Scientific paper

LEL-a Newly Designed Molecular Descriptor

Dragan Stevanović,[a,b] Aleksandar Ilić,[a] Cristina Onişor[c] and Mircea V. Diudea[c,*]

[a] *Faculty of Sciences and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia*

[b] *University of Primorska - FAMNIT, Glagoljaška 8, 6000 Koper, Slovenia*

[c] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, Arany Janos Str. 11, 400084, Cluj, Romania*

** Corresponding author: E-mail: diudea@gmail.com*

Received: 28-11-2008

Abstract

A correlating study of topological indices TIs provided by TOPOCLUJ software package and LEL, a newly proposed index built up on the eigenvalues of Laplacian matrix, on thirteen properties of octanes, revealed good correlating ability of a dozen selected TIs, all related to the Wiener index, and of LEL as well. LEL describes well the properties which are well accounted by the majority of the selected molecular descriptors: octane number MON, entropy S, volume MV, or refraction MR, particularly the acentric factor AF parameter, but also more difficult properties like boiling point, melting point and logP. LEL is the best correlated with the WK (Wiener-type number, taken the reciprocal of entries in the combinatorial D_p matrix, of higher rank, calculated by TOPOCLUJ software) indices. In a second set of polycyclic aromatic hydrocarbons, LEL was proved to be as good as the Randić χ index (a connectivity index) and better than the Wiener index (a distance based index). In addition, it is well defined mathematically and shows interesting relations in particular classes of graphs, these recommending LEL as a new, powerful topological index. The actual study proved the considered TIs are basic topological descriptors in prediction of various molecular properties, with good perspective in QSPR/QSAR studies.

Keywords*:* Topological indices, QSPR, QSAR, octanes, polycyclic aromatic hydrocarbons

1. Introduction

Elucidation of the relationship between molecular structures and their properties is a challenge to chemists for more than a century. A chemical structure can be quantified in various ways, one of the most popular, in the last decade being that which makes use of Graph Theory, particularly topological indices TIs, which are numerical descriptors encoding topological attributes of a molecular graph. They were used both in graph discriminating analysis as well as in quantitative structure-property relationship QSPR studies for modeling a variety of physico-chemical properties or in quantitative structure-activity relationship QSAR studies to predict biological activities. Nowadays, their number becomes uncountable, as a consequence of the explosive development of computational technology. Randić[1] has outlined some desirable attributes for the topological indices in the view of preventing their hazardous proliferation, among which: direct structural interpretation, good correlation with at least one property, good discrimination of isomers, simplicity, locally defined, and generalized to higher analogues, are the most important. There are several commercial software packages which calculate more than one thousand TIs: POLLY,[2] MOLCONN,[3] CODESSA,[4,5] DRAGON [6] and TOPOCLUJ.[7]

Alkanes represent an interesting class of compounds as a starting point for the application of molecular modeling procedures. Many properties of the alkanes vary function of molecular mass or branching, and alkanes can be described by using a single type of (carbon) atom. There are properties well accounted by a single molecular descriptor, e.g., octane number MON, entropy S, volume MV, refraction MR, etc. Other properties, such as, boiling point BP, heat of vaporization HV, total surface area TSA, partition coefficient LogP, density DENS, critical temperature CT, critical pressure CP and heat of formation DHF are notable exceptions, being not well modeled by any of the parameter sets.

Among the well modeled parameters, the acentric factor AF appears to be correlated more than 90% by eleven of twelve descriptors herein discussed. Definition of the best scored TIs, e.g., LEL, 2WD, 2WH, 1WK, 2WUCJD and PDS3[Sh[SZ]] will be given in the following section.

The purpose of the present report is to evaluate the relative performances of a pool of descriptors in relating the hydrocarbon molecular structures to a set of physical properties. It is of particular usefulness to predict values of some molecular properties, namely those which are difficult to measure or show health risk or for unavailable substances. In this respect, the newly designed index LEL and those provided by the TOPOCLUJ software are of basic importance. The sets of studied molecules were selected among the representative and sufficiently complex structures (octane isomers and polycyclic aromatic hydrocarbons (PAH)) and the correlations used were in monovariate regression, in view of a more direct interpretation of the results.

2. Description of Indices

In any process of molecular modeling, either quantum or correlating one, the need for a representation of molecular structure is critical and its role is significant to find appropriate predictive models. An information rich representation which is rapidly computed and readily manipulated is essential. This is the case of the topological indices, which are among the most used *molecular descriptors*. TIs are single number descriptors associated with a molecular graph representing a molecule, which does not depend on the numbering and pictorial representation of that graph. In this section, definitions for LEL and the best scored TIs, among the indices provided by the TOPOCLUJ software, are presented.

