NEW SPECTRAL INDICES FOR MOLECULE DESCRIPTION

Viviana Consonni* and Roberto Todeschini

Milano Chemometrics and QSAR Research Group – Dept. of Environmental Sciences, University of Milano-Bicocca, P.za della Scienza, 1 – 20126 Milano (Italy)

Two families of molecular descriptors are proposed, based on new functions of the molecule spectrum derived from any molecular matrix $\mathbf{M}(w)$, calculated by any weighting scheme w. The first class of molecular descriptors, denoted as $SpAD(\mathbf{M},w)$, is defined in terms of the absolute deviations of the eigenvalues and is a generalization of the concept of graph energy proposed by I. Gutman [1]. The second class, denoted as $SpMAD(\mathbf{M},w)$, is defined in terms of mean absolute deviations of the eigenvalues and, unlike the first class, is size-independent.

Some simple QSPR models were calculated, showing an encouraging capability of these descriptors to model physico-chemical properties of compounds.

* Corresponding author e-mail: viviana.consonni@unimib.it

1. Introduction

Molecular descriptors are used in mathematical chemistry for the design of chemical compounds with given physico-chemical/pharmacological/biological properties or for assessing their toxicological or environmental properties.

Spectral indices are among the huge number of molecular descriptors proposed in the literature [2]. They were originally used in the framework of the molecular orbital theory to study unsatured compounds [3; 4]. Then, they were generalized to study any compound, finding a lot of applications in modeling physico-chemical properties of molecules by QSAR/QSPR analysis [5].

Spectral indices are molecular descriptors defined in terms of the eigenvalues of a square graph-theoretical matrix $\mathbf{M}$ of size (n, n), n usually being the number of the molecule atoms or bonds [6-9]. The eigenvalues are the roots of the characteristic polynomial of the matrix $\mathbf{M}$ and the set of the eigenvalues is the matrix spectrum $\Lambda(\mathbf{M}) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Spectral indices can be the single eigenvalues or some function of them; they are derived from a two-dimensional representation of molecules (i.e. molecular graphs) or from a three-dimensional representation (i.e molecular geometries).

The most common eigenvalue functions used to derive spectral indices are given below in a general form which can be applied to any molecular matrix M(w), calculated with the weighting scheme w [10; 11]:

$$SpSum^k(\mathbf{M}, w) = \sum_{i=1}^{n} |\lambda_i|^k \qquad SpSum_+^k(\mathbf{M}, w) = \sum_{i=1}^{n^+} (\lambda_i^+)^k \qquad SpSum_-^k(\mathbf{M}, w) = \sum_{i=1}^{n^-} |\lambda_i^-|^k$$

$$MinSp(\mathbf{M}, w) = \min_i \{\lambda_i\} \qquad MaxSp(\mathbf{M}, w) = \max_i \{\lambda_i\}$$

$$MaxSpA(\mathbf{M}, w) = \max_i \{|\lambda_i|\} \qquad SpDiam(\mathbf{M}, w) = MaxSp - MinSp$$

where k is a parameter, usually taken equal to one; for negative k values, eigenvalues equal to zero must not be considered. For $k = 1$, $SpSum$ is the sum of the n absolute values of the spectrum eigenvalues; this quantity calculated on the adjacency matrix of simple graphs was called *graph energy* [1; 12; 13], the same quantity derived from the Laplacian matrix was called *Laplacian graph energy* [14; 15]. $SpSum_+$ is the sum of the n^+ positive eigenvalues, $SpSum_-$ is the sum of the absolute values of the n^- negative eigenvalues, $MinSp$ is the minimum eigenvalue, $MaxSp$ is the maximum eigenvalue, $MaxSpA$ is the maximum absolute value of the spectrum, and $SpDiam$ is the spectral diameter of the molecular matrix. Different weighting schemes w can be applied to characterize vertices or edges or both in order to better describe organic compounds containing heteroatoms and/or multiple bonds.

In Table 1, the most popular spectral indices derived from graph-theoretical matrices are collected.

Other spectral descriptors are the Burden eigenvalues derived from a modified connectivity matrix [28], the leading eigenvalue of the matrix calculated as the sum of adjacency and distance matrices [29], and eigenvalues derived from the path matrix [30].

