Topological Indices Based on Vertex, Distance, and Ring: On the Boiling Points of Paraffins and Cycloalkanes

Chenzhong Cao* and Hua Yuan

Department of Chemistry, Xiangtan Normal University, Xiangtan 411201, People's Republic of China

Received December 9, 2000

Vertex, distance, and ring (in cyclic compounds) are three essential structure elements of a molecular graph, based on which three new topological indices VDI, OEI, and RDI are proposed. Multiple regression analysis was carried out against the boiling points of 343 hydrocarbons(160 paraffins and 183 cycloalkanes) with VDI, OEI, RDI, and $N^{2/3}$(N is the number of carbon atoms) together, and a good QSPR model was obtained: $\ln(577 - \mathrm{Bp}) = 6.729609 - 0.154107N^{2/3} + 2.285632 \times 10^{-2}\mathrm{VDI} - 7.921410 \times 10^{-3}\mathrm{OEI} - 1.821962 \times 10^{-2}\mathrm{RDI}$ ($F = 6455.09$, $r = 0.9935$, rms = 6.44 °C, $n = 343$).

1. INTRODUCTION

As we know, the physicochemical and biochemical properties of compounds depend strongly on their molecular structures. For decades, people have been trying to quantify molecular structures of compounds on the basis of graph theory and relate it to their properties. As yet, much more than 100 topological indices, such as Randic index[1], Balaban index,[2] and Schultz index,[3] have been developed since Wiener reported the first well-known graph-theoretical descriptor, Wiener index.[4] Most of these indices were derived from distance matrix or vertex adjacency as well as edge adjacency matrices. But a molecular graph includes more structure elements than both distance and vertex. So a more overall consideration should be taken into molecular structure while defining topological indices.

In the earlier time, chemists tended to relate the properties of compounds with single structure descriptor.[5] Though its calculation is simple, the obtained result is not satisfactory because only one descriptor is not enough to express all the molecular structure information. To remedy this defect, chemists now are inclined to regress more structure descriptors against the properties of compounds, by which better correlation can be expected despite of the more complicated operation. For example, in the NBP (normal boiling point)-structure studies of alkanes, Liu[6] employed 10 structure descriptors and obtained an equation with high correlation coefficient and low standard deviation. For another example, Estrada[7] related the boiling points of cycloalkanes to molecular structures with six variables and got a good correlation result. It seems that the more variables that are employed, the better the result that can be attained. In fact, it is not the case. For instance, Cao[8] regressed only four variables against the boiling points of alkanes and alcohols and still obtained the approving result ($n = 328$, $r = 0.9990$, $s = 6.021$ °C). It implies that fewer descriptors may also get quite good QSPR results if the defined descriptors are elaborate.

In general, a good QSPR/QSAR equation should require a high correlation coefficient, low standard deviation with parameters as few as possible. But how many parameters are acceptable and what are they? For a long time, chemists have been concentrating on solving this problem. However it is not easy to give a satisfactory answer. Let us take a look into the molecular graph: the vertex, distance, and ring (in cyclic compounds) are three essential structure elements which can characterize most of the molecular structure information**.** The variables derived from the above three structure elements will improve the correlation between the properties of compounds and their molecular structures. In this paper we presented three new topological indices, OEI (odd−even index), VDI (vertex degree-distance index), and RDI (ring degree-distance index), and then carried out multiple regression analysis with these indices (plus carbon atom number N) against the boiling points of paraffins and cycloalkanes (including polycyclic and spirocycloalkanes with carbon atoms up to 10). A good result was obtained.

2. CALCULATIONS

In the hydrogen-suppressed graph, distance, D_{ij}, is the length of the shortest path between vertex i and j; vertex degree, V_i, is the number of vertices adjacent to vertex i; and ring degree, R_i, is the minimum number of vertices (adjacent to the ith vertex) that must be removed to transform the ith vertex into an acyclic one. Each of them are represented by the corresponding matrix, i.e., distance matrix-($N \times N$)D, vertex degree vector($1 \times N$)V, and ring degree vector($1 \times N$)R, respectively. Based on D, V, and R, we defined three new topological indices.

(1) Odd−even index (OEI)

$$\mathrm{OEI} = \sum_{i=1}^{N} \sum_{j \neq i}^{N} [(-1)^{D_{ij}-1} S] \qquad (1)$$

where N is the number of vertices in molecular graph. S is the derivative matrix from distance matrix D, whose elements are the squares of the reciprocal distances $(D_{ij})^{-2}$, i.e., $S = [1/D_{ij}^2]$ (when $i = j$, let $1/D_{ij}^2 = 0$). It means that the interaction between vertex i and j is in proportional to $(D_{ij})^{-2}$. However, the behavior of odd distance is different from that

* Corresponding author fax: 0086-0732-8291001; e-mail: czcao@xtnu.edu.cn.

10.1021/ci000467t CCC: $20.00
Published on Web 07/03/2001

of even distance between vertex i and j, thus a coefficient $(-1)^{D_{ij}-1}$ is introduced into the expression so as to obtain the positive and negative values for odd distance and even distance D_{ij} item, respectively.

(2) Vertex degree-distance index (VDI): The interaction of vertex i against j is determined not only by the distance between i and j, but also by their vertex degrees. So we defined VDI as

$$\mathrm{VDI} = \left(\prod_{i=1}^{N} f_i\right)^{1/N} \tag{2}$$

f_i is the elements of vector$(1 \times N)$ VS obtained by V-multiply-S

$$VS = [f_1, f_2, \cdots, f_N] \tag{3}$$

(3) Ring degree-distance index (RDI): Because of the rigidity of the ring, the freedom of vertex in ring is smaller than that in the chain. Thus, another index RDI was proposed to distinguish the cyclic and acyclic hydrocarbons.

$$\mathrm{RDI} = \left(\prod_{i=1}^{N} g_i\right)^{1/N} \tag{4}$$

g_i is the elements of vector $(1 \times N)$ RS obtained by R-multiply-S:

$$RS = [g_1, g_2, \cdots, g_N] \tag{5}$$

It should be pointed out that we employ eqs 2 and 3 other than expressions $(\sum_{i=1}^{N} f_i)/N$ and $(\sum_{i=1}^{N} g_i)/N$ to calculate VDI and RDI because the values obtained from the former are more distinguishable for isomers than that from the latter.

We take 1-methylbicyclo[1.1.0]butane for example to compute the indices OEI, VDI, and RDI. Figure 1 is the hydrogen-suppressed graph of 1-methylbicyclo[1.1.0]butane (where the digits are the random numberings of each vertex).