2. 1. LEL – an Index Built on the Laplacian Matrix

Let G = G(V,E) be a finite, undirected graph with n vertices V = {1,2,...,n} and m = |E| edges. The degree of a vertex u in V is denoted by d_u. Let G have the adjacency matrix A with eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$, and Laplacian matrix L = D-A, where D is the diagonal matrix of vertex degrees, with eigenvalues $\mu_1 \geq \mu_2 \geq ... \geq \mu_n$. The Laplacian-like energy, shortly LEL, of G is defined[8] as:

$$LEL = \sum_{i=1}^{n} \sqrt{\mu_i} \tag{1}$$

It has been shown to have a nice mathematical behavior. It is closely related to the coefficients c_k of the characteristic polynomial of the Laplacian matrix L:

$$\Lambda(G,x) = \sum_{k=0}^{n} (-1)^k c_k x^{n-k} \tag{2}$$

In particular,[9] for two graphs G and H of order n, if $c_k(G) \leq c_k(H)$ for k = 1,...n-1, then LEL(G) ≤ LEL(H). Furthermore, if a strict inequality $c_{k'}(G) < c_{k'}(H)$ holds for at least one value k', then LEL(G) < LEL(H). Using this relation, it has been proved[9] that, among trees of order n, the star S_n has the minimum, while the path P_n has the maximum LEL. Similarly, among unicyclic graphs of order n, the star with an edge between two of its leaves has the minimum LEL, and the cycle C_n has the maximum LEL.[10]

2. 2. Walk Indices or Wiener-type Indices of Higher Rank

A walk $w(1,n)$ in a graph G = G(V,E) (with V = V(G) being the set of vertices and E = E(G) the set of edges) is an alternating sequence of vertices and edges, $w(1,n) = (v_1, e_1, v_2, e_2, ..., v_{n-1}, e_m, v_n)$, $v_i \in V(G)$, $e_i \in E(G)$, $m \geq n - 1$,so that any two subsequent edges are adjacent: $(v_{i-1}, v_i) \in E(G)$. Revisiting vertices and edges is allowed.

If $V(w(1,n)) = \{v_1, v_2, ..., v_{n-1}, v_n\}$ is the set of vertices of the walk $w(1,n)$ and $E(w(1,n)) = \{e_1, e_2, ..., e_{m-1}, e_m\}$ the set of its edges, then $l(w(1,n)) = |E(w_{1,n})|$ represents the *length of walk* $w(1,n)$, which equals the number of traversed edges. If no other condition is imposed, the walk is called a random walk. If the walk starts and ends in the same vertex $v_n = v_1$ it is a closed walk, else it is open.[11–17]

Walks of length e, starting from the vertex $i \in V(G)$ can be counted by summing the entries in the i^{th} row of the e^{th} power of the adjacency matrix A:

$$^{e}W_i = \sum_{j \in V(G)} [A^e]_{ij} \tag{3}$$

$^{e}W_i$ is called the *walk degree* (of rank e) of vertex i (or atomic walk count[18,19]). Local and global invariants based on walks in graph were considered for correlating with physico-chemical properties.[19]

Weighted walk degrees can be easily calculated by means of the algorithm proposed by Diudea *et al.*[20] It evaluates a local (topological) property by iterative summation of vertex contributions over all vertices in row i:

$$M+I = {}^{0}W_M \tag{4}$$

$$[{}^{e+1}W_M]_{i,i} = \sum_j ([M]_{i,j}[{}^{e}W_M]_{j,j}) \tag{5}$$

$$[{}^{e+1}W_M]_{i,j} = [{}^{e}W_M]_{i,j} = [M]_{i,j} \tag{6}$$

where M is a square matrix and $^{e}W_M$ is the diagonal matrix of walk degrees (weighted by the property encoded by the matrix M). The diagonal entries $[{}^eW_M]_{i,i}$ represent the row sum of entries in the matrix M raised at power e $[M^e]_{i,j}$, or the (weighted by M) walk degrees ${}^eW_{M,i}$:

$$[{}^e W_M]_{ii} = \sum_j ([M^e]_{ij}) = {}^e W_{M,i} \tag{7}$$

The half sum of all local invariants ${}^eW_{M,i}$ in G defines a global invariant called the walk number ${}^eW_M$:

$${}^e W_M = {}^e W_M(G) = \frac{1}{2}\sum_i {}^e W_{M,i} \tag{8}$$

When M = A, then ${}^eW_A$ represents the molecular walk count:[18] when M = D (distance matrix), then ${}^eW_D$ is just the Wiener index,[21] of rank e. The extension of this idea resulted in Wiener-type indices of higher rank;[19] the info matrix M can be any square topological matrix. Among such matrices, the following ones are discussed in the present paper: D, H (of reciprocal topological distance entries, also called Harary matrix),[16,17,22] K (of reciprocal D_p matrix entries)[16,17,23] and UCJD (the unsymmetrical Cluj matrix on distances).[24–26] For definition of these matrices and derived indices the reader is invited to consult the refs.[16,17,27]