Based on the 3D-geometrical representation of molecules, eigenvalue-based decriptors are WHIM descriptors [31], EVA descriptors [32] and EEVA descriptors [33].

In this paper, two new general spectral functions are proposed by which several molecular descriptors can be derived both from graph-theoretical matrices and matrices based on molecule geometry.

Table 1. Some spectral indices defined in the literature. A is the number of vertices (atoms) of the molecular graph, λ the graph eigenvalues, and n^+ the number of positive eigenvalues.

Descriptor name	Formula	Graph-theoretical matrix	Ref.
Lovasz-Pelikan index	λ_1	adjacency matrix	[16]
graph energy	$\sum_{i=1}^{A}\left|\lambda_i\right|$	adjacency matrix	[1]
VAA1 index	$\sum_{i=1}^{n^+}\lambda_i$	adjacency matrix	[17]
quasi-Wiener index	$A\cdot\sum_{i=1}^{A-1}\dfrac{1}{\lambda_i}$	Laplace matrix	[18; 19]
second Mohar index	$\dfrac{4}{A\cdot\lambda_{A-1}}$	Laplace matrix	[20]
characteristic root index	$\sum_{i=1}^{n^+}\lambda_i$	χ matrix	[21; 22]
leading eigenvalue of the distance matrix	λ_1	distance matrix	[23]
folding degree index	$\dfrac{\lambda_1}{A}$	geometric distance / topological distance quotient matrix	[24; 25]
VAD1 index	λ_A	distance matrix	[17]
Estrada index	$\sum_{i=1}^{A}e^{\lambda_i}$	any matrix $\mathbf{M}$	[26; 27]

2. Theory

Let $\mathbf{M}(w)$ be a molecular matrix (n, n), calculated with the weighting scheme w, and $\{\lambda_i\}$ the set of the corresponding eigenvalues. The *SpAD* function is here proposed as the sum of the absolute deviation of each eigenvalue from the average eigenvalue $\bar{\lambda}$:

$$SpAD\left(\mathbf{M}, w\right) = \sum_{i=1}^{n}\left|\lambda_i - \bar{\lambda}\right|$$

where n is the size of the considered molecular matrix, e.g. the number of graph vertices (atoms). Spectral indices based on this function obviously depend on the size of the matrix, i.e. on the molecule size.

Therefore, a second spectral function is also proposed as mean absolute deviation (SpMAD) with the aim of eliminating the direct dependence of SpAD on the molecule size:

$$SpMAD\left(\mathbf{M}, w\right) = \frac{\sum_{i=1}^{n}\left|\lambda_i - \bar{\lambda}\right|}{n}$$

The two functions can be applied to any square graph-theoretical matrix, such as adjacency matrix A, vertex distance matrix D, detour matrix Δ, resistance matrix Ω, Laplace matrix L, quotient matrices, Barisz and Barysz-like matrices [34]. Moreover, they can be applied to the weighted covariance matrix $(3, 3)$ of the atom spatial coordinates, as defined in the WHIM descriptor theory [31]. Diverse spectral indices are

obtained depending on the chosen molecular matrix and the weighting scheme applied to describe molecules. For isomeric series of molecules, SpAD and SpMAD indices calculated from any graph matrix M differ by a constant and are correlated one.

It is noteworthy that the SpAD function applied to the adjacency matrix A is exactly the graph enegy proposed by I. Gutman [1]. In effect, the average eigenvalue of the adjacency matrix is equal to zero and thus:

$$SpAD(\mathbf{A}) = \sum_{i=1}^{A} \left| \lambda_i - \overline{\lambda} \right| = \sum_{i=1}^{A} \left| \lambda_i \right|$$

where A is the number of graph vertices.

For the same reason, $SpAD$ values obtained from graph-theoretical matrices whose diagonal elements are zero, as it usually is the case of matrices from unweighted molecular graphs, coincide with the corresponding graph energies.

3. QSPR models

When new molecular descriptors are going to be proposed, one of the rules for their acceptance is that they are required to model some properties of molecules. Thus, to evaluate the potential of the new spectral indices in QSAR/QSPR analysis, a number of them were calculated and their modelling power evaluated by using two benchmark data sets.

