

The Overall Wiener Index—A New Tool for Characterization of Molecular Topology

Danail Bonchev[†]

Program for Theory of Complex Natural Systems, Texas A&M University, Galveston, Texas 77553-1675

Received July 22, 2000

Recently, the concept of overall connectivity of a graph G , $TC(G)$, was introduced as the sum of vertex degrees of **all** subgraphs of G . The approach of more detailed characterization of molecular topology by accounting for all substructures is extended here to the concept of overall distance $OW(G)$ of a graph G , defined as the sum of distances in all subgraphs of G , as well as the sum of e th-order terms, ${}^eOW(G)$, with e being the number of edges in the subgraph. Analytical expressions are presented for $OW(G)$ of several basic classes of graphs. The overall distance is analyzed as a measure of topological complexity in acyclic and cyclic structures. The potential usefulness of the components of this generalized Wiener index in QSPR/QSAR is evaluated by its correlation with a number of properties of C3–C8 alkanes and by a favorable comparison with models based on molecular connectivity indices.

INTRODUCTION

The Wiener number, W , is one of the best descriptors of molecular topology. It was defined in 1947 for (saturated) acyclic hydrocarbons as the sum of the bonds between all pairs of non-hydrogen atoms in the molecule.¹ Wiener proposed a formula for calculating his “path number” by simply counting the number of atoms n_r and n_l on both sides of each bond:

$$W = \sum_{\text{all bonds}} n_r n_l \quad (1)$$

In 1971, Hosoya² revealed the graph theoretical origin of the Wiener number as the half sum of the distance matrix of molecular graph G :

$$W(G) = 1/2 \sum_{i=1}^n \sum_{j=1}^n d_{ij}(G) = 1/2 \sum_{i=1}^n d_i(G) \quad (2)$$

The second sum in eq 2 contains another graph-invariant, $d_i(G)$, the distance of vertex i (also called distance degree or distance sum), which is the total distance from vertex i to all other $(n-1)$ vertices of G and is calculated as the sum over all d_{ij} entries in the i th row of the distance matrix.

The definition of Hosoya thus opened the door to Wiener-number based systematic studies of molecular topology. The Wiener number was used as a measure of *molecular branching* and *cyclicity*, which allows one to identify the basic topological patterns in acyclic and cyclic molecules.^{3–9} A third essential topological feature of molecules, their *centrality*, was defined proceeding from a Wiener-number-based generalized definition of graph center.^{10–13} The minimum of the Wiener number was advanced as a stability criterion in crystal growth and in localizing crystal vacancies and defects.^{14–18} A modification of the Wiener number was developed for infinite polymer chains and was shown to correlate well with polymer electronic characteristics and

physicochemical properties.^{18–24} Included in software packages (e.g. cf. refs 25 and 26), the Wiener number is nowadays among the most widely used molecular descriptors for quantitative structure–property and structure–activity relationships (QSPR/QSAR).

Intensive work has been done on the theory of the Wiener number and distance matrix; properties have been defined as theorems, formulas for different classes of structures have been derived, etc.^{27–32} (See also the special issue of MATCH devoted to the 50th anniversary of the Wiener number³³). A number of new topological indices have been constructed as extensions of the Wiener number.^{34,35} Gutman made another extension of the Wiener formula (1) to cyclic molecules in his Szeged index³⁶ by specifying n_r and n_l as the number of vertices of graph G nearer to the left and right vertices i and j of the edge $\{ij\}$. Perhaps, the most important generalization of the distance-matrix-based Wiener number has been done by Klein and Randić,³⁷ who introduced the concept of *resistance distance* and resistance matrix by analogy with the theory of electrical networks. The sum of resistance matrix entries of a molecule, later called the Kirchhoff number,³⁸ reduces to the Wiener number for acyclic molecules. Thus, the Wiener number may be regarded as a specific case of the Kirchhoff number for acyclic structures. Evidence in favor of this conclusion is the formula derived by Klein et al.,³⁹ according to which the Kirchhoff number of any molecular graph is equal to the sum over the inverse nonzero eigenvalues of the Laplacian matrix (The Laplacian matrix is defined as the difference between the diagonal matrix of vertex degrees and the adjacency matrix.^{40,41}). The same formula was derived earlier for the Wiener number,⁴² but its validity is restricted to acyclic graphs. The extensions and generalizations of the Wiener number, already done and those yet to be developed, have been recently analyzed in detail by Klein et al.³⁹ A question may arise whether, after so many efforts, there would be room for any further essential extension of the Wiener number. This study aims to demonstrate that the potential for such extensions of this seminal topological index, and,

[†] Corresponding author phone: (713)529-2849; e-mail: bonchevd@aol.com.

generally, of the graph-distance-based molecular descriptors, is far from being exhausted. Moreover, it will be shown that the newly constructed *overall* Wiener index lacks the pitfalls of the Wiener number while preserving the best of its features.

BACKGROUND

In a recent series of papers,^{43–46} a new approach to the topological characterization of molecules was advanced. It proceeds from the total number of subgraphs $K(G)$ of molecular graph G and their presentation in a **complete** series of classes of order e , with $e = 0, 1, 2, \dots, E$ being the constant number of edges in each subgraph of a certain class, and E being the number of edges in the entire graph. Denoting the number of subgraphs in the e th class by eK one constructs the ordered sequence $K'(G)$:

$$K'(G) = K\{{}^0K, {}^1K, {}^2K, \dots, {}^EK\} \quad (3)$$

The next step in the procedure makes use of eq 3 to define the **overall** graph-invariant, $X(G)$, and the respective complete series $X'(G)$ by calculating the *partial* e th order invariant, ${}^eX(G)$, for each class of subgraphs having the same number of edges e . The graph-invariant selected in refs 43–46 was the vertex degree, a_i , and the resulted topological descriptor was termed *overall connectivity* and denoted by $TC(G)$. The abbreviation “TC” stands for *topological complexity*, because this first overall topological index was shown to be an adequate measure of structural complexity, increasing regularly with the size of the molecule, as well as with its degree of branching and cyclicity.

$$TC'(G) = TC\{{}^0TC, {}^1TC, {}^2TC, \dots, {}^ETC\} \quad (4)$$

$$TC(G) = \sum_{e=0}^E {}^eTC(G); {}^eTC(G) = \sum_{i=1}^{{}^eK} a_i \quad (5)$$

It should be mentioned that the idea of characterizing the molecular structure by a series of subgraphs of increasing size is not new. It was first advocated by Smolenski⁴⁷ and then by Gordon⁴⁸ in the 1960s. It is an essential part of the cluster expansion approach of Klein.⁴⁹ The fruitfulness of this idea was proven by the tremendous success of molecular connectivity series of indices of Randić, Kier, and Hall.^{50–52} However, it was only in recent publications^{43–46} that the potential of the detailed substructure presentation of molecule was realized in full, both with the introduction of the overall topological index and with that of the complete series of substructure classes. Indeed, a variety of overall molecular descriptors can be constructed proceeding from different subgraph invariants. Trinajstić⁵³ and co-workers made use of squared vertex degrees, thus producing an overall-type version of the Zagreb group indices.⁵⁴ In this study, we use the Wiener numbers of each subgraph, extending thus this seminal molecular descriptor to its most complete version, the overall Wiener index.

THE OVERALL WIENER NUMBER

Definition 1. The overall Wiener number $OW(G)$ of any graph G is defined as the sum of the Wiener numbers $W_i(G_i)$

of all K subgraphs of G :

$$OW(G) = \sum_{i=1}^K W_i(G_i \subset G) \quad (6)$$

As a consequence of eq 2, one may alternatively define $OW(G)$ as the sum of all distances, d_{ij} , or as the sum of all vertex distances (distance degrees), d_i , in all subgraphs of G . Therefore, $OW(G)$ has the meaning of the overall distance of the graph G .