2. 3. Indices Designed on Layer/shell Matrices

Define a layer of vertices located at distance k to the vertex i as:[19,28,29]

$$G(i)_k = \left\{ v \mid v \in V(G); \quad d_{iv} = k \right\} \tag{9}$$

The partition of G with respect to i will be:

$$G(i) = \left\{ G(i)_k \; ; \; k \in [0,1,..,ecc_i] \right\} \tag{10}$$

with ecc_i being the *eccentricity* of i (*i.e.*, the largest distance from i to the other vertices of G). The entries in the layer matrix (of vertex property) LM, is defined as:

$$[\mathrm{LM}]_{i,k} = \underset{v|d_{i,v}=k}{\Omega} p_v \tag{11}$$

with the most used operation being the summation. The zero columns are just the column of vertex properties $[\mathrm{LM}]_{i,0} = p_i$. Any atomic/vertex property can be considered as p_i. More over, any square matrix M can be taken as info matrix, in supplying local/vertex properties as row sum *RS*, column sum *CS* or diagonal entries given by the *Walk* matrix.[16,17]

Layer matrix is a collection of the above defined entries:

$$\mathrm{LM} = \left\{ [\mathrm{LM}]_{i,k} ; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{12}$$

with $d(G)$ being the diameter of the graph (*i.e.*, the largest distance in G).

Define the entries in the shell matrix (of pair vertex property) SM as:[29]

$$[\mathrm{SM}]_{i,k} = \underset{v|d_{i,v}=k}{\Omega} [\mathrm{M}]_{i,v} \tag{13}$$

with the most used operation being the summation.

Shell matrix is a collection of the above defined entries:

$$\mathrm{SM} = \left\{ [\mathrm{SM}]_{i,k} ; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{14}$$

Table 1. Topological indices for octanes

Molecule	LEL	1WD	2WD	1WW	1WH	2WH	1WK	2WK	1/1WK	1/2WK	2WUCJD	PDS3[Sh[Sz]]
1	9.153	84	1848	84	13.743	48.279	10.564	29.040	0.095	0.034	1596	120.000
2	9.120	79	1628	79	14.100	51.050	10.862	31.153	0.092	0.032	1396	78.600
3	9.115	76	1512	76	14.267	52.495	10.981	32.125	0.091	0.031	1284	94.320
4	9.114	75	1476	75	14.317	52.947	11.014	32.411	0.091	0.031	1248	94.320
5	9.108	72	1360	72	14.483	54.377	11.133	33.373	0.090	0.030	1136	110.040
6	9.065	71	1316	71	14.767	56.500	11.433	35.426	0.087	0.028	1112	78.600
7	9.079	70	1280	70	14.733	56.317	11.367	35.024	0.088	0.029	1072	110.040
8	9.082		1312	71	14.650	55.560	11.300	34.454	0.088	0.029	1102	94.320
9	9.088	74	1420	74	14.467	53.939	11.167	33.343	0.090	0.030	1206	78.600
10	9.056	67	1176	67	15.033	58.878	11.633	37.107	0.086	0.027	978	110.040
11	9.074	68	1208	68	14.867	57.482	11.467	35.847	0.087	0.028	1004	125.760
12	9.073	67	1172	67	14.917	57.924	11.500	36.124	0.087	0.028	968	141.480
13	9.049	64	1072	64	15.250	60.792	11.800	38.493	0.085	0.026	880	125.760
14	9.023	63	1032	63	15.417	62.042	11.967	39.621	0.084	0.025	850	78.600
15	9.031	66	1128	66	15.167	59.771	11.767	37.911	0.085	0.026	940	141.480
16	9.020	62	1000	62	15.500	62.799	12.033	40.191	0.083	0.025	820	125.760
17	9.044	65	1096	65	15.167	59.889	11.733	37.791	0.085	0.026	906	141.480
18	8.971	58	868	58	16.000	67.000	12.500	43.750	0.080	0.023	706	125.760

The zero column, $[SM]_{i,0} = 1$, in case of zero diagonal square info matrix but any other pair vertex property (written as diagonal entries) can be considered.

The index PDS3 [Sh[SZ]] in Table 1, is calculated as columns sum up to distance three on the shell matrix of Szeged matrix, taken as info matrix. For Szeged matrices and indices the reader is invited to consult refs.[16,17,27]

The TOPOCLUJ software package[7] is designed to calculate topological descriptors from topological matrices and/or polynomials. Several weighting schemes including group electronegativity, group mass and partial charges are enabled. Topological indices derived from the matrices: adjacency, connectivity, distance, detour, distance-path, detour-path, Cluj, their reciprocal matrices, walk-matrices, walk-operated matrices, layer- and shell-matrices were successfully used in correlating studies and graph discriminating analysis during the last decade.[17,30] The values of the best scored TIs for octane isomers are listed in Table 1.

Data for the octanes properties and correlation coefficients, in monovariate linear regression, with the best scored topological indices, among those provided by TOPOCLUJ software, are shown in Table 2. The inter-correlation of indices is presented in Table3 while in Table 4 that of the properties of the octane isomers.