As the goal of this paper was to obtain a first rough estimate of the predictive capability of the new descriptors, only univariate models were calculated. All of the calculations were performed by the statistical package MobyDigs [35]. The best correlations, estimated by Multiple Linear Regression (MLR), were chosen by using the leave-one-out procedure of cross-validation (Q^2) and checked by boostrap validation technique (Q^2_{BOOT}). Together with these modelling parameters, the fitting power (R^2) and the standard deviation error in prediction (SDEP) were also reported.

3.1 Molecular descriptors

The molecular descriptors used in QSPR analysis are spectral indices $SpAD$ and $SpMAD$ derived from some molecular matrices defined in Table 2. Most of these matrices are derived from unweighted molecular graphs; in this case, the symbol of spectral indices contains only the matrix they are derived from. Barisz and Burden matrices are, on the contrary, derived from weighted molecular graphs. The applied weighting schemes are based on atomic mass (m), atomic van der Waals volume (v), Sanderson electronegativity (e), and atomic polarizability (p). The weighting scheme based on the atomic weight Z was only applied to derive the Barisz matrix. Moreover, the Burden matrix was also derived from unweighted graphs. The WHIM covariance matrix was calculated from the 3D (xyz)-spatial atomic coordinates both on the unweighted and weighted molecular geometry.

Table 2. List of the new descriptors studied in this work. The symbol No. represents the total number of spectral indices derived from each matrix; the notation w indicates that descriptors are calculated by using different weighting schemes; D distinguishes between 2D and 3D descriptors.

Molecular descriptor	Matrix	No.	D	Information
$SpAD(\mathbf{A})$ and $SpMAD(\mathbf{A})$	adjacency matrix $\mathbf{A}$	2	2D	topostructural
$SpAD(\mathbf{D})$ and $SpMAD(\mathbf{D})$	distance matrix $\mathbf{D}$	2	2D	topostructural
$SpAD(\mathbf{L})$ and $SpMAD(\mathbf{L})$	Laplacian matrix $\mathbf{L}$	2	2D	topostructural
$SpAD(\mathbf{D/\Delta})$ and $SpMAD(\mathbf{D/\Delta})$	distance/detour quotient matrix $\mathbf{D/\Delta}$	2	2D	topostructural
$SpAD(\mathbf{D^Z}, w)$ and $SpMAD(\mathbf{D^Z}, w)$	Barysz matrices $\mathbf{D^Z}(w)$	10	2D	topochemical
$SpAD(\mathbf{B}, w)$ and $SpMAD(\mathbf{B}, w)$	Burden matrices $\mathbf{B}(w)$	10	2D	topochemical
$SpAD(\mathbf{W}, w)$ and $SpMAD(\mathbf{W}, w)$	weighted WHIM covar. matrices $\mathbf{W}(w)$	10	3D	geometrical
$SpAD(\mathbf{G/D})$ and $SpMAD(\mathbf{G/D})$	distance/distance quotient matrix $\mathbf{G/D}$	2	3D	geometrical

For comparison purposes, some other well-known topological descriptors were also calculated and evaluated in the QSPR analysis; they are reported in Table 3. All the descriptors were calculated by the DRAGON package [36] on the basis of the minimum energy molecular geometries optimized by HyperChem package (PM3 semiempirical method).

Table 3. List of additional descriptors used for model comparison.

Descriptor	Definition	Ref.
W	Wiener index	[37]
J	Balaban index	[38]
χ_1	Randic connectivity index	[39]
χ_2	second order connectivity index	[40]
M_1	first Zagreb index	[41]
M_2	second Zagreb index	[41]
SMTI	Schultz molecular topological index	[42]
GMTI	Gutman molecular topological index	[43]
H	Harary index	[44]
TI_1	first Mohar index	[20]
TI_2	second Mohar index	[20]
W^*	quasi-Wiener index	[19]
$SpMax(\mathbf{M})$	leading eigenvalue ($\mathbf{M} = \mathbf{A}, \mathbf{D}, \mathbf{L}, \mathbf{D/\Delta}, \mathbf{G/D}, \mathbf{D^Z}(w), \mathbf{B}(w)$)	

3.2 Data sets

The present study was performed using two diverse data sets taken from the literature: 1) the octane isomer data set, which is a well-known benchmark data set, and 2) the PCB (polychlorobiphenyls) data set, comprised of 209 compounds.