Definition 2. The e th-order overall Wiener number ${}^eOW(G)$ of any graph G is defined as the sum of the Wiener numbers $W_j({}^eG_j)$ of all eK subgraphs ${}^eG_j \subset G$ that have e edges:

$${}^eOW(G) = \sum_{j=1}^{{}^eK} W_j({}^eG_j \subset G) \quad (7)$$

Definition 3. The e th-order overall Wiener number ${}^eOW(G)$ can be presented as a sum of terms, ${}^eOW_k(G)$, representing the sum of the Wiener numbers in the subgraphs of specified type. For acyclic graphs these are the path ($k = p$), cluster ($k = c$), and pathcluster ($k = pc$) type, as defined by Kier and Hall:⁵¹

$${}^eOW(G) = {}^eOW_p(G) + {}^eOW_c(G) + {}^eOW_{pc}(G) = \sum_{j=1}^{{}^eK_p} W_{pj} + \sum_{l=1}^{{}^eK_c} W_{cl} + \sum_{m=1}^{{}^eK_{pc}} W_{pcm} \quad (7a)$$

Definition 4. The overall Wiener number vector $OW'(G)$ of any graph G is the sequence of all ${}^eOW(G)$ s listed in an ascending order of the number of edges e :

$$OW'(G) = OW\{{}^0OW, {}^1OW, {}^2OW, \dots, {}^EOW\} \quad (8)$$

or in more detail for acyclic graphs:

$$OW'(G) = OW\{{}^0OW, {}^1OW, {}^2OW, {}^3OW_p, {}^3OW_c, \dots, {}^EOW_p, {}^EOW_c, {}^EOW_{pc}\} \quad (8a)$$

The following corollaries follow directly from Definition 2.

Corollary 1: The E th-order overall Wiener number, ${}^EOW(G)$, is the Wiener number $W(G)$ itself:

$${}^EOW(G) = W \quad (9)$$

Corollary 2: The zero-order overall Wiener index, ${}^0OW(G)$, is equal to zero:

$${}^0OW(G) = 0 \quad (10)$$

Corollary 3: The first-order overall Wiener index, ${}^1OW(G)$, is equal to the number of graph edges E :

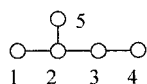
$${}^1OW = E \quad (11)$$

Table 1. Overall Wiener Index OW, and Its e th-Order Components, Calculated for C3–C8 Alkanes

no.	molecule ^a	¹ OW	² OW	³ OW	⁴ OW	⁵ OW	⁶ OW	⁷ OW	OW
1	C3	2	4						6
2	nC4	3	8	10					21
3	2MC3	3	12	9					24
4	nC5	4	12	20	20				56
5	2MC4	4	16	29	18				67
6	22MMC3	4	24	36	16				80
7	nC6	5	16	30	40	35			126
8	2MC5	5	20	39	58	32			154
9	3MC5	5	20	49	56	31			161
10	23MMC4	5	24	58	72	29			188
11	22MMC4	5	28	66	70	28			197
12	nC7	6	20	40	60	70	56		252
13	2MC6	6	24	49	78	102	52		311
14	3MC6	6	24	59	96	98	50		333
15	3EC5	6	24	69	114	93	48		354
16	24MMC5	6	28	58	116	128	48		384
17	23MMC5	6	28	78	130	123	46		411
18	22MMC5	6	32	76	130	124	46		414
19	33MMC5	6	32	96	144	118	44		440
20	223MMMC4	6	36	105	178	143	42		510
21	nC8	7	24	50	80	105	112	84	462
22	2MC7	7	28	59	98	137	164	79	572
23	3MC7	7	28	69	116	168	158	76	622
24	4MC7	7	28	69	136	165	156	75	636
25	3EC6	7	28	79	154	195	148	72	683
26	25MMC6	7	32	68	116	204	208	74	709
27	24MMC6	7	32	78	154	229	200	71	771
28	22MMC6	7	36	86	150	229	202	71	781
29	23MMC6	7	32	88	170	225	198	70	790
30	34MMC6	7	32	98	188	252	192	68	837
31	2M3EC5	7	32	98	208	248	188	67	848
32	33MMC6	7	36	106	204	249	190	67	859
33	3M3EC5	7	36	126	238	270	180	64	921
34	224MMMC5	7	40	95	208	316	236	66	968
35	234MMMC5	7	36	107	224	310	232	65	981
36	223MMMC5	7	40	125	256	332	226	63	1049
37	233MMMC5	7	40	135	272	327	222	62	1065
38	2233MMMMC4	7	48	162	356	429	252	58	1312

^a M and E stand for methyl and ethyl, respectively; the numbers denote the position of the methyl and ethyl branches.

Example:



$e = 1$: 1–2, 2–3, 3–4, 2–5; ${}^1OW = 4 \times 1 = 4$

$e = 2$: 1–2–3, 2–3–4, 1–2–5, 5–2–3;
 ${}^2OW = 4 \times 4 = 16$

$e = 3$: 1–2–3–4, 5–2–3–4, 1–2–3–5;
 ${}^3OW = 2 \times 10 + 9 = 29$

$e = 4$: the entire graph; ${}^4OW = 18$

$OW = 4 + 16 + 29 + 18 = 67$; $OW' = 67(4, 16, 29, 18)$

FORMULAS FOR THE OVERALL WIENER NUMBER OF SOME CLASSES OF GRAPHS

Denote the total number of vertices in graph G by n , the total number of edges by E , the number of edges in a subgraph by e , and the number of subgraphs having e edges each by eK . The following formulas were derived for the overall Wiener number of the specified classes of graphs G :

Linear Graphs, P_n

$${}^eOW(P_n) = e(e+1)(e+2)(n-e)/6;$$

$$OW = (n+3)(n+2)(n+1)n(n-1)/120 \quad (12)$$

Monocyclic Graphs, C_n

$${}^eOW(C_n) = e(e+1)(e+2)n/6 \text{ (for } e = 1, 2, \dots, n-1);$$

$${}^EOW(C_n) = W(C_n) \quad (13)$$

$$OW(C_n) = (n^5 + 2n^4 + 2n^3 - 2n^2 - an)/24;$$

$$a \text{ (} n \text{ even)} = 0; a \text{ (} n \text{ odd)} = 3 \quad (14)$$

Star Graphs, S_n

$${}^eOW(S_n) = \frac{E! e^2}{e! (E-e)!}; OW = \sum_{i=0}^{n-2} \frac{(n-1)!(n-i-1)^2}{i! (n-i-1)!} \quad (15)$$

THE OVERALL WIENER INDEX AS A COMPLEXITY MEASURE

The Wiener number W has been shown a long time ago^{3–9} to be a reliable measure of the two most essential features of molecular topology—the branching of molecular skeleton and the presence of atomic rings in it. More specifically, it has been shown that for isomeric molecules this structural descriptor decreases with the increasing number of branches, their length, centrality (more central positions), clustering (grouping the branches closer to each other), and multiplicity

