 \quad F=1460$

$$bp = -36 \ (\pm 26) + 50.0 \ (\pm 5.0) \ \varepsilon - 0.47 \ (\pm 0.23) \ \varepsilon^2 \qquad (22)$$

$N=21 \quad R_{fit}=0.998 \quad S_{fit}(N\text{-}I\text{-}1)=7.3 \quad F=1848$

The multivariate regression with two descriptors (ε and *ws*) also does not represent the improvement over the models (10) and (18) [14]:

$$bp = -32 \ (\pm 22) + 48.0 \ (\pm 3.5) \ \varepsilon + 3.6 \ (\pm 1.6) \ ws \qquad (23)$$

$N = 21 \quad R_{fit} = 0.998 \quad S_{fit}(N\text{-}I\text{-}1) = 7.2 \quad F = 1944$

where *ws* is a Wiener-like index [31], obtained from the quotient matrix $\mathbf{\Delta/D}$. Δ is the detour matrix [32,33] and **D** is the distance matrix [34].

We also carried out the multivariate regression with two, three and four descriptors using the CROMRsel procedure and the obtained fitted model were cross-validated. Descriptors we considered were $^{v}M_2$, χ, ε and three distance indices (the Wiener sum index *WS*, *ws*, the detour index ω). The *WS*

index is a Wiener-like index, obtained from the quotient matrix **D/Δ** [35] and ω is equal to the half-sum of the elements of the detour matrix [32,36]. We considered χ, ε, WS, ws and ω indices because they have been used in the previous structure-boiling point studies of benzenoid hydrocarbons [13,14]. Below we give the best obtained models followed by the best models containing the $^vM_2$ index – in the case of the two-descriptor model the best model contains the $^vM_2$ index:

(i) The two-descriptor model

$$bp = -40\ (\pm 22) + 6.42\ (\pm 0.46)\ ^vM_2 - 0.021\ (\pm 0.014)\ ws \qquad (24)$$

$N=21$ $R_{fit}=0.998$ $S_{fit}(N)=6.5$ $S_{fit}(N\text{-}I\text{-}1)=7.0$ $R_{cv}=0.996$ $S_{cv}(N)=8.4$ $S_{cv}(N\text{-}I\text{-}1)=8.9$ $F=2048$

(ii) The three-descriptor models

$$bp = -46\ (\pm 18) + 52.7\ (\pm 3.3)\ \varepsilon - 0.167\ (\pm 0.044)\ ws + 0.037\ (\pm 0.012)\ \omega \qquad (25)$$

$N=21$ $R_{fit}=0.999$ $S_{fit}(N)=5.3$ $S_{fit}(N\text{-}I\text{-}1)=5.9$ $R_{cv}=0.997$ $S_{cv}(N)=7.5$ $S_{cv}(N\text{-}I\text{-}1)=8.2$ $F=1936$

$$bp = -47\ (\pm 24) + 37.9\ (\pm 3.4)\ ^vM_2 - 0.32\ (\pm 0.45)\ ws - 0.024\ (\pm 0.015)\ WS \qquad (26)$$

$N=21$ $R_{fit}=0.998$ $S_{fit}(N)=6.3$ $S_{fit}(N\text{-}I\text{-}1)=7.0$ $R_{cv}=0.996$ $S_{cv}(N)=8.6$ $S_{cv}(N\text{-}I\text{-}1)=9.4$ $F=1338$

(iii) The four-descriptor models

$$bp = -62\ (\pm 22) + 59.2\ (\pm 5.8)\ \varepsilon - 0.72\ (\pm 0.54)\ WS - 0.242\ (\pm 0.071)\ ws$$
$$+ 0.057\ (\pm 0.019)\ \omega \qquad (27)$$

$N=21$ $R_{fit}=0.999$ $S_{fit}(N)=5.0$ $S_{fit}(N\text{-}I\text{-}1)=5.7$ $R_{cv}=0.997$ $S_{cv}(N)=7.6$ $S_{cv}(N\text{-}I\text{-}1)=8.5$ $F=1518$

$$bp = -88\ (\pm 26) + 8.62\ (\pm 0.91)\ ^vM_2 - 1.77\ (\pm 0.68)\ WS - 0.207\ (\pm 0.073)\ ws$$
$$+ 0.051\ (\pm 0.020)\ \omega \qquad (28)$$

$N=21$ $R_{fit}=0.999$ $S_{fit}(N)=5.3$ $S_{fit}(N\text{-}I\text{-}1)=6.1$ $R_{cv}=0.997$ $S_{cv}(N)=7.7$ $S_{cv}(N\text{-}I\text{-}1)=8.7$ $F=1330$

The model (27) possesseses the lowest values of the standard errors of estimate, but the linear models based on the variable Zagreb M_2 index (10) and the variable connectivity index (18) are also very good models with the standard errors of estimate for the fit and cross-validated models of 6.8 (7.2) $^\circ$C and 8.0 (8.4) $^\circ$C, respectively.