Comprehensive studies of numerous physico-chemical properties of the 18 octane isomers have been already published based on the idea that the properties of a series of isomeric compounds, being independent of molecular size, are suitable for the first evaluation of the modeling capabilities of different mathematical descriptors. In this paper, models for melting point (MP, 14 data), motor octane number (MON, 16 data) and entropy (S, 18 data) were evaluated [45; 46].

The second data set was chosen because, unlike the octane isomers, PCB are organic compounds containing heteroatoms and a different number of non-hydrogen atoms, thus being well described by weighted molecular graphs. For this data set, melting point (MP, 81 data) and water solubility (logSw, 136 data) were modelled by the selected molecular descriptors [47].

3.3 Results

A total number of 32 molecular descriptors were calculated for 18 octane isomers and 67 for 209 PCBs.

These molecular descriptors include both the new proposed spectral indices and some other well-known topological descriptors, together with an extended set of leading eigenvalues.

The greater number of molecular descriptors for PCBs is due to the fact that different weighting schemes were applied since PCBs, unlike octane isomers, contain heteroatoms and multiple bonds. Moreover, descriptors from the Barisz matrix were not calculated for octane isomers since the Barisz matrix coincides with the distance matrix in the case of saturated hydrocarbons.

Three physico-chemical properties were modelled for octanes and two for PCBs. For each property, its correlation with each single molecular descriptor was evaluated by different statistical parameters. All the correlations were ranked according to the leave-one-out Q^2, i.e. the predictive power, and the 20 best correlations were reported in the paper. Some correlations, based on spectral indices calculated from the same matrix but using a different weighting scheme, were omitted when all the statistical parameters were the same.

In Table 4, 5 and 6 the best 20 univariate models for melting point (Table 4), motor octane number (Table 5) and entropy (Table 6) for the octane isomer are collected.

In the case of octane isomers, the good modelling power of the graph energy $SpAD(\mathbf{A})$ was confirmed [48]; in effect, this descriptor gives the best correlation with the melting point. More in general, graph energies calculated from adjacency, distance and Laplace matrices gave good results for the studied properties.

It must be also noted that for melting point only the graph energy and the first two connectivity indices showed some modeling power.

Table 4 – The 20 best univariate models for the melting point of the octane isomers

ID	Models	Q^2	$Q^2 boot$	R^2	SDEP
1	$SpAD(\mathbf{A})$	67.1	70.1	77.7	3.409
2	χ_2	64.6	68.5	77.6	3.538
3	χ_1	53.0	56.9	67.3	4.078
4	M_1	37.2	40.0	51.9	4.712
5	$SpAD(\mathbf{B})$	30.0	34.6	46.0	4.974
6	$SpAD(\mathbf{L})$	29.1	30.9	42.3	5.005
7	$MaxSp(\mathbf{L})$	19.4	23.8	36.3	5.337
8	$MaxSp(\mathbf{G/D})$	16.8	20.9	33.5	5.423
9	$MaxSp(\mathbf{W})$	15.8	20.6	31.6	5.456
10	$MaxSp(\mathbf{D})$	15.4	20.6	31.2	5.467
11	$SpAD(\mathbf{D})$	15.4	21.2	31.2	5.467
12	$SpAD(\mathbf{W})$	15.2	20.1	31.2	5.473
13	$MaxSp(\mathbf{A})$	14.1	18.9	30.8	5.509
14	$MaxSp(\mathbf{B})$	12.7	17.4	29.4	5.556
15	W	12.5	17.5	29.2	5.560
16	GMTI	12.5	17.2	29.2	5.560
17	TI_1	12.5	17.7	29.2	5.560
18	M_2	8.7	12.4	25.1	5.681
19	SMTI	8.0	13.4	26.0	5.701
20	J	5.6	9.9	23.6	5.777

Table 5 – The 20 best univariate models for the motor octane number of the octane isomers