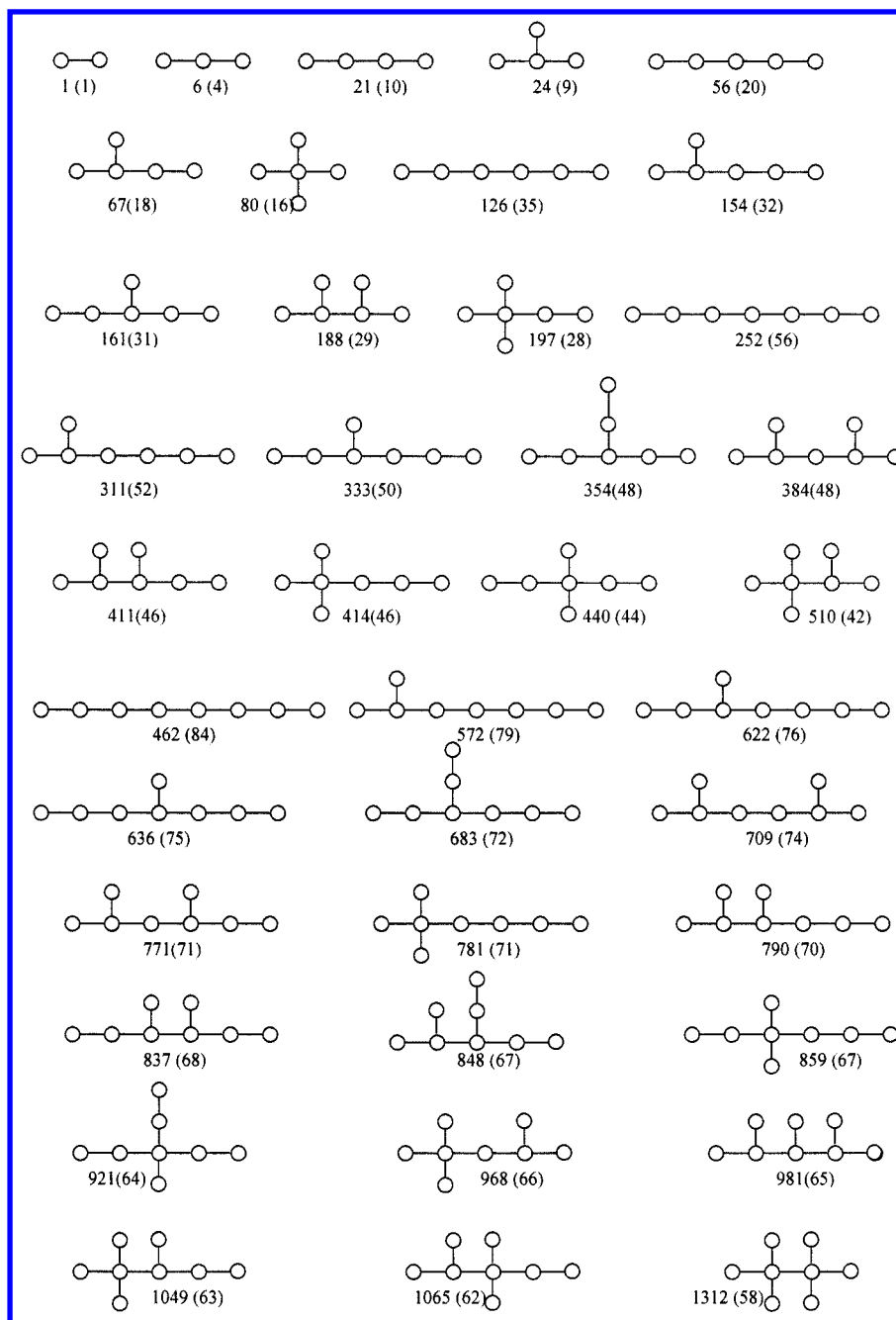


Figure 1. Overall Wiener number vs Wiener number of C2–C8 alkanes.

(higher number of branches at the same atom). Similarly, the Wiener number decreases in isomeric molecules with the increasing number of cycles in molecular skeleton, as well as with cycle interconnectedness, e.g. it decreases in the sequence of two cycles connected by a bridge, by a shared atom, by a shared bond, etc. However, two pitfalls prevent the use of this number as a measure of molecular complexity.^{43–46,55–60} The first one is its degeneracy (having two or more structures with the same Wiener number), which increases fast with the number of atoms and the presence of cycles. More importantly, the Wiener number shows opposing complexity trends when used for a mixed set of isomeric and nonisomeric compounds; it increases with size and decreases with branching and cyclicity. It will be shown in this section that the overall Wiener index preserves the basic features of the original Wiener number, while eliminating its pitfalls; the new descriptor of molecular topology

increases with both size and branching and cyclicity and is much less degenerate.

Table 1 and Figure 1 elucidate these points for acyclic molecules having up to eight non-hydrogen atoms. No two molecules from the selected set have the same overall Wiener number, in contrast to four such pairs having the same *W*-value. The basic branching patterns are illustrated below with overall Wiener number inequalities for selected sets of isomers:

(i) *number of branches*: linear < monobranched < dibranched < tribranched < tetrabranched < ...

(ii) *branch centrality*: 2M < 3M < 4M < ...; 2,2MM < 3,3MM < ...

(iii) *branch length*: 3M < 3E < ...; 2,3MMC6 < 2,3MEC5; 3,3MMC6 < 3,3MEC5;

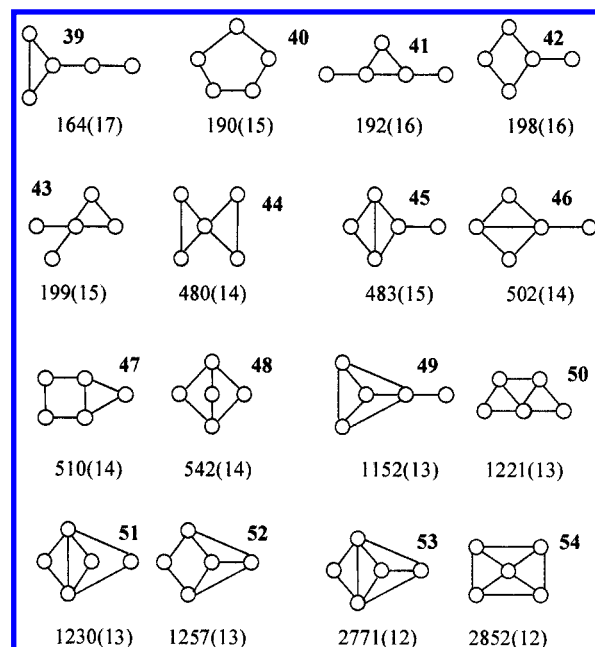
(iv) *branch clustering*: 2,5MMC6 < 2,4MMC6 < 2,3MMC6;

Table 2. Overall Wiener Index OW, and Its ϵ th-Order Components, Calculated for Cyclic Graphs Having Five Vertices

no.	¹ OW	² OW	³ OW	⁴ OW	⁵ OW	⁶ OW	⁷ OW	⁸ OW	OW
39	5	24	52	66	17				164
40	5	20	50	100	15				190
41	5	28	71	72	16				192
42	5	24	69	84	16				198
43	5	32	79	68	15				199
44	6	40	122	200	98	14			480
45	6	40	133	200	89	15			483
46	6	44	151	202	89	14			506
47	6	36	121	236	97	14			510
48	6	36	138	252	96	14			542
49	7	60	255	442	282	93	13		1152
50	7	44	233	492	320	104	13		1213
51	7	60	261	498	303	98	13		1240
52	7	68	222	532	331	108	13		1281
53	8	76	385	968	826	392	104	12	2771
54	8	72	364	1024	868	400	108	12	2856

(v) *branch multiplicity*: 2,3MMC4 < 2,2MMC4;
2,3MMC5 < 2,2MMC5

Besides these five clearly expressed branching patterns, there are more subtle cases in which two or more patterns are combined. In analyzing such cases one should bear in mind that the relative importance of these branching factors varies with the increase of the number of atoms. The latter increases the role of branch centrality, which becomes

**Figure 2.** Overall Wiener number vs Wiener number of cyclic graphs having five vertices.

dominant relative to branch multiplicity (e.g., 2,3MMC6 > 2,2MMC6, compared to 2,3MMC5 < 2,2MMC5), and even

Table 3. Path (p), Cluster (c), and Path Cluster (pc) Terms of Overall Wiener Index OW of C3–C8 Alkanes

no.	molecule ^a	³ W _p	³ W _c	⁴ W _p	⁴ W _c	⁴ W _{pc}	⁵ W _p	⁵ W _c	⁵ W _{pc}	⁶ W _p	⁶ W _c	⁶ W _{pc}	⁷ W _p	⁷ W _c	⁷ W _{pc}
1	C3														
2	nC4	10	0												
3	2MC3	0	9												
4	nC5	20	0	20	0	0									
5	2MC4	20	9	0	0	18									
6	22MMC3	0	36	0	16	0									
7	nC6	30	0	40	0	0	35	0	0						
8	2MC5	30	9	40	0	18	0	0	32						
9	3MC5	40	9	20	0	36	0	0	31						
10	23MMC4	40	18	0	0	72	0	29	0						
11	22MMC4	30	36	0	16	54	0	0	28						
12	nC7	40	0	60	0	0	70	0	0	56	0	0			
13	2MC6	40	9	60	0	18	70	0	32	0	0	52			
14	3MC6	50	9	60	0	36	35	0	63	0	0	50			
15	3EC5	60	9	60	0	54	0	0	93	0	0	48			
16	24MMC5	40	18	80	0	36	0	0	128	0	0	48			
17	23MMC5	60	18	40	0	90	0	29	94	0	0	46			
18	22MMC5	40	36	60	16	54	0	0	124	0	0	46			
19	33MMC5	60	36	20	16	108	0	0	118	0	0	44			
20	223MMMC4	60	45	0	16	162	0	87	56	0	42	0			
21	nC8	50	0	80	0	0	105	0	0	112	0	0	84	0	0
22	2MC7	50	9	80	0	18	105	0	32	112	0	52	0	0	79
23	3MC7	60	9	80	0	36	105	0	63	56	0	102	0	0	76
24	4MC7	60	9	100	0	36	70	0	95	56	0	100	0	0	75
25	3EC6	70	9	100	0	54	70	0	125	0	0	148	0	0	72
26	25MMC6	50	18	80	0	36	140	0	64	0	0	208	0	0	74
27	24MMC6	60	18	100	0	54	70	0	159	0	0	200	0	0	71
28	22MMC6	50	36	80	16	54	105	0	124	0	0	202	0	0	71
29	23MMC6	70	18	80	0	90	70	29	126	0	0	198	0	0	70
30	34MMC6	80	18	80	0	108	35	29	188	0	0	192	0	0	68
31	2M3EC5	80	18	100	0	108	0	29	219	0	0	188	0	0	67
32	33MMC6	70	36	80	16	108	35	0	214	0	0	190	0	0	67
33	3M3EC5	90	36	60	16	162	0	0	270	0	0	180	0	0	64
34	224MMMC5	50	45	120	16	72	0	0	316	0	0	236	0	0	66
35	234MMMC5	80	27	80	0	144	0	58	252	0	0	232	0	0	65
36	223MMMC5	80	45	60	16	180	0	87	245	0	42	184	0	0	63
37	233MMMC5	90	45	40	16	216	0	87	240	0	42	180	0	0	62
38	2233MMMC4	90	72	0	32	324	0	261	168	0	252	0	0	58	0

^a M and E stand for methyl and ethyl, respectively; the numbers denote the position of the methyl and ethyl branches.