Its calculation steps are as follows:

(a) The distance matrix D and its derivative matrices

$$D = \begin{bmatrix} 0 & 1 & 2 & 1 & 2 \\ 1 & 0 & 1 & 1 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 1 \\ 2 & 2 & 2 & 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & 1 & 0.25 & 1 & 0.25 \\ 1 & 0 & 1 & 1 & 0.25 \\ 0.25 & 1 & 0 & 1 & 0.25 \\ 1 & 1 & 1 & 0 & 1 \\ 0.25 & 0.25 & 0.25 & 1 & 0 \end{bmatrix}$$

$$(-1)^{D_{ij}-1} S = \begin{bmatrix} 0 & 1 & -0.25 & 1 & -0.25 \\ 1 & 0 & 1 & 1 & -0.25 \\ -0.25 & 1 & 0 & 1 & -0.25 \\ 1 & 1 & 1 & 0 & 1 \\ -0.25 & -0.25 & -0.25 & 1 & 0 \end{bmatrix}$$

(b) The vertex-degree vector V

$$V = [2, 3, 2, 4, 1]$$

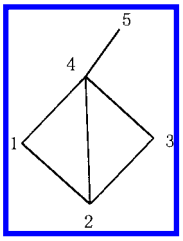


Figure 1. Structure of 1-methylbicyclo[1.1.0]butane.

Table 1. Interrelations of $N^{2/3}$, OEI, RDI, and VDI

	$N^{2/3}$	OEI	RDI	VDI
$N^{2/3}$	1.0000			
OEI	0.8800	1.0000		
RDI	0.3452	6.7246×10^{-3}	1.0000	
VDI	4.3218×10^{-2}	0.2578	0.8095	1.0000

Table 2. Regression Results with Different Variables

variables	F	r	rms
	Single Variable		
$N^{2/3}$	6871.39	0.9761	11.39
VDI	0.95	0.0527	51.74
OEI	2222.96	0.9311	19.79
RDI	22.88	0.2507	50.57
	Two Variables		
$N^{2/3}$, VDI	3434.06	0.9761	11.40
$N^{2/3}$, RDI	4205.78	0.9804	10.74
$N^{2/3}$, OEI	6860.51	0.9878	8.75
VDI, RDI	57.44	0.5025	42.73
VDI, OEI	1611.90	0.9511	17.40
RDI, OEI	2369.34	0.9659	14.29
	Three Variables		
$N^{2/3}$, VDI, RDI	6485.97	0.9914	7.32
$N^{2/3}$, VDI, OEI	5646.91	0.9901	7.73
$N^{2/3}$, RDI, OEI	4635.35	0.9880	8.61
VDI, RDI, OEI	1612.63	0.9667	14.02
	Four Variables		
$N^{2/3}$, VDI, RDI, OEI	6455.09	0.9935	6.44

(c) The ring-degree vector R

$$R = [1, 2, 1, 2, 0]$$

(d) The V-multiply-S vector VS

$$VS = [7.75, 8.25, 7.75, 8, 5.75]$$

(e) The R-multiply-S vector RS

$$RS = [4.25, 4, 4.25, 4, 3]$$

(f) The values of OEI, VDI, and RDI

$$\mathrm{OEI} = 1 \times 12 - 0.25 \times 8 = 10.000000$$

$$\mathrm{VDI} = (7.75 \times 8.25 \times 7.75 \times 8 \times 5.75)^{1/5} = 7.439830$$

$$\mathrm{RDI} = (4.25 \times 4 \times 4.25 \times 4 \times 3)^{1/5} = 3.869045$$

All the calculations can be performed quickly by computer. Likewise, the values of OEI, VDI, and RDI for 160 paraffins and 183 cycloalkanes were calculated and listed in Table 3. Obviously, the RDI index is zero for paraffins because of no cyclic atoms in their molecules.

3. RESULT AND DISCUSSION

In this section, we take 160 paraffins and 183 cycloalkanes (see Figure 3) with known boiling points (Bp) as the data

Table 3. Variables $N^{2/3}$, VDI, OEI, and RDI and Boiling Points (Bp) of 343 Paraffins and Cycloalkanes