We also considered the intercorrelation between the descriptors used in building up models (24) – (28). The intercorrelation matrix reflecting the pairwise linear correlation between $^vM_2$, ε, WS, ws and ω computed for 21 benzenoid is given in Table 2.

Table 2. The intercorrelation matrix

	$^vM_2$	ε	WS	ws	ω
$^vM_2$	1.000	0.999	0.946	0.981	0.973
ε		1.000	0.936	0.983	0.974
WS			1.000	0.912	0.920
ws				1.000	0.997
ω					1.000

The intercorrelation-degree is appraised by the correlation coefficient R. Pairs of indices with R ≥ 0.97 are regarded highly correlated, those with 0.90 ≤ R ≤ 0.97 appreciably correlated, those with 0.50 ≤ R ≤ 0.89 weakly correlated and the pairs of descriptors with low values of R (< 0.50) not correlated [37]. It appears that, according to the above classification, all the considered descriptors are either highly correlated ($^{v}M_2, \varepsilon$; $^{v}M_2, ws$; $^{v}M_2, \omega, \varepsilon$, ws; ε, ω, ws, ω) or appreciably correlated ($^{v}M_2, WS$; ε, WS; WS, ws; WS, ω). However, as Randić [*e.g.*, 38] pointed out, the intercorrelation criterion should not be always used for filtering descriptors to be used in building up the QSPR models.

Conclusions

The variable Zagreb M_2 index was used in the structure-boiling point modeling of benzenoid hydrocarbons. The obtained model is practically identical to the model based on the variable vertex-connectivity index and this is due to close relationship between the formulas for the two indices.

Comparative analysis of models based on several descriptors favoured the multivariate models with three and four descriptors. The best models with three and four descriptors did not include $^{v}M_2$ index. However, the next best three- and four-descriptor models contain the $^{v}M_2$ index. The best two-descriptor model contains $^{v}M_2$ index. The standard errors of estimate for the fit and cross-validated models listed in this report are in the 5.0 °C–9.4 °C range and this is a very good result since it shows that the boiling points of benzenoid hydrocarbons can be predicted within an error range of 0.8– 4.3%.

Acknowledgments.

References and Notes

1. Randić, M. Novel graph theoretical approach to heteroatoms in QSAR. *Chemometrics Intel. Lab. Syst.* **1991**, *10*, 213-227

2. Randić, M. On computation of optimal parameters for multivariate analysis of structure-property relationship. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 970-980

3. Randić, M.; Plavšić, D.; Lerš, N. Variable connectivity index for cycle-containing structures. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 657-662

4. Randić, M.; Basak, S.C. Multiple regression analysis with optimal molecular descriptors. *SAR QSAR Environ. Res.* **2000**, *11*, 1-23

5. Randić, M.; Pompe, M. The variable molecular descriptors based on the distance matrix. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 575-581

6. Randić, M.; Basak, S.C. On use of the variable connectivity index $^{1}\chi^{f}$ in QSAR: Toxicity of aliphatic ethers. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 614-618

7. Randić, M.; Pompe, M. The variable connectivity index $^{1}\chi^{f}$ versus traditional molecular descriptors: A comparative study of $^{1}\chi^{f}$ against descriptors of CODESSA. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 631-638

8. Nikolić, S.; Plavšić, D.; Trinajstić, N. On the Balaban-like topological indices. *MATCH − Comm. Math. Comput. Chem*. 2001, *44*, 361-386

9. Amić, D.; Lučić, B.; Nikolić, S.; Trinajstić, N. Predicting inhibition of microsomal *p*-hydroxylation of aniline by aliphatic alcohols: A QSAR approach based the weighted path numbers. *Croat. Chem. Acta* **2001**, *74*, 237-250

10. Amić, D.; Basak, S.C.; Lučić, B.; Nikolić, S.; Trinajstić, N. Structure-water solubility of aliphatic alcohols using the weighted path numbers. *SAR QSAR Environ. Res*. **2002**, *13*, 281-295