Table 4. Comparison of Multilinear Regression Statistics (Correlation Coefficient, Standard Deviation, and the Fischer Ratio) for the C3–C8 Alkane Properties Models with 1, 2, 3, 4, and 5 Parameters Produced by the Sets of Topological Indices $OW(ext)/OW(\chi)/OC$

properties	<i>R</i>	<i>s</i>	<i>F</i>
Boiling Point			
1	0.9850/0.9850/0.9542/0.9850	7.24/7.24/12.5/7.24	1172/1172/366/1172
2	0.9933/0.9933/0.9927/0.9933	4.92/4.92/5.13/4.92	1289/1289/1186/1289
3	0.9971/0.9960/0.9959/0.9973	3.30/3.84/3.90/3.16	1930/1419/1378/2104
4	0.9992/0.9981/0.9968/0.9992	1.82/2.69/3.48/1.78	4778/2175/1296/4990
5	0.9993/0.9980/0.9972/0.9993	1.70/2.60/3.31/1.60	4393/1811/1151/4914
Heat of Formation			
1	0.9788/0.9788/0.9922/0.9788	6.18/6.18/3.77/6.18	823/823/2273/823
2	0.9925/0.9925/0.9961/0.9979	3.73/3.73/2.69/1.95	1158/1158/2251/4112
3	0.9989/0.9989/0.9972/0.9991	1.50/1.50/2.32/1.30	4876/4876/2021/6469
4	0.9991/0.9991/0.9987/0.9992	1.39/1.39/1.56/1.23	4253/4253/3288/5430
5	0.9992/0.9991/0.9991/0.9995	1.33/1.36/1.37/1.02	3697/3531/3479/6324
Heat of Vaporization			
1	0.9687/0.9687/0.9255/0.9687	1.57/1.57/2.40/1.57	549/549/215/549
2	0.9910/0.9910/0.9841/0.9910	0.86/0.86/1.14/0.86	959/959/536/959
3	0.9931/0.9922/0.9891/0.993	0.77/0.81/0.96/0.79	814/718/509/763
4	0.9940/0.9937/0.9921/0.9941	0.73/0.75/0.83/0.72	677/644/516/692
5	0.9945/0.9937/0.9930/0.9950	0.70/0.76/0.79/0.67	579/500/453/633
Heat of Atomization			
1	1.0000/1.0000/0.9890/1.0000	1.46/1.46/57.3/1.46	255 ^a /255 ^a /1612/255 ^a
2	1.0000/1.0000/0.9972/1.0000	0.85/0.85/29.4/0.85	354 ^a /354 ^a /3127/354 ^a
3	1.0000/1.0000/0.9994/1.0000	0.39/0.39/13.5/0.37	1200 ^a /1200 ^a /9915/1300 ^a
4	1.0000/1.0000/0.9997/1.0000	0.36/0.36/9.18/0.33	1050 ^a /1050 ^a /1.61 ^a /1230 ^a
5	1.0000/1.0000/0.9999/1.0000	0.34/0.35/5.78/0.30	873 ^a /873 ^a /3.24 ^a /1190 ^a
Surface Tension			
1	0.8868/0.8868/0.9358/0.8868	0.85/0.85/0.64/0.85	114/114/218/114
2	0.9643/0.9153/0.9803/0.9538	0.49/0.86/0.37/0.56	199/77/369/151
3	0.9879/0.9879/0.9906/0.9879	0.29/0.29/0.26/0.29	391/391/506/393
4	0.9950/0.9927/0.9921/0.9942	0.19/0.23/0.24/0.21	690/475/438/598
5	0.9958/0.9942/0.9938/0.9963	0.18/0.21/0.22/0.17	641/464/432/727
Molar Volume			
1	0.9888/0.9888/0.9649/0.9888	2.17/2.17/3.83/2.17	1320/1320/405/1320
2	0.9990/0.9928/0.9966/0.9987	0.69/1.78/1.22/0.76	6777/994/2122/5467
3	0.9997/0.9996/0.9993/0.9997	0.38/0.41/0.55/0.39	15000/12700/6884/14300
4	0.9998/0.9997/0.9996/0.9998	0.32/0.41/0.42/0.34	16100/9263/9207/14000
5	0.9999/0.9997/0.9997/0.9998	0.23/0.41/0.36/0.33	23500/7447/9955/11400
Molar Refraction			
1	0.9994/0.9994/0.9771/0.9994	0.151/0.151/0.927/0.151	25700/25700/653/25700
2	1.0000/0.9997/0.9996/0.9999	0.044/0.112/0.123/0.055	15.2 ^a /2.35 ^a /1.93 ^a /9.64 ^a
3	1.0000/0.9999/0.9999/1.0000	0.041/0.051/0.062/0.042	11.4 ^a /7.44 ^a /5.17 ^a /11.2 ^a
4	1.0000/0.9999/0.9999/1.0000	0.041/0.051/0.053/0.041	8.69 ^a /5.62 ^a /5.24 ^a /8.78 ^a
5	1.0000/0.9999/1.0000/1.0000	0.042/0.051/0.044/0.041	6.74 ^a /4.45 ^a /6.07 ^a /7.02 ^a
Critical Volume			
1	0.9856/0.9856/0.9614/0.9856	0.0124/0.0124/0.0201/0.0124	1217/1217/439/1217
2	0.9918/0.9918/0.9834/0.9921	0.0095/0.0095/0.0134/0.0093	1055/1055/515/1094
3	0.9934/0.9923/0.9903/0.9931	0.0086/0.0093/0.0104/0.0088	847/726/577/813
4	0.9942/0.9924/0.9930/0.9945	0.0082/0.0094/0.0090/0.0080	704/533/585/748
5	0.9947/0.9924/0.9936/0.9948	0.0080/0.0095/0.0087/0.0079	597/414/497/612
Critical Pressure			
1	0.9317/0.9317/0.9062/0.9317	1.37/1.37/1.59/1.37	237/237/165/237
2	0.9706/0.9549/0.9661/0.9747	0.92/1.14/0.99/0.86	285/181/245/332
3	0.9849/0.9689/0.9800/0.9913	0.67/0.96/0.77/0.51	368/174/275/644
4	0.9934/0.9727/0.9901/0.9941	0.45/0.91/0.55/0.43	623/145/412/691
5	0.9950/0.9742/0.9920/0.9955	0.40/0.90/0.50/0.37	638/119/397/705
Critical Temperature			
1	0.9795/0.9795/0.9548/0.9795	10.3/10.3/15.3/10.3	851/851/371/851
2	0.9842/0.9842/0.9918/0.9850	9.21/9.21/6.63/8.99	542/542/1060/569
3	0.9948/0.9911/0.9932/0.9950	5.38/7.05/6.15/5.28	1081/625/824/1120
4	0.9961/0.9957/0.9948/0.9968	4.71/4.94/5.44/4.28	1061/964/793/1283
5	0.9983/0.9959/0.9962/0.9982	3.23/4.94/4.76/3.25	1810/771/831/1790

^a Multiply these numbers by 10⁴.

to the number of branches (first examples appear in alkanes having more than eight carbon atoms).