no.	compound[a]	$N^{2/3}$	VDI	OEI	RDI	Bp_{exp}	Bp_{calc}	Δ
1	1	1.0000	0.0000	0.0000	0.0000	−161.50	−140.31	−21.19
2	2	1.5874	1.0000	2.0000	0.0000	−88.60	−82.84	−5.76
3	3	2.0801	2.1634	3.5000	0.0000	−42.10	−43.66	1.56
4	4	2.5198	2.9131	5.2222	0.0000	−0.50	−5.02	4.52
5	2m3	2.5198	3.3677	4.5000	0.0000	−11.70	−14.47	2.77
6	5	2.9240	3.4010	6.8194	0.0000	36.00	30.95	5.05
7	2m4	2.9240	3.8514	6.4444	0.0000	27.80	23.66	4.14
8	22mm3	2.9240	4.5895	5.0000	0.0000	9.50	7.77	1.73
9	6	3.3019	3.7591	8.4967	0.0000	68.70	64.46	4.24
10	2m5	3.3019	4.1451	7.9167	0.0000	60.30	57.54	2.76
11	3m5	3.3019	4.1925	8.2639	0.0000	63.30	58.40	4.90
12	22mm4	3.3019	4.8758	7.1667	0.0000	49.70	45.64	4.06
13	23mm4	3.3019	4.6112	7.8889	0.0000	58.00	51.86	6.14
14	7	3.6593	4.0367	10.1183	0.0000	98.50	95.07	3.43
15	2m6	3.6593	4.3801	9.6739	0.0000	90.00	89.56	0.44
16	3m6	3.6593	4.4146	9.8161	0.0000	92.00	89.72	2.28
17	22mm5	3.6593	5.0050	8.5139	0.0000	79.20	77.98	1.22
18	23mm5	3.6593	4.8234	9.5833	0.0000	89.80	84.24	5.56
19	24mm5	3.6593	4.7465	8.8889	0.0000	80.50	82.39	−1.89
20	33mm5	3.6593	5.0972	9.2083	0.0000	86.10	79.67	6.43
21	3e5	3.6593	4.4611	9.9583	0.0000	93.50	89.76	3.74
22	223mmm4	3.6593	5.4937	8.8333	0.0000	80.90	73.65	7.25
23	8	4.0000	4.2597	11.7808	0.0000	125.70	123.40	2.30
24	2m7	4.0000	4.1290	11.2400	0.0000	117.60	122.81	−5.21
25	3m7	4.0000	4.6001	11.5178	0.0000	118.00	118.90	−0.90
26	4m7	4.0000	4.5940	11.3128	0.0000	117.70	118.22	−0.52
27	22mm6	4.0000	5.1237	10.3511	0.0000	106.80	109.08	−2.28
28	23mm6	4.0000	4.9583	11.2156	0.0000	115.60	114.03	1.57
29	24mm6	4.0000	4.9306	10.8683	0.0000	109.40	113.05	−3.65
30	25mm6	4.0000	4.9008	10.9311	0.0000	109.10	113.59	−4.49
31	33mm6	4.0000	5.1973	10.6356	0.0000	112.00	109.35	2.65
32	34mm6	4.0000	4.9992	11.3578	0.0000	117.70	114.12	3.58
33	3e6	4.0000	4.6388	11.5906	0.0000	118.50	118.76	−0.26
34	223mmm5	4.0000	5.5893	10.4028	0.0000	109.80	104.27	5.53
35	224mmm5	4.0000	5.4861	9.3611	0.0000	99.20	101.47	−2.27
36	233mmm5	4.0000	5.6310	10.7500	0.0000	114.80	105.12	9.68
37	234mmm5	4.0000	5.3465	10.7778	0.0000	113.50	108.28	5.22
38	2m3e5	4.0000	5.0075	11.1528	0.0000	115.60	113.28	2.32
39	3m3e5	4.0000	5.2877	11.1250	0.0000	118.20	110.19	8.01
40	2233mmmm4	4.0000	6.2627	10.0000	0.0000	106.50	95.40	11.10
41	9	4.3267	4.4434	13.4120	0.0000	150.80	149.42	1.38
42	2m8	4.3267	4.7280	12.9433	0.0000	142.80	145.03	−2.23
43	3m8	4.3267	4.7549	13.1247	0.0000	144.00	145.38	−1.38
44	4m8	4.3267	4.7479	13.0553	0.0000	142.40	145.21	−2.81
45	22mm7	4.3267	5.2272	11.8617	0.0000	132.70	136.31	−3.61
46	23mm7	4.3267	5.0775	12.8617	0.0000	140.50	141.28	−0.78
47	24mm7	4.3267	5.0474	12.3094	0.0000	133.50	139.67	−6.17
48	25mm7	4.3267	5.0562	12.7194	0.0000	136.00	141.00	−5.00
49	26mm7	4.3267	5.0290	12.3061	0.0000	134.00	139.84	−5.84
50	33mm7	4.3267	5.2906	12.4172	0.0000	137.30	137.61	−0.31
51	34mm7	4.3267	5.1101	12.9344	0.0000	140.60	141.21	−0.61
52	35mm7	4.3267	5.0849	12.7922	0.0000	136.00	140.97	−4.97
53	44mm7	4.3267	5.2851	12.0072	0.0000	135.20	136.24	−1.04
54	3e7	4.3267	4.7903	13.2367	0.0000	143.00	145.42	−2.42
55	4e7	4.3267	4.7880	13.1672	0.0000	142.10	145.20	−3.10
56	223mmm6	4.3267	5.6355	12.1150	0.0000	133.60	133.07	0.53
57	224mmm6	4.3267	5.5843	11.4206	0.0000	126.50	131.14	−4.64
58	225mmm6	4.3267	5.5546	11.6883	0.0000	124.00	132.39	−8.39
59	233mmm6	4.3267	5.6716	12.2572	0.0000	137.70	133.20	4.50
60	234mmm6	4.3267	5.4611	12.6322	0.0000	139.00	136.64	2.36
61	235mmm6	4.3267	5.3953	12.3478	0.0000	131.30	136.31	−5.01
62	244mmm6	4.3267	5.6199	11.5628	0.0000	130.70	131.28	−0.58
63	334mmm6	4.3267	5.7164	12.3994	0.0000	140.50	133.25	7.25
64	2m3e6	4.3267	5.1211	12.8650	0.0000	138.00	140.86	−2.86
65	2m4e6	4.3267	5.0925	12.7228	0.0000	133.80	140.65	−6.85
66	3m3e6	4.3267	5.3702	12.6322	0.0000	140.60	137.56	3.04
67	3m4e6	4.3267	5.1594	13.0072	0.0000	140.40	140.97	−0.57
68	2233mmmm5	4.3267	6.3158	11.7917	0.0000	140.20	124.96	15.24
69	2234mmmm5	4.3267	6.0175	11.4722	0.0000	133.00	126.89	6.11
70	2244mmmm5	4.3267	6.1265	9.7083	0.0000	122.30	119.42	2.88
71	2334mmmm5	4.3267	6.0825	12.1667	0.0000	141.50	128.69	12.81
72	22mm3e5	4.3267	5.6903	11.8472	0.0000	133.80	131.57	2.23
73	23mm3e5	4.3267	5.7641	12.5417	0.0000	142.00	133.27	8.73

Table 3 (Continued)