Table 2. Data for octanes properties: boiling point BP, motor octane number MON, heat of vaporizat-ion HV, molar volume MV, entropy S, total surface area TSA, acentric factor AF, molar refraction MR, partition coefficient n-octanol/water Log P , density D, critical temperature CT, critical pressure CP and heat of formation DHF and their correlation coefficients, in monovariate linear regression, with the selected topological indices

Mol.	BP	MON	HV	MV	S	TSA	AF	MR	LogP	DENS	CT	CP	DHF
1	9.15	–	34.41	162.61	111.67	415.30	0.40	39.19	3.67	0.70	296.20	24.64	4.14
2	9.12	23.10	33.81	163.65	109.84	407.85	0.38	39.23	3.61	0.70	288.00	24.80	3.06
3	9.11	35.00	33.89	161.85	111.26	397.34	0.37	39.10	3.61	0.71	292.00	25.60	3.29
4	9.11	39.00	33.89	162.12	109.32	396.04	0.37	39.12	3.61	0.70	290.00	25.60	4.00
5	9.11	52.40	33.61	160.08	109.43	379.04	0.36	38.94	3.61	0.71	292.00	25.74	3.59
6	9.06	77.40	32.26	164.29	103.42	405.11	0.34	39.25	3.65	0.70	279.00	25.60	2.56
7	9.08	78.90	33.20	160.41	108.02	384.93	0.35	38.98	3.54	0.71	293.00	26.60	4.23
8	9.08	69.90	32.59	163.09	106.98	388.11	0.34	39.13	3.54	0.70	282.00	25.80	2.80
9	9.09	55.70	32.64	164.72	105.72	395.08	0.36	39.26	3.54	0.69	279.00	25.00	2.50
10	9.06	83.40	32.47	160.89	104.74	389.79	0.32	39.01	3.65	0.71	290.84	27.20	3.17
11	9.07	81.70	33.28	158.65	106.59	376.91	0.34	38.85	3.54	0.72	298.00	27.40	4.97
12	9.07	88.10	–	158.81	106.06	368.10	0.33	38.84	3.54	0.72	295.00	27.40	5.08
13	9.05	88.70	32.79	157.04	101.48	366.99	0.31	38.72	3.65	0.73	305.00	28.90	4.76
14	9.02	99.90	32.01	159.52	101.31	371.75	0.30	38.93	3.58	0.72	294.00	28.20	4.09
15	9.03	100.00	31.01	165.10	104.09	392.19	0.31	39.26	3.58	0.69	271.15	25.50	3.13
16	9.02	99.40	32.34	157.30	102.06	377.40	0.29	38.76	3.58	0.73	303.00	29.00	4.52
17	9.04	95.90	32.73	158.85	102.39	368.93	0.32	38.87	3.48	0.72	295.00	27.60	4.32
18	8.97	–	31.42	138.60	93.06	390.47	0.26	–	3.62	0.82	270.80	24.50	4.88
LEL	*0.673*	*–0.941*	*0.887*	0.647	*0.950*	0.509	*0.991*	0.462	0.175	–0.646	0.243	–0.399	–0.319
1WD	0.559	–0.958	0.715	0.627	0.878	0.721	0.966	0.672	*0.701*	–0.612	0.054	–0.577	–0.459
2WD	0.561	*–0.966*	0.803	0.595	0.864	0.732	0.956	0.648	*0.719*	–0.580	0.053	–0.583	–0.424
1WW	*0.540*	–0.957	0.795	0.628	0.878	0.721	0.966	0.669	0.258	–0.613	0.041	–0.579	–0.453
1WH	–0.573	0.953	–0.826	–0.676	–0.929	–0.622	*–0.992*	–0.616	–0.679	0.666	–0.111	0.505	0.428
2WH	–0.557	0.948	–0.816	–0.687	–0.929	–0.625	*–0.991*	–0.633	–0.173	0.676	–0.098	0.510	0.443
1WK	–0.589	0.949	–0.837	–0.682	*–0.940*	–0.587	*–0.995*	–0.588	–0.671	0.675	–0.139	0.477	0.410
2WK	–0.569	0.945	–0.825	–0.694	–0.939	–0.594	*–0.995*	–0.610	–0.153	0.685	–0.121	0.486	0.429
1/1WK	*0.599*	–0.957	*0.843*	0.653	0.930	0.611	*0.992*	0.585	0.187	–0.644	0.130	–0.501	–0.398
1/2WK	0.595	–0.963	0.836	0.623	0.910	0.648	0.983	0.599	0.211	–0.612	0.104	–0.537	–0.395
2WUCJD	*0.551*	–0.963	0.794	0.592	0.855	*0.744*	0.950	0.658	*0.716*	–0.576	0.040	–0.592	–0.431
PDS3 [Sh[SZ]]	0.432	0.561	–0.010	*–0.984*	–0.500	*–0.896*	–0.600	*–0.991*	–0.122	*0.985*	*0.894*	*0.948*	*0.866*