Figure 2 and Table 2 provide the opportunity to analyze complexity trends in cyclic molecules with a representative sample of 16 cyclic graphs having five vertices. Once again, there are no degenerate overall Wiener values, in contrast to the Wiener number high degeneracy (there are two pairs,

two triplets, and even one quadruplet of structures having the same Wiener number values). Two basic complexity patterns are clearly manifested. The topological complexity of cyclic molecules increases with the following:

(i) the number of cycles: $OW(39-43) < OW(44-48) < OW(49-52) < OW(53,54)$

(ii) the degree of connectedness of a pair of cycles: $OW(44) < OW(45-47) < OW(48)$

A third pattern, that of complexity increasing with the increase in cycle size at a constant number of vertices may also be identified, e.g. when comparing the tricyclic structures: $OW(49-50) < OW(51-52)$. The same pattern is also observed in the bicyclic structures, although in these cases it acts in parallel with pattern (ii): $OW(44-46) < OW(47) < OW(48)$. More subtle cases again appear when several complexity factors are simultaneously present. Such is the case with the five monocyclic structures **39-43** for which the cycle size is intermingled with a variety of branching factors. Another mixed case is that of the tricyclic graphs **49-51**, but the complexity of molecules containing cycles generally requires a more detailed study.

QUANTITATIVE STRUCTURE-PROPERTY RELATIONSHIPS WITH THE OVERALL WIENER INDEX

C3-C8 Alkanes. The potential of the overall Wiener index OW and its components (the e th order eOW indices) for structure-property and structure-activity studies can be evaluated by performing comparative modeling of selected series of compounds against molecular connectivity indices, the most widely used molecular descriptors in QSPR/QSAR. In this study, we apply multiple regression analysis (MRA) in a QSPR modeling of 10 physicochemical properties of C3-C8 alkanes. The latter are used as a standard for evaluating topological descriptors of molecular structures. The properties are those modeled in our preceding study of overall connectivity indices,⁴⁶ namely boiling points,⁶¹ T_B in °C; critical temperatures,⁶¹ T_c in °C; critical pressures,⁶¹ P_c in atm; critical volume,⁶² V_c in L/mol; molar volume, V_m in cm³/mol;⁵¹ molecular refraction, R_m in cm³/mol;⁵¹ surface tension,⁶¹ ST in dyn/cm; the heat of formation in gaseous state, $\Delta H_f(g)$ in kJ/mol;⁶³ the heat of vaporization, ΔH_v in kJ/mol;⁶³ the heat of atomization, ΔH_a in kcal/mol.⁵¹ The values of the molecular connectivity indices are taken from Kier and Hall's monograph.⁵¹ Two different sets of overall Wiener indices were used in the modeling, the first one included the OW index and the its e th-order components, eOW , for $e = 1, 2, \dots, 7$. The extended set of overall Wiener indices (Table 3) included also the path (p), cluster (c), and pathcluster (pc) terms of order 2 to 7. The best models have been selected from the computer program after exhaustive testing of all variables. The total number of evaluated five-variable models with the extended set of overall Wiener indices exceeded 20,000.

Table 4 summarizes the statistics (correlation coefficient, r , standard deviation, s , and the Fischer ratio, F) of the overall Wiener and molecular connectivity models of the C3-C8 alkanes derived with one to five variables. It was found worthwhile to extend the comparison by including the statistics for the respective models derived earlier with overall connectivity indices,⁶⁴ which showed high performance in the MRA modeling. The comparison of the first two columns of Table 4 shows that considerably better statistics is obtained when using the Kier and Hall⁵¹ grouping of isomeric molecules into path, cluster, and pathcluster classes. All five-variable models decreased their standard deviation after including these eOW_k terms. The improvement was consider-

able for critical pressure ($0.90 \Rightarrow 0.409$), molar volume ($0.41 \Rightarrow 0.23$), critical temperature ($4.94 \Rightarrow 3.23$), boiling point ($2.60 \Rightarrow 1.70$), molar refraction ($0.051 \Rightarrow 0.041$), and critical volume ($0.0095 \Rightarrow 0.0080$).

Comparing columns 1, 3, and 4 in Table 4 one finds that the overall Wiener set of indices compares favorably to molecular connectivity indices in 44 out of 50 pairs of models, including all models with four and five variables. For five of the properties the decrease in standard deviation is a considerable one: heats of atomization – from 5.78 to 0.34, boiling point – from 3.31 to 1.70, molar volume – from 0.36 to 0.23, critical temperature – from 4.76 to 3.23, and critical pressure – from 0.50 to 0.40. Yet, overall connectivity indices provide even better results than overall Wiener index in 26 models, whereas in 14 cases the same statistics resulted (the 10 one-variable models are all the same), and in the remaining 10 models the overall Wiener indices prevail. The standard deviations provided by these two sets of overall indices are rather close, the only more significant difference being that for heat of formation, with $s(OC) = 1.02 < s(OW) = 1.33$, and for molar volume, with $s(OW) = 0.23 < s(OC) = 0.33$.

The best overall Wiener models obtained for the 10 examined alkane properties are given below. The first-order overall Wiener index, which equals the number of edges, ${}^1OW = E$, takes part in all 10 models. The next most frequently used index is the second-order one, which is equal to the number of two-edge subgraphs, 2K , factored by four, the Wiener number of one such subgraph, ${}^2OW = 4({}^2K)$.

$$T_B = -128.6(\pm 2.2) + 47.66(\pm 0.84){}^1OW - 2.090(\pm 0.089){}^2OW - 0.2058(\pm 0.0152){}^4OW - 0.0833(\pm 0.0100){}^5OW_p - 0.0314(\pm 0.0129){}^6OW_p \quad (16)$$

$$n = 38, r = 0.9993, s = 1.70, F = 4393$$

$$\Delta H_f = 68.03(\pm 1.19) + 14.43(\pm 0.35){}^1OW + 2.044(\pm 0.087){}^2OW - 0.1649(\pm 0.0273){}^3OW - 0.0260(\pm 0.0048){}^5OW_{pc} - 0.0425(\pm 0.0081){}^5OW_c \quad (17)$$

$$n = 38, r = 0.9992, s = 1.33, F = 3697$$

$$\Delta H_v = 3.218(\pm 0.948) + 7.714(\pm 0.656){}^1OW - 0.00474(\pm 0.00133)OW - 0.6513(\pm 0.1567){}^2OW + 0.0215(\pm 0.0059){}^4OW + 0.1107(\pm 0.0496){}^3OW_c \quad (18)$$

$$n = 38, r = 9945, s = 0.70, F = 579$$

$$\Delta H_a = 397.15(\pm 0.45) + 278.46(\pm 0.16){}^1OW + 0.4620(\pm 0.0211){}^2OW - 0.00500(\pm 0.00041)OW + 0.0182(\pm 0.0031){}^4OW_p + 0.0132(\pm 0.0019){}^5OW_p \quad (19)$$

$$n = 38, r = 1.0000, s = 0.34, F = 9,190,000$$

$$ST = 7.720(\pm 0.415) + 2.829(\pm 0.100){}^1OW - 0.3396(\pm 0.0138){}^2OW + 0.0611(\pm 0.0034){}^3OW - 0.0510(\pm 0.0022){}^4OW_p - 0.00883(\pm 0.00154){}^7OW_{pc} \quad (20)$$

$$n = 33, r = 9958, s = 0.18, F = 641$$

$$V_m = 50.54(\pm 0.31) + 14.93(\pm 0.07)^1 OW + 0.6942(\pm 0.0182)^2 OW - 0.1836(\pm 0.0045)^3 OW - 0.0263(\pm 0.0034)^5 OW_c + 0.0405(\pm 0.0067)^6 OW_c \quad (21)$$

$$n = 32, r = 0.9999, s = 0.23, F = 23500$$

$$R_m = 6.517(\pm 0.089) + 4.734(\pm 0.025)^1 OW + 0.00657(\pm 0.00409)^2 OW - 0.0115(\pm 0.0017)^3 OW_p - 0.00100(\pm 0.00090)^4 OW + 0.000455(\pm 0.000353)^5 OW \quad (22)$$

$$n = 33, r = 1.0000, s = 0.042, F = 67400$$

$$V_c = 0.0831(\pm 0.0082) + 0.0583(\pm 0.0018)^1 OW - 0.000243(\pm 0.00037)^4 OW + 0.000218(\pm 0.000050)^6 OW - 0.0000638(\pm 0.0000357)^6 OW_{pc} - 0.000216(\pm)^7 OW_{pc} \quad (23)$$