no.	compound[a]	$N^{2/3}$	VDI	OEI	RDI	Bp_{exp}	Bp_{calc}	Δ
74	24mm3e5	4.3267	5.4753	12.2222	0.0000	136.70	135.07	1.63
75	33ee5	4.3267	5.4625	12.9167	0.0000	145.00	137.62	7.38
76	10	4.6416	4.5978	15.0680	0.0000	174.10	173.56	0.54
77	2m9	4.6416	4.8603	14.5433	0.0000	166.80	169.44	−2.64
78	3m9	4.6416	4.8854	14.7968	0.0000	167.80	170.02	−2.22
79	4m9	4.6416	4.8789	14.6309	0.0000	165.70	169.55	−3.85
80	5m9	4.6416	4.8775	14.7665	0.0000	165.10	170.00	−4.90
81	22mm8	4.6416	5.3165	13.6058	0.0000	155.00	162.10	−7.10
82	23mm8	4.6416	5.1804	14.5094	0.0000	164.00	166.33	−2.33
83	24mm8	4.6416	5.1517	14.9027	0.0000	153.00	167.88	−14.88
84	25mm8	4.6416	5.1545	14.2977	0.0000	157.00	165.89	−8.89
85	26mm8	4.6416	5.1625	14.2316	0.0000	158.50	165.60	−7.10
86	27mm8	4.6416	5.1368	14.1466	0.0000	160.00	165.56	−5.56
87	33mm8	4.6416	5.3739	13.9686	0.0000	161.20	162.74	−1.54
88	34mm8	4.6416	5.2094	14.6214	0.0000	166.00	166.43	−0.43
89	35mm8	4.6416	5.1821	14.2741	0.0000	159.00	165.55	−6.55
90	36mm8	4.6416	5.1886	14.5486	0.0000	159.50	166.39	−6.89
91	44mm8	4.6416	5.3668	13.8297	0.0000	160.00	162.36	−2.36
92	45mm8	4.6416	5.2058	14.5519	0.0000	162.40	166.23	−3.83
93	3e8	4.6416	4.9189	14.8844	0.0000	168.00	169.99	−1.99
94	4e8	4.6416	4.9170	14.8541	0.0000	164.00	169.91	−5.91
95	223mmm7	4.6416	5.6842	13.7056	0.0000	158.00	158.93	−0.93
96	224mmm7	4.6416	5.6339	12.8061	0.0000	147.70	156.42	−8.72
97	225mmm7	4.6416	5.6424	13.4211	0.0000	147.00	158.38	−11.38
98	226mmm7	4.6416	5.6159	12.8722	0.0000	148.20	156.81	−8.61
99	233mmm7	4.6416	5.7158	13.9833	0.0000	160.00	159.54	0.46
100	234mmm7	4.6416	5.5243	14.1533	0.0000	163.00	161.93	1.07
101	235mmm7	4.6416	5.5000	14.2161	0.0000	159.70	162.36	−2.66
102	236mmm7	4.6416	5.4715	14.0078	0.0000	155.70	161.95	−6.25
103	244mmm7	4.6416	5.6651	12.8789	0.0000	152.00	156.36	−4.36
104	245mmm7	4.6416	5.5003	14.0111	0.0000	157.00	161.68	−4.68
105	246mmm7	4.6416	5.4401	13.2506	0.0000	145.00	159.75	−14.75
106	255mmm7	4.6416	5.6739	13.6989	0.0000	152.80	159.00	−6.20
107	334mmm7	4.6416	5.7544	14.0561	0.0000	164.00	159.41	4.59
108	335mmm7	4.6416	5.7045	13.5667	0.0000	155.70	158.27	−2.57
109	344mmm7	4.6416	5.7542	13.8511	0.0000	164.00	158.74	5.26
110	345mmm7	4.6416	5.5628	14.4311	0.0000	164.00	162.47	1.53
111	2m3e7	4.6416	5.2211	14.4556	0.0000	166.00	165.78	0.22
112	2m4e7	4.6416	5.1935	14.2439	0.0000	160.00	165.35	−5.35
113	2m5e7	4.6416	5.1965	14.3828	0.0000	159.70	165.77	−6.07
114	3m3e7	4.6416	5.4457	14.3583	0.0000	163.80	163.34	0.46
115	3m4e7	4.6416	5.2563	14.6639	0.0000	167.00	166.12	0.88
116	3m5e7	4.6416	5.2258	14.5911	0.0000	160.00	166.17	−6.17
117	4m3e7	4.6416	5.2532	14.5283	0.0000	167.00	165.71	1.29
118	4m4e7	4.6416	5.4458	14.0839	0.0000	166.00	162.44	3.56
119	4p7	4.6416	4.9170	14.6883	0.0000	161.80	169.38	−7.58
120	4ip7	4.6416	5.2210	14.5217	0.0000	159.50	165.99	−6.49
121	2233mmmm6	4.6416	6.2890	13.3789	0.0000	160.00	152.01	7.99
122	2234mmmm6	4.6416	6.0652	13.4067	0.0000	155.00	154.27	0.73
123	2235mmmm6	4.6416	5.9987	13.3272	0.0000	148.70	154.64	−5.94
124	2244mmmm6	4.6416	6.1958	11.9900	0.0000	153.00	148.22	4.78
125	2245mmmm6	4.6416	5.9764	12.9800	0.0000	147.90	153.70	−5.80
126	2255mmmm6	4.6416	6.1318	12.5256	0.0000	137.00	150.66	−13.66
127	2334mmmm6	4.6416	6.1259	13.8961	0.0000	164.00	155.32	8.68
128	2335mmmm6	4.6416	6.0316	13.2644	0.0000	153.00	154.12	−1.12
129	2344mmmm6	4.6416	6.1026	13.5489	0.0000	162.00	154.38	7.62
130	2345mmmm6	4.6416	5.8603	13.9867	0.0000	158.00	158.17	−0.17
131	3344mmmm6	4.6416	6.3744	13.6633	0.0000	170.50	152.14	18.36
132	22mm3e6	4.6416	5.7351	13.6394	0.0000	159.00	158.22	0.78
133	22mm4e6	4.6416	5.6808	13.3550	0.0000	147.00	157.79	−10.79
134	23mm3e6	4.6416	5.9204	15.0178	0.0000	169.00	161.01	7.99
135	23mm4e6	4.6416	5.5749	14.3617	0.0000	164.00	162.13	1.87
136	24mm3e6	4.6416	5.5804	14.1567	0.0000	164.00	161.40	2.60
137	24mm4e6	4.6416	5.7488	13.6394	0.0000	158.00	158.09	−0.09
138	25mm3e6	4.6416	5.5131	14.0772	0.0000	157.00	161.78	−4.78
139	33mm4e6	4.6416	5.8100	13.9239	0.0000	165.00	158.45	6.55
140	34mm3e6	4.6416	5.8430	14.2711	0.0000	170.00	159.28	10.72
141	33ee6	4.6416	5.5336	14.5039	0.0000	166.30	162.99	3.31
142	34ee6	4.6416	5.3034	14.7367	0.0000	162.00	165.92	−3.92
143	2m3ip6	4.6416	5.5425	14.0144	0.0000	163.00	161.29	1.70
144	22334mmmmm5	4.6416	6.6909	13.0833	0.0000	166.00	147.08	18.92
145	22344mmmmm5	4.6416	6.6071	12.0417	0.0000	159.30	144.35	14.95
146	223mmm3e5	4.6416	6.3864	13.4583	0.0000	168.00	151.33	16.67
147	224mmm3e5	4.6416	6.0856	12.7917	0.0000	155.30	152.01	3.29