ID	Models	Q2	Q2boot	R2	SDEP
1	$SpAD(\mathbf{D})$	90.5	90.9	92.8	7.395
2	$MaxSp(\mathbf{D})$	90.5	90.7	92.8	7.395
3	TI_1	89.1	89.5	91.6	7.930
4	W	89.1	89.5	91.6	7.930
5	GMTI	89.1	89.3	91.6	7.930
6	SMTI	87.0	87.6	89.8	8.653
7	$MaxSp(\mathbf{W})$	86.4	87.4	88.8	8.847
8	$SpAD(\mathbf{W})$	84.9	85.8	87.6	9.321
9	$MaxSp(\mathbf{B})$	83.8	84.2	88.0	9.672
10	$MaxSp(\mathbf{G/D})$	82.5	83.5	86.0	10.043
11	J	81.2	81.6	86.0	10.405
12	$MaxSp(\mathbf{A})$	80.5	81.1	85.4	10.591
13	M_2	77.2	77.1	83.9	11.449
14	M_1	71.4	72.9	77.7	12.838
15	$MaxSp(\mathbf{L})$	63.7	65.5	71.7	14.449
16	$SpMAD(\mathbf{L})$	58.6	60.1	67.4	15.444
17	TI_2	55.1	58.1	66.4	16.084
18	χ_1	50.2	53.5	60.4	16.933
19	$SpMAD(\mathbf{B})$	43.5	46.6	54.7	18.033
20	χ_2	21.9	27.3	36.9	21.209

Table 6 – The 20 best univariate models for the entropy of the octane isomers

ID	Models	Q2	Q2boot	R2	SDEP
1	M_1	87.3	88.3	91.1	1.563
2	M_2	83.9	85.4	88.7	1.759
3	*SpAD*(**L**)	79.3	81.1	84.9	1.997
4	*MaxSp*(**A**)	75.9	77.5	83.9	2.156
5	*MaxSp*(**B**)	75.4	77.2	83.4	2.180
6	χ_1	74.8	77.0	82.1	2.203
7	J	74.5	76.3	82.2	2.216
8	*MaxSp*(**L**)	71.9	73.9	80.6	2.327
9	*MaxSp*(**D**)	68.1	70.0	78.6	2.481
10	*SpAD*(**D**)	68.1	70.0	78.6	2.481
11	W	66.2	68.2	77.1	2.553
12	GMTI	66.2	68.1	77.1	2.553
13	TI_1	66.2	68.4	77.1	2.553
14	SMTI	61.6	63.5	73.3	2.720
15	*SpAD*(**B**)	61.5	64.2	72.3	2.725
16	χ_2	54.0	58.8	68.7	2.979
17	*SpAD*(**W**)	52.0	54.6	67.4	3.042
18	*MaxSp*(**W**)	51.1	53.3	66.9	3.070
19	*MaxSp*(**G/D**)	47.7	50.5	61.2	3.176
20	TI_2	44.7	47.9	57.4	3.265

Concerning the motor octane number, several descriptors gave a good predictive power. In particular, the new descriptor based on the distance matrix, together with the corresponding leading eigenvalue, gave the best result. The first Mohar index, Wiener index and Gutman molecular topological index also showed a good predictive behaviour for motor octane number.

Concerning the entropy of the octane isomers, the first Zagreb index M_1 was the best modelling descriptor ($Q^2 = 87.3$); however, it can be noted that also the second Zagreb index M_2 as well as *SpAD*(**L**) calculated on the Laplace matrix gave acceptable results ($Q^2 = 83.9$ and $Q^2 = 79.3$, respectively).

In Table 7 and 8, the best 20 univariate models for melting point (Table 7) and water solubility (Table 8) for the PCB compounds are collected.

Concerning the first property, the predictive power of the new proposed descriptors (from Barisz and distance matrices) resulted good enough, if compared with the best descriptors, i.e. the leading eigenvalues from the distance/distance quotient matrix (**G/D**) and the distance/detour quotient matrix (**D/Δ**).

In the case of the PCB water solubility, several descriptors performed well and once again the spectral indices appeared among the best predictive descriptors. For both properties, because of the congeneric data set, several *MaxSp*, *SpAD* and *SpMAD* descriptors based on the Barisz, Burden, and WHIM matrices gave similar results for different weighting schemes and were not reported in the tables.

4. Conclusions

Two new classes of molecular descriptors were proposed, defined in terms of functions of the spectrum of a molecular matrix. They are founded on a simple theoretical basis and are closely related to some well-

known indices used in graph theory to describe molecular structures. In particular, the *SpAD* function is an extension of the graph energy to any molecular matrix having an average eigenvalue different from zero, as it often is the case of matrices derived from weighted molecular graphs.