$$n = 38, r = 0.9947, s = 0.0080, F = 597$$

$$P_c = 51.28(\pm 0.59) - 4.931(\pm 0.165)^1 OW + 0.0655(\pm 0.0088)^3 OW_p + 0.0268(\pm 0.0030)^4 OW + 0.0153(\pm 0.0027)^5 OW_p - 0.0296(\pm 0.0024)^6 OW_c \quad (24)$$

$$n = 38, r = 0.9950, s = 0.40, F = 638$$

$$T_c = -18.34(\pm 4.36) + 65.10(\pm 1.63)^1 OW - 0.0325(\pm 0.0048) OW - 3.172(\pm 0.209)^2 OW - 0.5726(\pm 0.0372)^4 OW_p + 0.1926(\pm 0.0150)^5 OW_{pc} \quad (25)$$

$$n = 38, r = 0.9983, s = 3.23, F = 1810$$

Isomeric Octanes. In the previous section the verification of the new set of indices was performed in the manner that is more instructive for practical applications. This includes nonisomeric compounds, and proceeding beyond one-variable correlation to an extended number of variables, so as to produce the least standard deviation. For the set of 38 compounds we consider five variables as an optimal choice (for some of the properties investigated the sixth variable was always insignificant). Some authors argue⁶⁵⁻⁶⁷ that the best verification should be the usage of a series of isomeric compounds, for which the indices tested should show a considerable discriminating power. Another argument used is that in nonisomeric series the correlation coefficient easily goes up to 0.97 only as a function of the number of (non-hydrogen) atoms.⁶⁸ The last is certainly true, but for practical applications we are primarily interested in models producing low standard deviation, the correlation coefficient playing only a secondary role. On the other hand, the models derived for an isomeric series of compounds have no direct practical application, because they cannot be used for predictions. For example, models derived for octanes are not good for nonanes, not referring to decanes.⁶⁹ Therefore, the modeling of isomeric series is basically of theoretical value. For this reason, having already shown in the previous section the

applicability of the overall Wiener indices in QSPR, it was also of interest to examine how well they would perform for an isomeric series. Another motivation was the fact that octane series is frequently used for molecular connectivity type of modeling, including some of its recent generalizations.^{68,70}

In Table 5, a comparison is made between the best model statistics of the octane properties models produced by the series of overall Wiener, molecular connectivity, and overall connectivity indices. As shown in Table 4, two versions of overall Wiener sets of indices were used, the eOW and eOW_k ones. The smaller set of 18 compounds restricted to four the maximum number of variables to be included in a model. The overall Wiener index performed even better than in the nonisomeric series of C3-C8 alkanes. It showed better statistics than the other two sets of indices. Out of 40 pairs of models compared (10 properties times four models with one, two, three, and four variables for each property) the overall Wiener indices compared favorably to molecular connectivity ones in 31 cases. It outperformed also the overall connectivity indices in 22 cases, underperformed it in nine cases, and in another nine cases the quality of the models was equal. Particularly strong is the dominance of overall Wiener indices in the models with three and four variables. The decrease in the standard deviation in models with four variables is particularly strong for molar volume, $s(OW) = 0.15 < s(OC) = 0.34 < s(\chi) = 0.35$, and critical temperature, $s(OW) = 0.87 < s(\chi) = 1.13 < s(OC) = 1.23$. Molecular connectivity indices show better performance (7 out of 10 cases) only for one-variable correlations, with the Randić index (first-order molecular connectivity) playing the major role. This indicates a better potential use of the Randić index in combination with other indices outside the closed set of molecular connectivity indices.

A summary of the performance of overall Wiener, overall connectivity, and molecular connectivity sets of indices is shown in Table 6.

Worth mentioning is also that two of the one-variable correlations obtained with overall-type indices show the best correlation ever for molar volume and molar refraction:

$$R_m = 39.861(\pm 0.033) - 0.0124(\pm 0.0005)^3 OW_p \quad (26)$$

$$n = 17, r = 0.9887, s = 0.028, F = 655$$

$$V_m = 171.634(\pm 0.651) - 0.2622(\pm 0.0158)^3 TCl_p \quad (27)$$

$$n = 16, r = 0.9754, s = 0.55, F = 274$$

As mentioned in a recent publication of Estrada,⁷⁰ the best single-variable correlations with octanes' R_m and V_m , out of 36 compared topological indices, have been obtained with Estrada's edge-connectivity index ϵ^c . The correlation coefficients of the Estrada models, $r(R_m) = 0.9306$ and $r(V_m) = 0.9392$, however, are considerably lower than those of eqs 26 and 27.

Summarizing the results, one may conclude that both types of overall topological indices examined show a good potential for QSPR and QSAR studies, with the statistics of

Table 5. Comparison of Multilinear Regression Statistics (Correlation Coefficient, Standard Deviation, and the Fischer Ratio) for Octane Properties Models with 1, 2, 3, and 4 Parameters Produced by the Sets of Topological Indices $OW(ext)/OW/\chi/OC$

properties	<i>R</i>	<i>s</i>	<i>F</i>
Boiling Point			
1	0.8144/0.8144/0.8807/0.7204	5.61/5.61/2.98/4.38	32/32/56/17
2	0.9892/0.9454/0.9647/0.9579	0.96/2.12/1.72/1.87	341/63/101/83
3	0.9924/0.9886/0.9855/0.9898	0.83/1.01/1.14/0.96	305/201/158/225
4	0.9963/0.9930/0.9918/0.9948	0.61/0.83/0.89/0.71	431/229/196/311
Heat of Formation			
1	0.7808/0.7808/0.9107/0.7561	3.34/3.34/2.21/3.50	25/25/78/12
2	0.9690/0.9690/0.9526/0.9690	1.36/1.36/1.68/1.36	116/116/74/115
3	0.9786/0.9711/0.9745/0.9795	1.18/1.47/1.28/1.15	106/40/88/110
4	0.9827/0.9710/0.9771/0.9827	1.10/1.42/1.26/1.10	92/54/69/92
Heat of Vaporization			
1	0.9232/0.9232/0.9300/0.9232	0.62/0.62/0.59/0.62	92/92/102/92
2	0.9701/0.9605/0.9621/0.9703	0.41/0.47/0.46/0.40	120/89/93/121
3	0.9813/0.9788/0.9775/0.9790	0.33/0.35/0.36/0.35	122/107/100/108
4	0.9841/0.9803/0.9796/0.9855	0.32/0.35/0.36/0.30	100/80/77/110
Heat of Atomization			
1	0.7878/0.7878/0.9124/0.7623	0.79/0.79/0.53/0.83	26/26/79/22
2	0.9677/0.9677/0.9517/0.9679	0.34/0.34/0.41/0.33	110/110/72/111
3	0.9778/0.9690/0.9742/0.9787	0.29/0.34/0.31/0.28	102/72/87/106
4	0.9827/0.9696/0.9762/0.9823	0.26/0.35/0.31/0.27	92/51/66/90
Surface Tension			
1	0.6474/0.4901/0.7646/0.6474	0.69/0.79/0.58/0.69	11/4.7/21/11
2	0.9883/0.9883/0.9764/0.9824	0.14/0.14/0.20/0.18	293/293/143/193
3	0.9930/0.9857/0.9855/0.9908	0.12/0.16/0.16/0.13	305/148/147/233
4	0.9947/0.9938/0.9887/0.9944	0.10/0.11/0.15/0.11	283/239/131/267
Molar Volume			
1	0.9733/0.7227/0.9555/0.9754	0.63/1.88/0.70/0.55	251/15/162/274
2	0.9952/0.9883/0.9796/0.9883	0.28/0.43/0.52/0.40	666/273/154/272
3	0.9960/0.9934/0.9903/0.9920	0.26/0.34/0.37/0.34	497/300/203/246
4	0.9988/0.9951/0.9924/0.9928	0.15/0.30/0.35/0.34	1121/278/179/190
Molar Refraction			
1	0.9887/0.7630/0.9693/0.9887	0.028/0.121/0.046/0.028	655/21/233/655
2	0.9963/0.9817/0.9883/0.9963	0.017/0.037/0.031/0.017	950/186/272/950
3	0.9976/0.9952/0.9947/0.9972	0.014/0.020/0.022/0.015	912/449/377/757
4	0.9981/0.9962/0.9953/0.9978	0.013/0.018/0.021/0.014	801/392/291/682
Critical Volume			
1	0.7319/0.7234/0.6993/0.7342	0.0108/0.0110/0.0114/0.0108	18/18/15/19
2	0.8257/0.8257/0.7812/0.8166	0.0093/0.0093/0.0103/0.0095	16/16/12/15
3	0.8563/0.8563/0.8459/0.8497	0.0088/0.0088/0.0091/0.0090	13/13/12/12
4	0.8780/0.8572/0.8743/0.8697	0.0084/0.0091/0.0086/0.0087	11/9.0/11/10
Critical Pressure			
1	0.7206/0.5792/0.7394/0.7206	1.03/1.21/1.00/1.03	17/8.1/19/17
2	0.9748/0.8127/0.9687/0.9748	0.34/0.89/0.38/0.34	143/15/114/143
3	0.9926/0.8230/0.9867/0.9888	0.19/0.90/0.26/0.24	310/9.8/171/204
4	0.9958/0.8691/0.9906/0.9902	0.15/0.81/0.22/0.23	388/10/171/164
Critical Temperature			
1	0.4835/0.4058/0.6465/0.4732	8.79/9.17/7.66/8.84	4.9/3.2/11/4.6
2	0.8890/0.8565/0.9565/0.9047	4.75/5.35/3.02/4.42	28/21/81/34
3	0.9832/0.8854/0.9929/0.9872	1.96/4.99/1.27/1.71	136/17/326/179
4	0.9970/0.9005/0.9948/0.9938	0.87/4.84/1.13/1.23	532/14/312/262