Legend of indices : LEL = Laplacian-like energy
1WD and 2WD = Wiener numbers of higher rank (1 and 2)
1WW = Hyper-Wiener number
1WH and 2WH = Harary numbers of higher rank (1 and 2)
1WK and 2WK = Wiener-type number, taken the reciprocal of entries in the D_p matrix, of higher rank (1 and 2).
1/1WK and 1/2WK = global reciprocal of 1WK and 2WK
2WUCJD = Wiener-type number (of rank 2), calculated on Cluj matrix on distance
PDS3 [Sh[SZ]] = Columns sum up to distance 3 on the Shell matrix calculated on SZ matrix

Table 3. Intercorrelation matrix for the best scored indices in octanes

Variable	LEL	1WD	2WD	1WW	1WH	2WH	1WK	2WK	1/1WK	1/2WK	2WUCJD	PDS3 [Sh[SZ]]
LEL	1.00	0.96	0.95	0.96	–0.99	–0.98	–0.99	–0.99	0.99	0.98	0.94	–0.34
1WD		1.00	1.00	1.00	–0.99	–0.99	–0.98	–0.98	0.99	0.99	1.00	–0.43
2WD			1.00	1.00	–0.98	–0.98	–0.97	–0.97	0.98	0.99	1.00	–0.42
1WW				1.00	–0.99	–0.99	–0.98	–0.98	0.99	0.99	1.00	–0.43
1WH					1.00	1.00	1.00	1.00	–1.00	–1.00	–0.98	0.40
2WH						1.00	1.00	1.00	–1.00	–0.99	–0.98	0.41
1WK							1.00	1.00	–1.00	–0.99	–0.97	0.38
2WK								1.00	–1.00	–0.99	–0.97	0.39
1/1WK									1.00	1.00	0.98	–0.38
1/2WK										1.00	0.99	–0.39
2WUCJD											1.00	–0.43
PDS3[Sh[S]]												1.00

Table 4. Intercorrelation matrix for the selected molecular properties of octanes

Variable	BP	MON	HV	MV	S	TSA	AF	MR	LogP	DENS	CT	CP	DHF
BP	1.00	–0.32	0.11	0.12	0.62	0.06	0.63	–0.31	0.18	–0.15	0.75	0.08	0.33
MON		1.00	0.08	–0.38	–0.62	–0.39	–0.66	0.24	–0.09	0.37	0.05	0.43	0.49
HV			1.00	–0.02	0.09	–0.32	0.03	–0.09	–0.25	0.00	0.16	0.16	0.37
MV				1.00	0.73	0.29	0.68	–0.90	–0.03	–1.00	0.13	–0.04	–0.61
S					1.00	0.41	0.95	–0.68	0.07	–0.74	0.30	–0.28	–0.32
TSA						1.00	0.55	0.07	0.53	–0.25	–0.46	–0.82	–0.61
AF							1.00	–0.56	0.15	–0.67	0.19	–0.44	–0.39
MR								1.00	0.15	0.92	–0.51	–0.35	0.29
LogP									1.00	0.05	–0.07	–0.28	–0.24
DENS										1.00	–0.18	–0.01	0.58
CT											1.00	0.71	0.50
CP												1.00	0.50
DHF													1.00

Data for a second set of 82 polycyclic aromatic hydrocarbons are included in Table 5 while the correlations are given in Table 6.

Table 5. Data for polycyclic aromatic hydrocarbons (PAH): melting point MP, boiling point BP, partition coefficient n-octanol/water Log P and the corresponding LEL, Wiener W and Randić χ indices

No.	Molecule	MP	BP	LogP	LEL	W	χ
1	naphtalene	81	218	3.35	13.341	109	4.966
2	1-methylnaphthalene	–22	245	3.87	14.572	140	5.377
3	2-methylnaphthalene	35	241	4	14.575	144	5.36
4	1-ethylnaphthalene	–14	259	4.39	15.837	182	5.915
5	2-ethylnaphthalene	–7	258	4.38	15.841	190	5.898
6	2-6-dimethylnaphthalene	110	262	4.31	15.808	186	5.754
7	2-7-dimethylnaphthalene	97	262	–	15.808	185	5.754
8	1-7-dimethylnaphthalene	–14	263	4.44	15.805	180	5.771
9	1-5-dimethylnaphthalene	80	269	4.31	15.802	176	5.788
10	1-2-dimethylnaphthalene	–4	271	4.31	15.803	178	5.788
11	1-3-7-trimethylnaphthalene	14	280	–	17.037	226	6.165
12	2-3-5-trimethylnaphthalene	25	285	–	17.035	224	6.182
13	2-3-6-trimethylnaphthalene	101	286	4.73	17.038	230	6.165
14	phenalene	85	–	–	17.919	210	6.449
15	1-phenylnaphthalene	45	334	–	21.739	412	7.949
16	2-phenylnaphthalene	104	360	–	21.744	436	7.933
17	anthracene	216	340	4.5	19.197	279	6.933