Table 3 (Continued)

no.	compound[a]	$N^{2/3}$	VDI	OEI	RDI	Bp_{exp}	Bp_{calc}	Δ
148	234mmm3e5	4.6416	6.1787	13.8333	0.0000	169.40	154.60	14.80
149	2m33ee5	4.6416	5.8975	14.2083	0.0000	174.00	158.55	15.45
150	24mm3ip5	4.6416	5.4747	13.1667	0.0000	157.00	159.14	−2.14
151	11	4.9461	4.7296	16.7039	0.0000	195.90	195.86	0.04
152	12	5.2415	4.8436	18.3563	0.0000	216.30	216.62	−0.32
153	13	5.5288	4.9432	19.9949	0.0000	235.40	235.90	−0.50
154	14	5.8088	5.0312	21.6453	0.0000	253.70	253.90	−0.20
155	15	6.0822	5.1095	23.2855	0.0000	270.60	270.68	−0.08
156	16	6.3496	5.1796	24.9346	0.0000	287.00	286.40	0.60
157	17	6.6115	5.2429	26.5759	0.0000	301.80	301.10	0.70
158	18	6.8683	5.3002	28.2241	0.0000	316.10	314.90	1.20
159	19	7.1204	5.3525	29.8661	0.0000	329.70	327.85	1.85
160	20	7.3681	5.4004	31.5136	0.0000	343.00	340.03	2.97
161	c3	2.0801	4.0000	6.0000	2.0000	−32.80	−34.87	2.07
162	c4	2.5198	4.5000	7.0000	2.2500	12.60	5.83	6.77
163	1mc3	2.5198	4.8455	7.0000	1.8612	0.70	−2.79	3.49
164	bc110b	2.5198	6.7454	9.5000	4.1231	8.00	7.31	0.68
165	c5	2.9240	5.0000	7.5000	2.5000	49.30	38.75	10.55
166	11mc3	2.9240	5.8730	7.5000	1.7826	20.60	20.68	−0.08
167	12mc3	2.9240	5.4547	8.2222	1.7826	32.60	29.12	3.48
168	1ec3	2.9240	5.0057	8.9444	1.4147	35.90	34.18	1.72
169	1mc4	2.9240	5.1808	8.2222	2.1046	36.30	35.72	0.58
170	s22p	2.9240	7.1895	10.0000	3.5947	39.00	33.18	5.82
171	mbc110b	2.9240	7.4398	10.0000	3.8690	33.50	32.79	0.71
172	bc210p	2.9240	6.8770	10.0000	4.0889	46.00	41.89	4.11
173	bc111p	2.9240	6.8989	10.0000	4.0696	36.00	41.44	−5.44
174	c6	3.3019	5.2222	9.6667	2.6111	80.70	76.32	4.38
175	123mc3	3.3019	5.9512	9.6667	1.7321	63.00	59.69	3.31
176	1e1mc3	3.3019	5.9087	9.6667	1.4286	57.00	57.32	−0.32
177	1e2mc3	3.3019	5.5058	10.0417	1.4286	63.00	63.61	−0.61
178	1pc3	3.3019	5.1137	10.4167	1.0497	69.00	66.19	2.81
179	1ipc3	3.3019	5.6258	10.3889	1.1783	58.30	61.27	−2.97
180	112mc3	3.3019	6.3022	8.9444	1.7321	52.60	52.53	0.07
181	12mc4	3.3019	5.7394	9.6667	2.0129	62.00	64.81	−2.81
182	13mc4	3.3019	5.6499	9.3194	2.0129	59.00	64.45	−5.45
183	11mc4	3.3019	6.0657	8.9444	2.0129	53.60	58.02	−4.42
184	1ec4	3.3019	5.2935	10.0417	1.6749	70.70	68.39	2.31
185	1mc5	3.3019	5.5724	8.9444	2.3494	71.80	66.97	4.83
186	bcpr	3.3019	6.7458	12.8889	2.7760	76.00	73.16	2.84
187	13mbcb	3.3019	7.9830	10.7222	3.7084	55.00	58.61	−3.61
188	s23hx	3.3019	7.1692	11.4444	3.5846	69.50	69.93	−0.43
189	bc220hx	3.3019	6.9460	11.4444	4.0633	83.00	76.89	6.11
190	bc310hx	3.3019	7.0281	10.7222	4.0852	81.00	73.27	7.73
191	mbc210p	3.3019	7.5259	10.7222	3.9069	60.50	65.85	−5.35
192	bc211hx	3.3019	7.1598	10.0000	4.1457	71.00	69.42	1.58
193	c7	3.6593	5.4444	12.0556	2.7222	118.40	110.61	7.79
194	1sbc3	3.6593	5.6938	12.0833	0.9365	90.30	92.54	−2.24
195	1122mc3	3.6593	7.0045	9.8889	1.6968	76.00	76.04	−0.04
196	1123mc3	3.6593	6.6900	10.6111	1.6969	78.00	82.46	−4.46
197	1tbc3	3.6593	6.4284	11.3333	1.0340	80.50	82.28	−1.78
198	11ec3	3.6593	5.9675	11.7083	1.2196	88.60	90.56	−1.96
199	12ec3	3.6593	5.5620	11.9411	1.2196	90.00	95.94	−5.94
200	1e23mc3	3.6593	5.9457	11.3611	1.4386	91.00	91.41	−0.41
201	1m2pc3	3.6593	5.5395	11.5939	1.1046	93.00	93.85	−0.85
202	1m1ipc3	3.6593	6.4073	11.3333	1.2196	81.50	84.18	−2.68
203	1bc3	3.6593	5.2234	12.1739	0.7892	98.00	96.77	1.23
204	1e3mc4	3.6593	5.6901	11.2189	1.6657	89.50	95.68	−6.18
205	1e2mc4	3.6593	5.7768	11.3611	1.6657	94.00	95.27	−1.27
206	1ipc4	3.6593	5.8221	11.3611	1.4228	92.70	92.64	0.06
207	1pc4	3.6593	5.3630	11.5939	1.2946	100.70	97.46	3.24
208	13mc5	3.6593	6.0090	10.2639	2.2475	91.30	93.63	−2.33
209	1ec5	3.6593	5.6422	10.6389	1.9319	103.50	96.33	7.17
210	12mc5	3.6593	6.0699	10.6111	2.2475	95.60	94.29	1.31
211	11mc5	3.6593	6.3538	9.8889	2.2475	87.90	88.36	−0.46
212	1mc6	3.6593	5.7200	10.9861	2.4730	101.00	101.51	−0.51
213	dcprm	3.6593	6.6367	13.8889	2.3998	102.00	101.85	0.15
214	11ms22p	3.6593	7.9764	12.3889	3.0282	78.00	86.86	−8.86
215	122mbcb	3.6593	8.1656	12.3889	3.2541	84.00	86.76	−2.76
216	s33h	3.6593	9.0714	12.7639	3.5947	96.50	81.08	15.42
217	s24h	3.6593	7.3218	12.3889	3.6609	98.50	99.68	−1.18
218	14mbc210p	3.6593	8.0749	11.6667	3.7818	74.00	89.68	−15.68
219	13mbc111p	3.6593	7.9659	11.3194	3.7202	71.50	89.00	−17.50
220	mbc310hx	3.6593	9.0489	11.6667	3.9490	92.00	80.22	11.78
221	6mbc310hx	3.6593	7.2602	12.0417	3.7597	103.00	99.89	3.11