Table 6. Summary of the Comparative Statistics of the Best Models of Ten Alkane Properties with 1, 2, 3, 4, and 5 Variables Obtained with the Three Sets of Topological Indices, OW , OC , and χ

no. variables	rank								
	OW			OC			χ		
	1	2	3	1	2	3	1	2	3
38 C3–C8 Alkanes									
1	0+8 ^a	0+2 ^b	0	0+8 ^a	0+2 ^b	0	2	0	8
2	2+3 ^a	3	2	3+3 ^a	3	1	2	1	7
3	4	5+1 ^b	0	5	4+1 ^b	0	1	0	9
4	2	8	0	8	2	0	0	0	10
5	2	8	0	8	2	0	0	0	10
total	10+11 ^a	24+3 ^b	2	24+11 ^a	11+3 ^b	1	5	1	44
18 C8 Alkanes									
1	0+2 ^a	4+3 ^b	1	1+2 ^a	1+3 ^b	3	7	0	3
2	4+3 ^a	2	1	2+3 ^a	4	1	1	1	8
3	7	2	1	2	8	0	1	0	9
4	8+1 ^a	1	0	1+1 ^a	5	3	0	3	7
total	19+6 ^a	9+3 ^b	3	6+6 ^a	18+3 ^b	7	9	4	27

^a Shared rank 1 with another descriptor. ^b Shared rank 2 with another descriptor.

their performance found in the present study to follow the order

(i) nonisomeric C3–C8 series:

overall connectivity > overall Wiener > molecular connectivity

(ii) isomeric octane series:

overall Wiener > overall connectivity > molecular connectivity

It should be mentioned that this conclusion is drawn for linear regression models, the comparison of nonlinear models remaining open for later studies.

Commenting more generally on the applicability of the overall Wiener set of indices (like that of the overall connectivity indices) one should mention that their calculation requires the prior generation of all subgraphs or at least (like molecular connectivity indices) such a generation up to a certain level, e.g., up to subgraphs having five to seven

edges. The computer program for substructure generation very recently developed⁷¹ is very helpful in such applications.

ACKNOWLEDGMENT

The comments of Dr. C. Rücker (Freiburg) and Dr. A. T. Balaban (Galveston) are gratefully acknowledged.

Note Added in Proof: Molar volume and molar refraction of alkanes have been modeled in 1999 by C. Rücker using the path number p_3 as a single variable. The results obtained have been communicated to Estrada, who has briefly mentioned them in ref 70. The standard deviation and correlation coefficients of Rücker's models are very close to those of our eqs 26 and 27, because of the linear proportionality of the 3OW_p and 3TCl_p variables to p_3 .