11. Lučić, B.; Miličević, A.; Nikolić, S.; Trinajstić, N. On variable Wiener index. *Ind. J. Chem.* **2003**, *42A*, 1279-1282

12. Miličević, A.; Nikolić, S. On variable Zagreb indices. *Croat. Chem. Acta* **2004**, *77*, 97-1001

13. Randić, M. Quantitative structure-property relationship. Boiling points of planar benzenoids. *New J. Chem*. **1996**, *20*, 1001-1009

14. Plavšić, D.; Trinajstić, N.; Amić, D.; Šoškić, M. Comparison between the structure-boiling point relationships with different descriptor for condensed benzenoids. *New J. Chem*. **1998**, *22*, 1075-1078

15. Harary, F. *Graph Theory, 2nd printing*; Addison-Wesley: Reading (MA, U.S.A.), **1971**

16. Trinajstić, N. *Chemical Graph Theory*, *2nd edition*; CRC Press: Boca Raton (FL, U.S.A.), **1992**

17. Gutman, I.; Trinajstić, N. Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons. *Chem. Phys. Lett*. **1972**, *17*, 535-537

18. Gutman, I.; Ruščić, B.; Trinajstić, N.; Wilcox, Jr., C.F. Graph theory and molecular orbitals. XII. Acyclic polyenes. *J. Chem. Phys*. **1975**, *62***,** 3399-3405

19. Devillers, J.; Balaban, A.T. Editors; *Topological Indices and Related Descriptors in QSAR and QSPR*; Gordon & Breach: Amsterdam (The Netherlands), **1999**

20. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Wiley-VCH: Weinheim (Germany) **2000**

21. Nikolić, S.; Kovačević, G.; Miličević, A.; Trinajstić, N. The Zagreb indices 30 years after. *Croat. Chem. Acta* **2003**, *76*, 113-124

22. Vukičević, D.; Trinajstić, N. Modified Zagreb M_2 index − Comparison with the Randić connectivity index for benzenoid systems. *Croat. Chem. Acta* **2003**, *76*, 183-187

23. Miličević, A.; Nikolić, S.; Trinajstić, N. On reformulated Zagreb indices. *Molecular Diversity* **2004**, in press

24. Clar, E. *Polycyclic Hydrocarbons*; Academic Press: London (U.K.) **1964**

25. Gutman, I.; Cyvin, S.J. *Introduction to the Theory Benzenoid Hydrocarbons*; Springer-Verlag: Berlin (Germany), **1989**

26. Lučić, B.; Trinajstić, N. Multivariate regression outperforms several robust architectures of neural networks. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 121-132

27. Lučić, B.; Trinajstić, N.; Sild, S.; Karelson, M.; Katritzky, A.R. A new efficient approach for variable selection based on multiregression: Prediction of gas chromatographic retention times and response factors. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 610-621

28. Lučić, B.; Amić, D.; Trinajstić, N. Nonlinear multivariate regression outperforms several concisely designed neural networks on three QSPR data sets. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 403-413

29. Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc*. **1975**, *97*, 6609-6615

30. Estrada, E. Edge adjacency relationships and a novel topological index related to molecular volume. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 31-33

31. Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc*. **1947**, *69*, 17-209

32. Amić, D.; Trinajstić. N. On the detour matrix. *Croat. Chem. Acta* **1995**, *68*, 53-62

33. Trinajstić. N.; Nikolić, S.; Lučić, B.; Amić, D.; Mihalić, Z. The detour matrix in chemistry. *J. Chem. Inf. Comput. Sci*. **1997,** *37*, 631-638

34. Mihalić, Z.; Veljan, D.; Amić, D.; Nikolić, S.; Plavšić, D.; Trinajstić. N. The distance matrix in chemistry. *J.Math. Chem*. **1992,** *11*, 223-258

35. Randić, M. On characterization of cyclic structures. *J. Chem. Inf. Comput. Sci*. **1997,** *37*, 1063-1071

36. Lukovits, I. The detour index. *Croat. Chem. Acta* **1996**, *69*, 873-883

37. Trinajstić, N.; Nikolić, S.; Basak, S.C.; Lukovits, I. Distances indices and their hypercounterparts: Intercorrelation and use in the structure-property modeling. *SAR QSAR Environ. Res*. **2001**, *12*, 31-54

38. Randić, M. The connectivity index 25 years after. *J. Mol. Graph. Model*. **2001**, *20*, 19-35