Table 3 (Continued)

no.	compound[a]	$N^{2/3}$	VDI	OEI	RDI	Bp_{exp}	Bp_{calc}	Δ
222	2mbc310hx	3.6593	7.3264	12.3889	3.7381	100.00	100.30	−0.30
223	bc320h	3.6593	7.1288	12.3889	4.0974	110.50	105.54	4.96
224	mbc211hx	3.6593	7.7347	10.9444	3.9740	81.50	92.38	−10.88
225	tc221026h	3.6593	8.9764	12.0000	5.5476	106.00	96.55	9.45
226	bc221h	3.6593	7.3287	10.9444	4.1947	105.50	98.78	6.72
227	bc410h	3.6593	7.0013	13.1111	3.9912	116.00	108.69	7.31
228	bc311h	3.6593	7.1163	12.3889	4.0512	110.00	105.28	4.72
229	tc410024h	3.6593	8.6910	14.1667	5.3986	105.00	106.52	−1.52
230	tc410013h	3.6593	8.9775	13.4444	5.6588	107.50	102.97	4.53
231	tc410027h	3.6593	8.4629	14.8889	5.2045	110.00	109.99	0.01
232	tec410h	3.6593	10.2560	15.9444	6.7084	104.00	107.54	−3.54
233	tc311024h	3.6593	8.8149	13.4444	5.4827	107.00	103.21	3.79
234	tec320h	3.6593	10.4756	14.5000	6.8399	108.50	100.89	7.61
235	c8	4.0000	5.5694	13.2778	2.7847	149.00	137.97	11.03
236	11223mc3	4.0000	7.3293	11.7778	1.6709	100.50	105.01	−4.51
237	11m2pc3	4.0000	6.1775	12.2711	1.1477	105.90	114.68	−8.78
238	ib2mc3	4.0000	5.9744	12.6461	0.9109	110.00	116.20	−6.20
239	112m2ec3	4.0000	6.9210	11.8056	1.4461	104.50	107.58	−3.08
240	11m2ipc3	4.0000	6.6228	11.8333	1.2516	94.40	109.21	−14.81
241	1nepec3	4.0000	6.3372	11.8611	0.7229	106.00	107.86	−1.86
242	p2ec3	4.0000	5.5961	13.4378	0.9933	108.00	123.70	−15.70
243	b2mc3	4.0000	5.5918	13.2956	0.8552	124.00	122.09	1.91
244	1spec3	4.0000	5.7208	13.7156	0.7401	117.70	121.31	−3.61
245	5msbc3	4.0000	6.1506	13.2778	0.7883	115.50	115.62	−0.12
246	1pec3	4.0000	5.3212	13.7400	0.6061	128.00	124.44	3.56
247	1m12ec3	4.0000	6.2566	12.7606	1.2516	108.90	116.51	−7.61
248	1133mc4	4.0000	6.9429	10.3889	1.9039	86.00	105.99	−19.99
249	1234mc4	4.0000	6.5416	12.5278	1.9039	114.50	118.13	−3.63
250	12ec4	4.0000	5.8236	13.1356	1.4451	119.00	123.99	−4.99
251	1sbc4	4.0000	5.8745	13.1356	1.1592	123.00	121.10	1.90
252	p3mc4	4.0000	5.7092	12.7156	1.3305	117.40	122.72	−5.32
253	123mc5	4.0000	6.4688	12.1528	2.1739	117.00	119.78	−2.78
254	124mc5	4.0000	6.4124	11.8056	2.1739	115.00	119.11	−4.11
255	1e1mc5	4.0000	6.3768	11.8056	1.9044	121.50	117.23	4.27
256	1e2mc5	4.0000	6.0880	12.1806	1.9044	124.70	121.61	3.09
257	1e3mc5	4.0000	6.0265	12.0383	1.9044	121.00	121.74	−0.74
258	113mc5	4.0000	6.6696	11.0833	2.1739	104.90	113.77	−8.87
259	112mc5	4.0000	6.7733	11.7778	2.1739	114.00	115.22	−1.22
260	1pc5	4.0000	5.6741	12.2711	1.5407	131.00	123.23	7.77
261	1ipc5	4.0000	6.0981	11.8333	1.6682	126.40	118.28	8.12
262	14mc6	4.0000	6.1117	12.3856	2.3742	121.80	125.98	−4.18
263	13mc6	4.0000	6.1234	12.1806	2.3742	122.30	125.13	−2.83
264	12mc6	4.0000	6.1701	12.5278	2.3742	126.60	125.89	0.71
265	11mc6	4.0000	6.4157	11.8056	2.3742	119.50	120.75	−1.25
266	1ec6	4.0000	5.7813	12.7606	2.0874	131.80	128.36	3.44
267	1mc7	4.0000	5.8868	13.2500	2.5934	134.00	133.13	0.87
268	bcprm	4.0000	6.5825	15.9311	2.1212	129.00	131.68	−2.68
269	bcb	4.0000	6.8454	14.9133	2.9490	136.00	132.13	3.87
270	s34o	4.0000	7.3485	13.5833	3.6742	128.00	128.19	−0.19
271	1223mbcb	4.0000	8.6346	13.5556	3.2212	105.00	110.87	−5.87
272	2244mbcb	4.0000	8.1535	12.8889	2.9503	104.00	111.23	−7.23
273	33mbc310hx	4.0000	7.8133	12.8611	3.4201	115.00	118.68	−3.68
274	2mbc320h	4.0000	7.4084	13.9306	3.8030	130.50	129.86	0.64
275	bc330o	4.0000	7.2920	13.2083	4.1365	137.00	131.21	5.79
276	s25o	4.0000	7.2621	14.3056	3.6310	125.00	131.28	−6.28
277	1mbc410h	4.0000	7.5600	13.9306	3.8952	125.00	129.06	−4.06
278	7mbc410h	4.0000	7.2421	14.3056	3.7349	138.00	132.33	5.67
279	bc420o	4.0000	7.0977	14.3056	4.0175	133.00	136.07	−3.07
280	1mbc221h	4.0000	7.8877	12.1111	4.0543	117.00	120.47	−3.47
281	2mbc221h	4.0000	7.5717	12.4861	3.8820	125.00	123.68	1.32
282	7mbc221h	4.0000	7.6195	12.8333	3.9209	128.00	124.75	3.25
283	1mtc2210h	4.0000	9.3442	13.1667	5.2728	111.00	119.22	−8.22
284	ds2121o	4.0000	8.7535	15.3889	4.3767	103.00	125.94	−22.94
285	ds2022o	4.0000	8.9710	16.7778	4.4855	115.00	129.54	−14.54
286	tc3210o	4.0000	8.8162	14.6111	5.3879	136.00	130.79	5.21
287	3mtc2210h	4.0000	9.0429	13.8889	5.0622	120.50	123.22	−2.72
288	bc510o	4.0000	7.0172	15.3750	3.9425	141.00	139.99	1.01
289	tc510035o	4.0000	8.5201	16.4306	5.1907	142.00	138.57	3.43
290	tc510024o	4.0000	8.5428	16.7778	5.2149	149.00	139.74	9.26
291	tc3300o	4.0000	8.8808	13.8889	5.4327	125.00	127.94	−2.94
292	tec330o	4.0000	10.0389	14.4167	6.6195	137.50	127.64	9.86
293	14mbc211hx	4.0000	8.1835	11.7639	3.8499	91.00	114.38	−23.38
294	c9	4.3267	5.6944	14.3750	2.8472	175.00	162.43	12.57
295	1shxc3	4.3267	5.7586	15.3617	0.5895	143.00	147.76	−4.76