REFERENCES AND NOTES

- Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20. Relation of the Physical Properties of the Isomeric Alkanes to Molecular Structure. *J. Phys. Chem.* **1948**, *52*, 1082–1089.
- Hosoya, H. Topological Index: A Newly Proposed Quantity Characterizing the Topological Nature of Structured Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332–2339.
- Bonchev, D.; Trinajstić, N. Information Theory, Distance matrix, and Molecular Branching. *J. Chem. Phys.* **1977**, *67*, 4517–4533.
- Bonchev, D.; Trinajstić, N. On Topological Characterization of Molecular Branching. *Intern. J. Quantum Chem. Symp.* **1978**, *12*, 293–303.
- Bonchev, D.; Mekenyan, O.; Trinajstić, N. Topological Characterization of Cyclic Structures. *Int. J. Quantum Chem.* **1980**, *17*, 845–893.
- Mekenyan, O.; Bonchev, D.; Trinajstić, N. On Algebraic Characterization of Bridged Polycyclic Compounds. *Int. J. Quantum Chem.* **1981**, *19*, 929–955.
- Mekenyan, O.; Bonchev, D.; Trinajstić, N. A Topological Characterization of Cyclic Structures With Acyclic Branches. *MATCH* **1981**, *11*, 145–168.
- Mekenyan, O.; Bonchev, D.; Trinajstić, N. Structural Complexity and Molecular Properties of Cyclic Molecules With Acyclic Branches. *Croat. Chem. Acta* **1983**, *56*, 237–261.
- Bonchev, D. Topological Order in Molecules. I. Molecular Branching Revisited. *Theochem* **1995**, *336*, 137–156.
- Bonchev, D.; Balaban A. T.; Mekenyan, O. Generalization of the Graph Center Concept, and Derived Topological Indexes. *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 106–113.
- Bonchev, D.; Balaban, A. T.; Randić, M. The Graph Center Concept for Polycyclic Graphs. *Intern. J. Quantum Chem.* **1981**, *19*, 61–82.
- Bonchev, D. The Concept for the Center of a Chemical Structure and Its Applications. *Theochem* **1989**, *185*, 155–168.
- Balaban, A. T.; Bonchev, D.; Seitz, W. A. Topological/Chemical Distances and Graph Centers in Molecular Graphs With Multiple Bonds. *Theochem* **1992**, *280*, 253–260.
- Bonchev, D.; Mekenyan, O.; Fritsche, H. An Approach to the Topological Modeling of Crystal Growth. *J. Cryst. Growth* **1980**, *49*, 90–96.
- Bonchev, D.; Mekenyan, O.; Fritsche, H. A Topological Approach to Crystal Vacancy Studies. I. *Phys. Stat. Sol. (a)* **1979**, *55*, 181–187.
- Mekenyan, O.; Bonchev, D.; Fritsche, H. A Topological Approach to Crystal Vacancy Studies. II. *Phys. Stat. Sol. (a)* **1979**, *56*, 607–614.
- Mekenyan, O.; Bonchev, D.; Fritsche, H. A Topological Approach to Crystal Defect Studies. *Z. Phys. Chem. (Leipzig)* **1984**, *265*, 959–967.
- Fritsche, H.-G.; Bonchev, D.; Mekenyan, O. The Optimum Topology of Small Clusters. *Z. Phys. Chem. (Leipzig)* **1989**, *270*, 467–476.
- Bonchev, D.; Mekenyan, O. A Topological Approach to the Calculation of the π -Electron Energy and Energy Gap of Infinite Conjugated Polymers. *Z. Naturforsch.* **1980**, *35a*, 739–747.
- Bonchev, D.; Mekenyan, O.; Polansky, O. E. A Topological Approach to the Predicting of the Electron Energy Characteristics of Conjugated Infinite Polymers. II. PPP– Calculations. *Z. Naturforsch.* **1981**, *36a*, 643–646.
- Bonchev, D.; Mekenyan, O.; Polansky, O. E., III. The Influence of Some Structural Modifications of Polymers. *Z. Naturforsch.* **1981**, *36a*, 647–650.
- Mekenyan, O.; Dimitrov, S.; Bonchev, D. Graph-Theoretical Approach to the Calculation of Physico-Chemical Properties of Polymers. *Eur. Polymer J.* **1983**, *12*, 1185–1193.
- Bonchev, D.; Mekenyan, O.; Kamenska, V. A Topological Approach to the Modeling of Polymer Properties (The TEMPO Method). *J. Math. Chem.* **1992**, *11*, 107–132.
- Balaban, T.-S.; Balaban, A. T.; Bonchev, D. A Topological Approach to the Predicting of Properties of Infinite Polymers. VI. Rational Formulas for the Normalized Wiener Index and a Comparison with Index J. *Theochem* **2000**, (in press).
- Mekenyan, O.; Karabunarliev, S.; Bonchev, D. The Microcomputer OASIS System for Predicting the Biological Activity of Chemical Compounds. *Comput. Chem.* **1990**, *14*, 193–200.
- MolconnZ. <http://www.eslc.vabiotech.com/molconn/manuals/>.
- Polansky, O. E.; Bonchev, D. The Theory of Wiener Number. *MATCH* **1986**, *21*, 133–186.
- Polansky, O. E.; Bonchev, D. Theory of the Wiener Number of Graphs. II. Transfer Graphs and Some of Their Metric Properties. *MATCH* **1990**, *25*, 3–40.
- Rouvray, D. H. The Role of Topological Distance Matrix in Chemistry. In *Mathematical and Computational Concepts in Chemistry*; Trinajstić, N., Ed.; Ellis Horwood: Chichester, U.K., 1986; pp 295–306.
- Gutman, I.; Yeh, Y.-N.; Lee, S.-L.; Luo, Y.-L. Some Results in the Theory of the Wiener Number. *Ind. J. Chem.* **1993**, *32A*, 651–661.
- Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Number: Development and Application. *Croat. Chem. Acta* **1995**, *68*, 105–129.
- Tratch, S. S.; Stankevich, M. I.; Zefirov, N. S. Combinatorial Models and Algorithms in Chemistry. The Expanded Wiener Number – A Novel Topological Index. *J. Comput. Chem.* **1990**, *11*, 899–908.
- MATCH (Communications in Mathematical Chemistry and in Computer Chemistry)*; 1997; Vol. 35, *Fifty Years of the Wiener Number*.
- Randić, M. Molecular Descriptors for Structure–Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478–483.
- Klein, D. J.; Lukovits, G.; Gutman, I. On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 50–52.
- Gutman, I. Formula for the Wiener Number of Trees and Its Extension to Graphs Containing Cycles. *Graph Theory Notes, NYAS* **1994**, *27*, 9–15.
- Klein, D. J.; Randić, M. Resistance Distance. *J. Math. Chem.* **1993**, *12*, 81–95.
- Bonchev, D.; Balaban, A. T.; Liu, X.; Klein, D. J. Molecular Cyclicity and Centricity of Polycyclic Graphs. I. Cyclicity Based on Resistance Distances or Reciprocal Distances. *Int. J. Quantum Chem.* **1994**, *50*, 1–20.
- Zhu, H.-Y.; Klein, D. J.; Lukovits, I. Extensions of the Wiener Number. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 420–428.
- Mohar, B. The Laplacian Spectrum of Graphs. In *Graph Theory, Combinatorics, and Applications*; Alavi, Y., Chartrand, C., Ollermann, O. R., Schwenk, A. J., Eds.; John Wiley & Sons: New York 1991; pp 871–898.
- Trinajstić, N.; Babić, D.; Nikolić, S.; Playšić, D.; Amić, D.; Mihalić, Z. The Laplacian Matrix in Chemistry. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 368–376.
- Mohar, B.; Babić, D.; Trinajstić, A. Novel Definition of the Wiener Number for Trees. *N. J. Chem. Inf. Comput. Sci.* **1993**, *33*, 153–154.
- Bonchev, D. Kolmogorov's Information, Shannon's entropy, and Topological Complexity of Molecules. *Bulg. Chem. Commun.* **1995**, *28*, 567–582.
- Bonchev, D. Novel Indices for the topological complexity of molecules. *SAR/QSAR Envir. Res.* **1997**, *7*, 23–44.
- Bonchev, D. Overall Connectivity and Topological Complexity: A New Tool for QSPR/QSAR. In *Topological Indices and Related Descriptors in QSAR and QSPR*. Devillers, J., Balaban, A. T., Eds.; Gordon and Breach: The Netherlands, 1999; pp 361–401.
- Bonchev, D. Overall Connectivities/Topological Complexities: A New Powerful Tool for QSPR/QSAR. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 934–941.
- Smolenski, E. A. Application of the Theory of Graphs to Calculation of the Additive Structural Properties of Hydrocarbons. *Russ. J. Phys. Chem.* **1964**, *38*, 700–704.
- Gordon, M.; Kennedy, J. W. The Graph-Like State of Matter. Part 2. TCGI Schemes for the Thermodynamics of Alkanes and the Theory of Inductive Inference. *J. Chem. Soc., Faraday Trans. 2* **1973**, *69*, 484–504.
- Klein, D. J. Chemical Graph-Theoretic Cluster-Expansions. *Int. J. Quantum Chem.* **1986**, *S20*, 153–171.
- Kier, L. B.; Hall, L. H.; Murray, W. J.; Randić, M. Molecular Connectivity. I: Relationship to Nonspecific Local Anesthesia. *J. Pharm. Sci.* **1975**, *64*, 1971–1974.
- Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York 1976.
- Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Analysis*; Research Studies Press: Chichester, 1986.
- Nikolić, S.; Tolić, I. M.; Trinajstić, N.; Baučić, I. On the Zagreb Indices As Complexity Indices. *Croat. Chem. Acta* **2000**, *73*, 909–921.

- (54) Gutman, I.; Rušćić, B.; Trinajstić, N.; Wilcox, C. W., Jr. Graph Theory and Molecular Orbitals. 12. Acyclic Polyenes. *J. Chem. Phys.* **1975**, *62*, 3399–3405.
- (55) Bertz, S. H. The First General Index of Molecular Complexity. *J. Am. Chem. Soc.* **1981**, *103*, 3599–3601.
- (56) Bertz, S. H. A Mathematical Model of Molecular Complexity. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: New York, 1983; pp 206–221.
- (57) Bonchev, D.; Polansky, O. E. On the Topological Complexity of Chemical Systems. In *Graph Theory and Topology in Chemistry*; King, R. B., Rouvray, D. H., Eds.; Elsevier: Amsterdam, 1987; pp 126–158.
- (58) Bonchev, D. The Problems of Computing Molecular Complexity. In *Computational Chemical Graph Theory*; Rouvray, D. H., Ed.; Nova Publications: New York, 1990; pp 34–67.
- (59) Bertz, S. H.; Sommer, T. J. Rigorous Mathematical Approaches to Strategic Bonds and Synthetic Analysis Based on Conceptually Simple New Complexity Indices. *Chem. Commun.* **1997**, 2409–2410.
- (60) Rücker, G.; Rücker, C. Walk Counts, Labyrinthicity, and Complexity of Acyclic and Cyclic Graphs and Molecules. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 99–106.
- (61) Needham, D. E.; Wei, I.-C.; Seybold, P. G. Molecular Modeling of the Physical Properties of the Alkanes. *J. Am. Chem. Soc.* **1988**, *110*, 4186–4194.
- (62) *Selected Values of Properties of Hydrocarbons and Related Compounds*; Research Project 44; American Petroleum Institute: Pittsburgh, Pennsylvania, 1977.
- (63) Garbalena, M.; Herndon, W. C. Optimum Graph-Theoretical Models for Enthalpic Properties of Alkanes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 37–42.
- (64) Bonchev, D.; Trinajstić, N. Overall Molecular Descriptors. 3. Overall Zagreb Indices. *SAR/QSAR Envir. Res.*, in press.
- (65) Randić, M.; Trinajstić, N. Viewpoint 4. Comparative Structure–Property Studies: The Connectivity Basis. *J. Mol. Struct. (THEOCHEM)* **1993**, *284*, 209–221.
- (66) Randić, M.; Trinajstić, N. In Search for Graph Invariants of Chemical Interest. *J. Mol. Struct. (THEOCHEM)* **1993**, *300*, 551–572.
- (67) Randić, M. Generalized Molecular Descriptors. *J. Math. Chem.* **1991**, *7*, 155–168.
- (68) Estrada, E.; Rodriguez, L. Edge-Connectivity Indices in QSPR/QSAR Studies. 1. Comparison to Other Topological Indices in QSPR Studies. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1037–1041.
- (69) Ivanciuc, O., private communication.
- (70) Estrada, E.; Edge-Connectivity Indices in QSPR/QSAR Studies. 2. Accounting for Long-Range Bond Contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1042–1048.
- (71) Rücker, G.; Rücker, C. On Finding Nonisomorphic Connected Subgraphs and Distinct Molecular Substructures. *J. Chem. Inf. Comput. Sci.* **2000**, in press.

CI000104T