No.	Molecule	MP	BP	LogP	LEL	W	χ
18	1-methylanthracene	86	363	–	20.426	334	7.343
19	2-methylanthracene	209	–	–	20.429	342	7.327
20	2-7-dimethylanthracene	241	370	–	21.66	413	7.72
21	2-6-dimethylanthracene	250	370	–	21.66	414	7.72
22	2-3-dimethylanthracene	252	–	–	21.658	408	7.737
23	9-10-dimethylanthracene	183	–	5.69	21.646	378	7.788
24	phenanthrene	101	338	4.52	19.194	271	6.949
25	1-methylphenanthrene	123	359	5.08	20.422	326	7.36
26	2-methylphenanthrene	56	355	5.24	20.425	334	7.343
27	3-methylphenanthrene	65	352	5.15	20.425	330	7.343
28	4-methylphenanthrene	50	–	–	20.422	322	7.36
29	9-methylphenanthrene	91	355	–	20.421	322	7.36
30	3-6-dimethylphenanthrene	141	363	–	21.656	396	7.737
31	4-5-methylenephenanthrene	116	359	–	21.195	300	7.433
32	tetracene	257	–	5.76	25.047	569	8.899
33	benzo[a]anthracene	162	–	5.91	25.043	553	8.916
34	chrysene	256	441	5.86	25.039	545	8.933
35	benzo[c]phenanthrene	68	–	–	25.038	529	8.933
36	triphenylene	199	439	5.49	25.032	513	8.949
37	pyrene	156	393	5	22.49	362	7.933
38	1-methylpyrene	70	410	–	23.717	428	8.343
39	2-methylpyrene	144	410	–	23.72	434	8.327
40	4-methylpyrene	148	410	–	23.717	424	8.343
41	2-7-dimethylpyrene	–	396	–	24.949	515	8.72
42	pentacene	271	–	–	30.894	1011	10.865
43	dibenzo[ai]anthracene	264	–	6.81	30.889	987	10.882
44	dibenzo[ah]anthracene	270	–	5.8	30.885	971	10.899
45	dibenzo[aj]anthracene	198	–	–	30.885	955	10.899
46	benzo[b]chrysene	294	–	–	30.885	971	10.899
47	dibenzo[ac]anthracene	205	–	–	30.877	907	10.916
48	pycene	–	519	–	30.881	963	10.916
49	benzo[a]pyrene	177	496	5.97	28.331	680	9.916
50	benzo[e]pyrene	179	493	–	28.325	652	9.933
51	perylene	278	–	6.25	28.326	654	9.933
52	coronene	360	–	6.5	34.906	1002	11.899
53	anthranthrene	261	–	–	31.621	839	10.899
54	benzo[ghi]perylene	283	–	6.9	31.617	815	10.916
55	dibenzo[ae]pyrene	234	–	–	34.163	1082	11.916
56	1-methylchrysene	161	–	–	26.265	620	9.343
57	6-methylchrysene	257	–	–	26.267	632	9.343
58	3-methylcholanthrene	180	–	6.75	29.54	804	10.327
59	indeno[1-2-3-cd]pyrene	163	–	–	31.599	845	10.916
60	pentaphene	263	–	–	30.889	979	10.882
61	hexaphene	308	–	–	36.734	1589	12.848
62	indano	–51	178	–	12.043	79	4.466
63	indene	–2	183	2.92	12.043	79	4.466
64	azulene	100	270	3.22	13.335	107	4.966
65	acenaphthene	96	279	3.92	16.624	166	5.949
66	acenaphthylene	93	270		16.624	166	5.949
67	fluorene	117	294	4.18	17.899	219	6.449
68	1-methylfluorene	87	318	4.97	19.128	267	6.86
69	2-methylfluorene	104	318	–	19.131	274	6.843
70	3-methylfluorene	88	316	–	19.131	272	6.843
71	4-methylfluorene	71	–	–	19.128	265	6.86
72	9-methylfluorene	47	–	–	19.125	262	6.877
73	1-2-benzofluorene	190	407	5.4	23.746	461	8.433
74	fluoranthene	111	383	5.2	22.466	364	7.949
75	2-3-benzofluorene	209	402	5.75	23.75	471	8.416
76	3-4-benzofluorene	125	406	–	23.745	453	8.433
77	benzo[ghi]fluoranthene	149	432	5.78	25.759	478	8.933

No.	Molecule	MP	BP	LogP	LEL	W	χ
78	benzo[k]fluoranthene	217	481	–	28.313	698	9.916
79	benzo[b]fluoranthene	168	481	–	28.307	676	9.933
80	benzo[j]fluoranthene	166	480	–	28.309	678	9.933
81	ovalene	473	–	–	47.307	2106	15.865
82	quaterryllene	483	–	–	58.242	4544	19.865

Table 6. Correlation of PAH properties with selected topological indices

Property	LEL	W	χ
MP (n = 80)	0.857	0.748	0.855
BP (n = 53)	0.989	0.955	0.988
LogP (n = 37)	0.945	0.905	0.948

3. Results and Discussion

Wiener index is perhaps the most studied topological descriptor. It was used to predict the thermodynamic properties of hydrocarbons since the pioneering works of Wiener. The pool of Wiener-type descriptors is still in growth.