Table 3 (Continued)

no.	compound[a]	$N^{2/3}$	VDI	OEI	RDI	Bp_{exp}	Bp_{calc}	Δ
296	1123mc5	4.3267	7.0783	13.1944	2.1184	134.00	139.32	−5.32
297	1122mc5	4.3267	7.4014	13.1667	2.1184	135.00	135.97	−0.97
298	1133mc5	4.3267	7.2260	11.7778	2.1184	118.20	132.88	−14.68
299	1m1pc5	4.3267	6.3273	13.3128	1.5599	145.00	142.77	2.23
300	12m1ec5	4.3267	6.7674	13.5694	1.8832	142.50	141.84	0.66
301	13ec5	4.3267	6.0513	13.7572	1.6741	148.20	147.91	0.29
302	12ec5	4.3267	6.1199	13.8300	1.6741	150.50	147.49	3.01
303	11ec5	4.3267	6.4247	13.5972	1.6741	151.00	143.69	7.31
304	1ibc5	4.3267	6.0651	13.4033	1.2921	148.00	143.57	4.43
305	1m3ipc5	4.3267	6.4119	13.3128	1.6741	141.00	142.84	−1.84
306	1tbc5	4.3267	6.7312	12.5278	1.4883	145.00	135.44	9.56
307	3pec4	4.3267	5.9404	14.7850	0.9884	148.70	147.13	1.57
308	1bc5	4.3267	5.7164	13.9172	1.2267	156.60	148.24	8.36
309	1m2pc5	4.3267	6.0757	13.8928	1.5599	149.50	147.24	2.26
310	1sbc5	4.3267	6.1287	13.6878	1.3867	154.30	144.66	9.64
311	11m2ec5	4.3267	6.6684	13.2847	1.8832	138.00	141.85	−3.85
312	11m3ec5	4.3267	6.6163	12.9378	1.8832	133.00	141.17	−8.17
313	113mc6	4.3267	6.7302	12.8750	2.3001	136.60	143.12	−6.52
314	124mc6	4.3267	6.4935	13.8022	2.3001	144.80	148.62	−3.82
315	135mc6	4.3267	6.4600	13.2500	2.3001	139.50	147.07	−7.57
316	1e2mc6	4.3267	6.1929	14.1772	2.0514	154.30	150.89	3.41
317	1e3mc6	4.3267	6.1449	14.0350	2.0514	150.00	150.88	−0.88
318	1pc6	4.3267	5.8046	14.3372	1.7084	156.70	152.55	4.15
319	1ipc6	4.3267	6.1888	14.0350	1.8296	154.80	148.72	6.08
320	112mc6	4.3267	6.8104	13.5694	2.3001	145.10	144.71	0.39
321	114mc6	4.3267	6.7084	13.2850	2.3001	136.00	144.75	−8.75
322	1m1ec6	4.3267	6.4472	13.8022	2.0514	152.00	147.13	4.87
323	1m4ec6	4.3267	6.1284	14.1044	2.0514	150.80	151.27	−0.47
324	1ec7	4.3267	5.9394	15.1044	2.2340	163.70	157.85	5.85
325	12mc7	4.3267	6.2983	14.6667	2.4975	157.00	154.96	2.04
326	2ebc221h	4.3267	7.4064	14.1356	3.2890	152.00	148.53	3.47
327	4ms25o	4.3267	7.5723	16.0694	2.8388	149.00	149.95	−0.95
328	bc430n	4.3267	7.2729	15.0000	4.0756	163.00	158.79	4.21
329	dc4m	4.3267	6.7582	16.3478	2.6271	161.80	157.13	4.67
330	15ms33h	4.3267	7.7325	15.7222	3.2053	132.20	150.07	−17.87
331	77mbc221h	4.3267	8.1707	14.2222	3.7204	148.20	144.69	3.51
332	c10	4.6416	5.7744	16.3722	2.8872	202.00	187.84	14.16
333	13ec6	4.6416	6.1727	15.8339	1.8250	172.00	174.89	−2.89
334	1m2ipc6	4.6416	6.5461	15.3267	1.8250	171.00	169.81	1.19
335	1m3ipc6	4.6416	6.4969	15.3894	1.8250	167.00	170.48	−3.48
336	1m4ipc6	4.6416	6.4772	15.3233	1.8250	170.00	170.44	−0.44
337	1bc6	4.6416	5.8361	16.0241	1.3908	180.90	175.41	5.49
338	1sbc6	4.6416	6.2176	15.8339	1.5479	179.30	172.44	6.86
339	1ibc6	4.6416	6.1551	15.4139	1.4553	171.30	170.99	0.31
340	1tbc6	4.6416	6.7632	14.8094	1.6464	171.50	164.77	6.73
341	1pc7	4.6416	5.9537	16.6256	1.8672	182.80	179.71	3.09
342	26mbc321o	4.6416	7.7939	16.4689	3.6762	164.50	175.58	−11.08
343	37mbc330o	4.6416	7.6516	12.4308	3.6153	166.00	163.42	2.58

[a] The molecular structures of compounds are showed in Figure 3. Compounds of no. 151−160 come from ref 9; the others come from ref 10.

set to test the QSPR applicability of the three indices, OEI, VDI, and RDI, in combination with another variable $N^{2/3}$ (N is the number of carbon atoms. $N^{2/3}$ is a useful variable correlating with the boiling points.[8] So $N^{2/3}$ is used here.). The following correlations have been carried out

$$Bp = a_0 + a_1 N^{2/3} + a_2 VDI + a_3 OEI + a_4 RDI \quad (6a)$$

$$\ln(Bp_0 - Bp) = b_0 + b_1 N^{2/3} + b_2 VDI + b_3 OEI + b_4 RDI \quad (7a)$$

where a_0 and b_0 are the intercept terms; a_1, a_2, a_3, a_4, b_1, b_2, b_3, and b_4 are the regression coefficients; and $Bp_0 = 577$, which is a parameter obtained by an optimization technique with the purpose of getting the lowest standard deviation for the examined compounds.