Table 7 – The 20 best univariate models for the melting point of the PCB compounds

ID	Models	Q2	Q2boot	R2	SDEP
1	*MaxSP*(**G/D**)	63.8	64.1	66.3	30.313
2	*MaxSP*(**D/Δ**)	63.4	63.6	65.9	30.471
3	*SpMAD*(**D^Z**, m)	63.1	63.2	65.6	30.595
4	*MaxSp*(**D^Z**, m)	63.0	63.3	65.6	30.620
5	GMTI	62.9	63.2	65.4	30.685
6	*SpAD*(**D**)	62.7	63.0	65.3	30.756
7	*MaxSp*(**D**)	62.7	63.2	65.3	30.756
8	SMTI	62.5	63.0	65.0	30.834
9	W	62.4	62.7	64.9	30.894
10	*SpMAD*(**D**)	61.3	61.8	63.9	31.313
11	*SpAD*(**A**)	60.4	60.8	63.1	31.703
12	*SpMAD*(**D/Δ**)	60.0	60.1	62.8	31.867
13	*SpAD*(**A**)	59.9	60.3	62.7	31.880
14	*SpAD*(**A**)	58.4	59.2	61.2	32.480
15	M_1	58.0	58.3	60.8	32.651
16	χ_1	57.8	58.0	60.7	32.724
17	*SpAD*(**L**)	57.7	58.1	60.5	32.769
18	M_2	57.3	57.7	60.2	32.924
19	χ_2	57.2	57.5	60.1	32.936
20	*SpMAD*(**D^Z**, Z)	56.5	56.9	59.2	33.221

Table 8 – The 20 best univariate models for the water solubility of the PCB compounds

ID	Models	Q2	Q2boot	R2	SDEP
1	*MaxSp*(**D/Δ**)	94.3	94.3	94.4	0.272
2	*MaxSp*(**D**)	94.2	94.3	94.4	0.272
3	*SpMAD*(**D**)	94.2	94.3	94.4	0.272
4	*MaxSp*(**D^Z**, Z)	94.2	94.2	94.4	0.273
5	*SpAD*(**D^Z**, Z)	94.2	94.2	94.4	0.273
6	*MaxSp*(**G/D**)	94.0	94.1	94.2	0.276
7	GMTI	93.8	93.9	94.0	0.282
8	SMTI	93.7	93.7	93.8	0.285
9	W	93.5	93.5	93.7	0.289
10	TI_1	93.4	93.5	93.6	0.290
11	*SpAD*(**D/Δ**)	92.9	93.0	93.1	0.301
12	χ_2	92.7	92.7	92.9	0.307
13	*SpAD*(**L**)	92.6	92.7	92.8	0.308
14	M_1	92.2	92.3	92.4	0.316
15	*SpMAD*(**D/Δ**)	92.1	92.2	92.4	0.318
16	χ_1	91.8	91.9	92.1	0.323
17	M_2	91.0	91.0	91.2	0.341
18	*SpMAD*(**D^Z**, v)	90.3	90.4	90.5	0.354
19	*SpAD*(**A**)	89.2	89.2	89.5	0.373
20	*SpMAD*(**D^Z**, p)	88.5	88.6	88.8	0.384

From the calculated QSPR models, it is apparent that the new descriptors contain some good structural information for property modelling since some of them appear in the best correlations of the modelled properties.

Multivariate models and other properties should be furtherly studied more in details.

References

1. Gutman I. (1978). The energy of a graph. *Ber. Math. Statist. Sekt. Forschungszentrum Graz*, *103*, 1-22.

2. Todeschini R. and Consonni V. (2000). *Handbook of Molecular Descriptors*. Wiley-VCH, Weinheim (Germany), 668 pp.

3. Graovac A., Gutman I. and Trinajstic N. (1977). *Topological Approach to the Chemistry of Conjugated Molecules*. Springer-Verlag, Berlin (Germany), 123 pp.

4. Trinajstic N. (1992). *Chemical Graph Theory*. CRC Press, Boca Raton (FL), 322 pp.

5. Ivanciuc O., Ivanciuc T. and Diudea M. V. (1999). Polynomials and Spectra of Molecular Graphs. *Roum. Chem. Quart. Rev.*, *7*, 41-67.