Other indices gained considerable attention, and revisited from time to time, to add new and important results. Among these, the Zagreb indices,[31] the Hosoya indices and polynomials[32] and the Cluj indices and polynomials[33] are the best known.

In the present work we selected twelve TIs, all but LEL being extensions of the Wiener index. Thirteen properties[34,35] were considered for correlating study, their values in octane isomers and the correlation with the best scored TIs (among more than eighth hundred TIs provided by TOPOCLUJ software package) being listed in Table 1. Other results will refer to a set of aromatic hydrocarbons (see below).

Table 2 contains properties from the benchmark data sets of Milano Chemometrics & QSAR Research Group[34], except HV (Heat of Vaporization) – which is taken from the Korea Thermophysical Properties Data Bank[35]. Nonanes and decanes data are taken from the Korea Thermophysical Properties Data Bank, while polyaromatic hydrocarbons (PAH) are taken from the Milano Chemometrics & QSAR Research Group. All correlations were made in monovariate linear regression, to have a more direct interpretation of the results and a good insight on the structure-property relationship.

Boiling point BP, one of the difficult to model properties, is the best correlated by the newly proposed descriptor LEL: R = 0.673. In n-alkanes, up to eicosane, the Pearson correlation index is R = 0.978. The index clearly bears size information because in the set of nonanes R = 0.525 while in decane isomers R = 0.265. The second well correlated among the selected indices, in the set of octanes, is 1/1WK, (0.599). Compare with the classical Wiener index for which R = 0.559.

Octane number MON, related to the motor combustion, is the best modeled by 2WD, a Wiener-type index of rank 2 (–0.966), LEL being this time second from the bottom but still highly correlated (–0.941).

Heat of vaporization HV, is the best modeled by LEL (0.887), next being 1/1WK, (0.843).

The three above properties are the worst described by PDS3 [Sh[SZ]] (0.432, 0.561, –0.010, respectively). This index, calculated on a shell-type matrix is practically uncorrelated with all the other indices herein discussed and fit the best in case of "recalcitrant" properties, when the other indices failed (see Table 2, last row, the italicized entries). The matrix of index inter-correlation is given in Table 3. It can be seen that all but the last are related higher than 0.94. LEL is well correlated with the WK indices.

Molecular volume MV, molecular refraction MR and density DENS are highly correlated properties (see Table 4). These are modestly described by the all indices except PDS3 [Sh[SZ]], as above mentioned.

Acentric function AF, is the property on which we focus the attention in the following. This parameter is a well described one by all indices excepting the last index. The best score was recorded by 1WK (–0.995); the newly introduced index LEL is also highly correlated to this parameter in heptanes (0.987), octanes (0.990), nonanes[35] (0.978) and decanes[35] (0.975).

A second set of aromatic hydrocarbons[34] (Table 5) was also investigated: melting point MP, boiling point BP and the partition coefficient n-octanol/water LogP, vs. three TIs, LEL, Wiener W and Randić χ index.[36] The results show LEL as good as χ index and better than Wiener index (Table 6). Within this set of molecules, the degeneracy is comparable for the three indices; LEL is correlated here with the Wiener index W by 0.914 while with the Randić index χ by 0.999.

4. Conclusions

A correlating study of topological indices provided by TOPOCLUJ software package and LEL, a newly proposed index, on thirteen properties of octanes revealed good correlating ability of a dozen selected TIs, all related to the Wiener index.

LEL is the best correlated with the WK indices. It describes well the properties which are well accounted by the majority of the selected molecular descriptors: octane number MON, entropy S, volume MV, or refraction MR, particularly the AF parameter, but also more difficult properties like boiling point, melting point and logP. Among the desirable attributes required by a good TI, LEL fulfils: good correlation with at least one property, good discrimination of isomers and simplicity. In addition, it is well defined mathematically and shows interesting relations in particular classes of graphs. This index and those provided by the TOPOCLUJ software as well, was proved to be of basic importance in QSAR/QSPR studies.

5. Acknowledgement

This work is supported in part by the research program P1-0285 of the Slovenian Agency for Research and the grants 144015G and 144007 of the Serbian Ministry of Science and Technological Development and in part by the Romanian GRANT ID_506.

6. References

1. M. Randić, *J. Math. Chem.* **1991**, *7*, 155–168.
2. S. C. Basak, D. K. Harris, V. R. Magnuson, *POLLY* 2.3, University of Minnesota, **1988.**
3. L. H. Hall, *MOLCONN* and *MOLCONN2*, Hall Associates Consulting, 2 Davis Street, Quincy, MA 02170.
4. A. R. Katritzky, V. S. Lobanov, M. Karelson, CODESSA Manual, University of Florida, Gainesville, Fl, **1995**.
5. M. Karelson, V. S. Lobanov, A. R. Katritzky, *Chem. Rev.* **1996**, *96*, 1027–1043.
6. R. Todeschini et al., Dragon software, http://www.talete.mi.it.
7. O. Ursu, M. V. Diudea, TOPOCLUJ 4.0, Babes-Bolyai University, **2005**.
8. J. Liu, B. A. Liu, *MATCH Commun. Math. Comput. Chem.* **2008,** *59*, 355–372.
9. D. Stevanović, *MATCH Commun. Math. Comput. Chem.* **2009**, *61*, 407–417.
10. D. Stevanović, A. Ilić, On the Laplacian coefficients of unicyclic graphs, *Lin. Algebra Appl.* **2009**, *430*, 2290–2300.
11. F. Harary, Graph Theory, Addison-Wesley, Reading, Ma, 1969.
12. N. Trinajstić, Chemical Graph Theory, CRC Press, Inc., Boca Raton, Florida, **1983**.
13. E. Cayley, *Philos. Mag.* **1874**, *67*, 444–446.
14. C. Berge, Graph Theory and Its Applications (in Romanian), Ed. Tehnică, Bucharest, **1969**.
15. M. Randić, W. L. Woodworth, A. Graovac, *Int. J. Quant. Chem.* **1983**, *24*, 435–452.
16. M. V. Diudea, I. Gutman, L. Jäntschi, Molecular Topology, NOVA, New York, **2002**.
17. M. V. Diudea, M. S. Florescu, P. V. Khadikar, Molecular Topology and Its Applications, EFICON, Bucharest, **2006**.
18. G. Rucker, C. Rucker, *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 683–695.
19. M. V. Diudea, *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 535–540.
20. M. V. Diudea, M. Topan, A. Graovac, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1071–1078.
21. H. Wiener, *J. Amer. Chem. Soc.* **1947**, *69***,** 17–20.
22. M. V. Diudea, *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 292–299.
23. M. V. Diudea, I. Gutman, *Croat. Chem. Acta,* **1998**, *71*, 21–51.
24. M. V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **1997**, *35*, 169–183.
25. M. V. Diudea, *J. Chem. Inf. Comput. Sci*. **1997**, *37*, 300–305.
26. M. V. Diudea, B. Prv, M. I. Topan, *J. Serb. Chem. Soc*. **1997**, *62*, 267–276
27. D. Janežič, S. Nikolić, N. Trinajstić, Graph Theoretical Matrices in Chemistry, MCM, Kragujevac, **2007**.
28. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1064–1071.
29. M. V. Diudea, O. Ursu, *Indian J. Chem.*, 42*A*, **2003**, 1283–1294.
30. G. Katona, M. V. Diudea, *Acta Univ. Cibiniensis*, **2002**, *5*, 19–36.
31. D. Janežič, A. Miličević, S. Nikolić, N. Trinajstić, and D. Vukičević, *Croat. Chem. Acta* **2007**, *80*, 541–545.
32. D. Janežič, B. Luičić, A. Miličević, S. Nikolić, N. Trinajstić, and D. Vukičević, *Croat. Chem. Acta* **2007**, *80*, 541–545.
33. M. V. Diudea, A. E. Vizitiu and D. Janežič, *J. Chem. Inf. Model*. **2007**, *47*, 864–874.
34. Milano Chemometrics & QSAR Research Group, Molecular Descriptors Dataset (available online at)
35. Korea Thermophysical Properties Data Bank, (available online at).
36. X. Li, Y. Shi, *MATCH Commun. Math. Comput. Chem.* **2008**, *59*, 127–156.

Povzetek

Deskriptorje dobljene s programskim paketom TOPOCLUJ in indeksom LEL osnovanim na Laplaceovi matriki smo uporabili za izgradnjo korelacijskih modelov, ki služijo za opis trinajstih različnih lastnosti oktanov. Deskriptor LEL zelo dobro opiše večino izbranih lastnosti, kot so: oktansko število, entropija, volumen molekule, indeks refrakcije in parameter AF. Deskriptor LEL je uporaben tudi za opis bolj kompleksnih lastnosti kot sta temperaturi vrelišča in tališča, ter logP. Opazili smo dobro korelacijo indeksa LEL z indeksi WK, izračunanimi s programom TOPOCLUJ. Ob korelacijski analizi drugega obravnavanega seta spojin – policiklični aromatski sistemi se je LEL izkazal tako dobro kot Randićev χ indeks in bolje kot Wienerjev indeks. Naša študija je pokazala, da je novi topološki indeks LEL uporaben za matematični opis določenih razredov grafov.