6. Hall G. G. (1981). Eigenvalues of Molecular Graphs. *Bull. Inst. Math. Appl.*, *17*, 70-72.

7. Hall G. G. (1992). Eigenvalue Distributions for the Graphs of Alternant Hydrocarbons. *J. Chem. Inf. Comput. Sci.*, *32*, 11-13.

8. Hall G. G. (1993). Eigenvalue Distributions in Alternant Hydrocarbons. *J. Math. Chem.*, *13*, 191-203.

9. Lee Shyi-Long and Yeh Yeong-Nan (1993). On Eigenvalues and Eigenvectors of Graphs. *J. Math. Chem.*, *12*, 121-135.

10. Ivanciuc O. and Balaban A. T. (1999). The Graph Description of Chemical Structures. In *Topological Indices and Related Descriptors in QSAR and QSPR.* (Devillers J. and Balaban A. T., Eds.), Gordon and Breach Science Publishers, Amsterdam (The Netherlands), pp. 59-167.

11. Ivanciuc O., Ivanciuc T. and Balaban A. T. (1999). Vertex- and Edge-Weighted Molecular Graphs and Derived Structural Descriptors. In *Topological Indices and Related Descriptors in QSAR and QSPR.* (Devillers J. and Balaban A. T., Eds.), Gordon and Breach Science Publishers, Amsterdam (The Netherlands), pp. 169-220.

12. Gutman I. (2001). The energy of a graph: Old and new results. In *Algebraic Combinatorics and Applications.* (Betten A., Kohnert A., Laue R. and Wassermann A., Eds.), Springer-Verlag, Berlin (Germany), pp. 196-211.

13. Gutman I. (2005). Topology and stability of conjugated hydrocarbons. The dependence of total π-electron energy on molecular topology. *J. Serb. Chem. Soc.*, *70*, 441-456.

14. Gutman I. and Zhou Bo (2006). Laplacian energy of a graph. *Linear Algebra and Its Applications*, *414*, 29-37.

15. Zhou Bo and Gutman I. (2007). On Laplacian Energy of Graphs. *MATCH Commun. Math. Comput. Chem.*, *57*, 211-220.

16. Lovasz L. and Pelikan J. (1973). On the Eigenvalue of Trees. *Period. Math. Hung.*, *3*, 175-182.

17. Balaban A. T., Ciubotariu D. and Medeleanu M. (1991). Topological Indices and Real Vertex Invariants Based on Graph Eigenvalues or Eigenvectors. *J. Chem. Inf. Comput. Sci.*, *31*, 517-523.

18. Gutman I., Yeh Yeong-Nan, Lee Shyi-Long and Luo Y. L. (1993). Some Recent Results in the Theory of the Wiener Number. *Indian J. Chem.*, *32A*, 651-661.

19. Mohar B., Babic D. and Trinajstic N. (1993). A Novel Definition of the Wiener Index for Trees. *J. Chem. Inf. Comput. Sci.*, *33*, 153-154.

20. Trinajstic N., Babic D., Nikolic S., Plavšic D., Amic D. and Mihalic Z. (1994). The Laplacian Matrix in Chemistry. *J. Chem. Inf. Comput. Sci.*, *34*, 368-376.

21. Saçan M. T. and Inel Y. (1993). Prediction of Aqueous Solubility of PCBs Related to Molecular Structure. *Tr. J. Chemistry*, *17*, 188-195.

22. Saçan M. T. and Inel Y. (1995). Application of the Characteristic Root Index Model to the Estimation of N-Octanol Water Partition Coefficients. Polychlorinated Biphenyls. *Chemosphere*, *30*, 39-50.

23. Schultz H. P., Schultz E. B. and Schultz T. P. (1990). Topological Organic Chemistry. 2. Graph Theory, Matrix Determinants and Eigenvalues, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.*, *30*, 27-29.

24. Randic M., Kleiner A. F. and DeAlba L. M. (1994). Distance/Distance Matrices. *J. Chem. Inf. Comput. Sci.*, *34*, 277-286.

25. Randic M. and Krilov G. (1999). On a Characterization of the Folding of Proteins. *Int. J. Quant. Chem.*, *75*, 1017-1026.

26. Estrada E. and Rodríguez-Velásquez J. A. (2005). Subgraph centrality in complex networks. *Phys. Rev. E*, *71*, ?

27. Estrada E. and Hatano N. (2007). Statistical-mechanical approach to subgraph centrality in complex networks. *Chem. Phys. Lett.*, *439*, 247-251.

28. Burden F. R. (1997). A Chemically Intuitive Molecular Index Based on the Eigenvalues of a Modified Adjacency Matrix. *Quant. Struct. -Act. Relat.*, *16*, 309-314.

29. Cash G. G. (1995). Heats of Formation of Polyhex Polycyclic Aromatic Hydrocarbons from their Adjacency Matrices. *J. Chem. Inf. Comput. Sci.*, *35*, 815-818.

30. Randic M. (1997). On Molecular Branching. *Acta Chim. Sloven.*, *44*, 57-77.

31. Todeschini R., Lasagni M. and Marengo E. (1994). New Molecular Descriptors for 2D- and 3D-Structures. Theory. *J. Chemom.*, *8*, 263-273.

32. Ferguson A. M., Heritage T. W., Jonathon P., Pack S. E., Phillips L., Rogan J. and Snaith P. J. (1997). EVA: A New Theoretically Based Molecular Descriptor for Use in QSAR\QSPR Analysis. *J. Comput. Aid. Molec. Des.*, *11*, 143-152.

33. Tuppurainen K. (1999). EEVA (Electronic Eigenvalue): A New QSAR/QSPR Descriptor for Electronic Substituent Effects Based on Molecular Orbital Energies. *SAR & QSAR Environ. Res.*, *10*, 39-46.

34. Janežic D., Milicevic A., Nikolic S. and Trinajstic N. (2007). *Graph Theoretical Matrices in Chemistry*. University of Kragujevac, Kragujevac (Serbia), 205 pp.

35. Todeschini R., Consonni V., Mauri A. and Pavan M. (2003). MOBYDIGS: Software for Regression and Classification Models by Genetic Algorithms. In *Chemometrics: Genetic Algorithms and Artificial Neural Networks.* (Leardi R., Ed.), Elsevier, Amsterdam (The Netherlands).

36. DRAGON - Software for the calculation of molecular descriptors 5.5. Talete S.r.l., Via V.Pisani 13, Milano (Italy).

37. Wiener H. (1947). Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.*, *69*, 17-20.

38. Balaban A. T. (1982). Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett.*, *89*, 399-404.

39. Randic M. (1975). On Characterization of Molecular Branching. *J. Am. Chem. Soc.*, *97*, 6609-6615.

40. Kier L. B. and Hall L. H. (1977). The Nature of Structure-Acitivity Relationships and their Relation to Molecular Connectivity. *Eur. J. Med. Chem.*, *12*, 307-312.

41. Gutman I., Ruscic B., Trinajstic N. and Wilcox Jr C. F. (1975). Graph Theory and Molecular Orbitals. XII. Acyclic Polyenes. *J. Chem. Phys.*, *62*, 3399-3405.

42. Schultz H. P. (1989). Topological Organic Chemistry. 1. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.*, *29*, 227-228.

43. Gutman I. (1994). Selected Properties of the Schultz Molecular Topological Index. *J. Chem. Inf. Comput. Sci.*, *34*, 1087-1089.

44. Plavšic D., Nikolic S., Trinajstic N. and Mihalic Z. (1993). On the Harary Index for the Characterization of Chemical Graphs. *J. Math. Chem.*, *12*, 235-250.

45. Randic M., Guo Xiaofeng, Oxley T., Krishnapriyan H. and Naylor L. (1994). Wiener Matrix Invariants. *J. Chem. Inf. Comput. Sci.*, *34*, 361-367.

46. Diudea M. V., Minailiuc O. M. and Katona G. (1997). Molecular Topology. 26. SP Indices: Novel Connectivity Descriptors. *Rev. Roum. Chim.*, *42*, 239-249.

47. Gramatica P., Navas N. and Todeschini R. (1998). 3D-Modelling and Prediction by WHIM Descriptors. Part 9. Chromatographic Relative Retention Time and Physico-Chemical Properties of Polychlorinated Biphenyls (PCBs). *Chemom. Intell. Lab. Syst.*, *40* , 53-63.

48. Gutman I., Vidovic D., Cmiljanovic N., Milosavljevic S. and Radenkovic S. (2003). Graph energy - A useful molecular structure-descriptor. *Indian J. Chem.*, *42A*, 1309-1311.