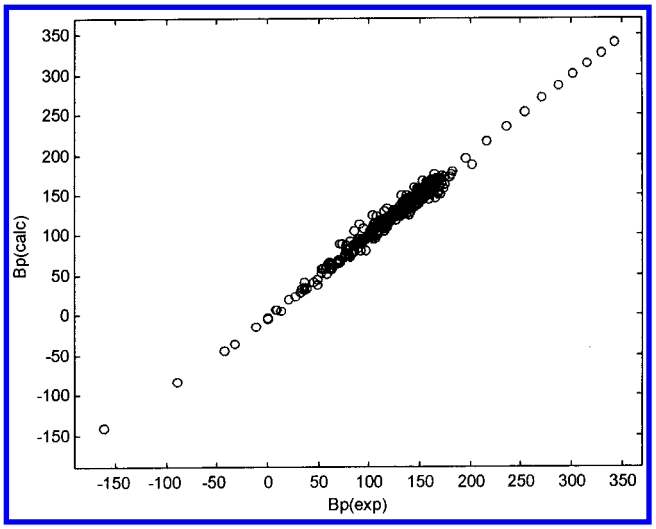
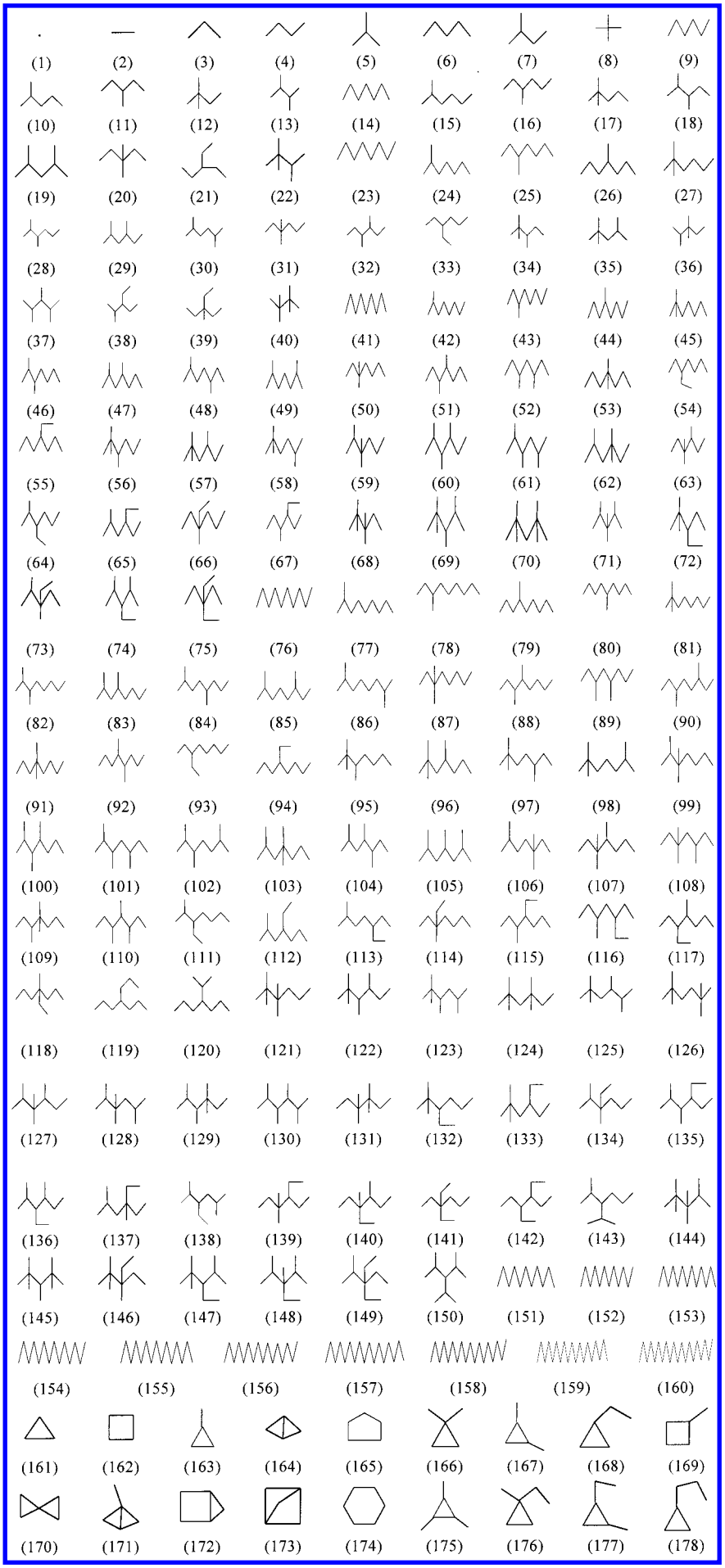


Figure 2. Plot of Bp_{calc} vs Bp_{exp} for 343 compounds.


(1) (2) (3) (4) (5) (6) (7) (8) (9)
(10) (11) (12) (13) (14) (15) (16) (17) (18)
(19) (20) (21) (22) (23) (24) (25) (26) (27)
(28) (29) (30) (31) (32) (33) (34) (35) (36)
(37) (38) (39) (40) (41) (42) (43) (44) (45)
(46) (47) (48) (49) (50) (51) (52) (53) (54)
(55) (56) (57) (58) (59) (60) (61) (62) (63)
(64) (65) (66) (67) (68) (69) (70) (71) (72)
(73) (74) (75) (76) (77) (78) (79) (80) (81)
(82) (83) (84) (85) (86) (87) (88) (89) (90)
(91) (92) (93) (94) (95) (96) (97) (98) (99)
(100) (101) (102) (103) (104) (105) (106) (107) (108)
(109) (110) (111) (112) (113) (114) (115) (116) (117)
(118) (119) (120) (121) (122) (123) (124) (125) (126)
(127) (128) (129) (130) (131) (132) (133) (134) (135)
(136) (137) (138) (139) (140) (141) (142) (143) (144)
(145) (146) (147) (148) (149) (150) (151) (152) (153)
(154) (155) (156) (157) (158) (159) (160)
(161) (162) (163) (164) (165) (166) (167) (168) (169)
(170) (171) (172) (173) (174) (175) (176) (177) (178)

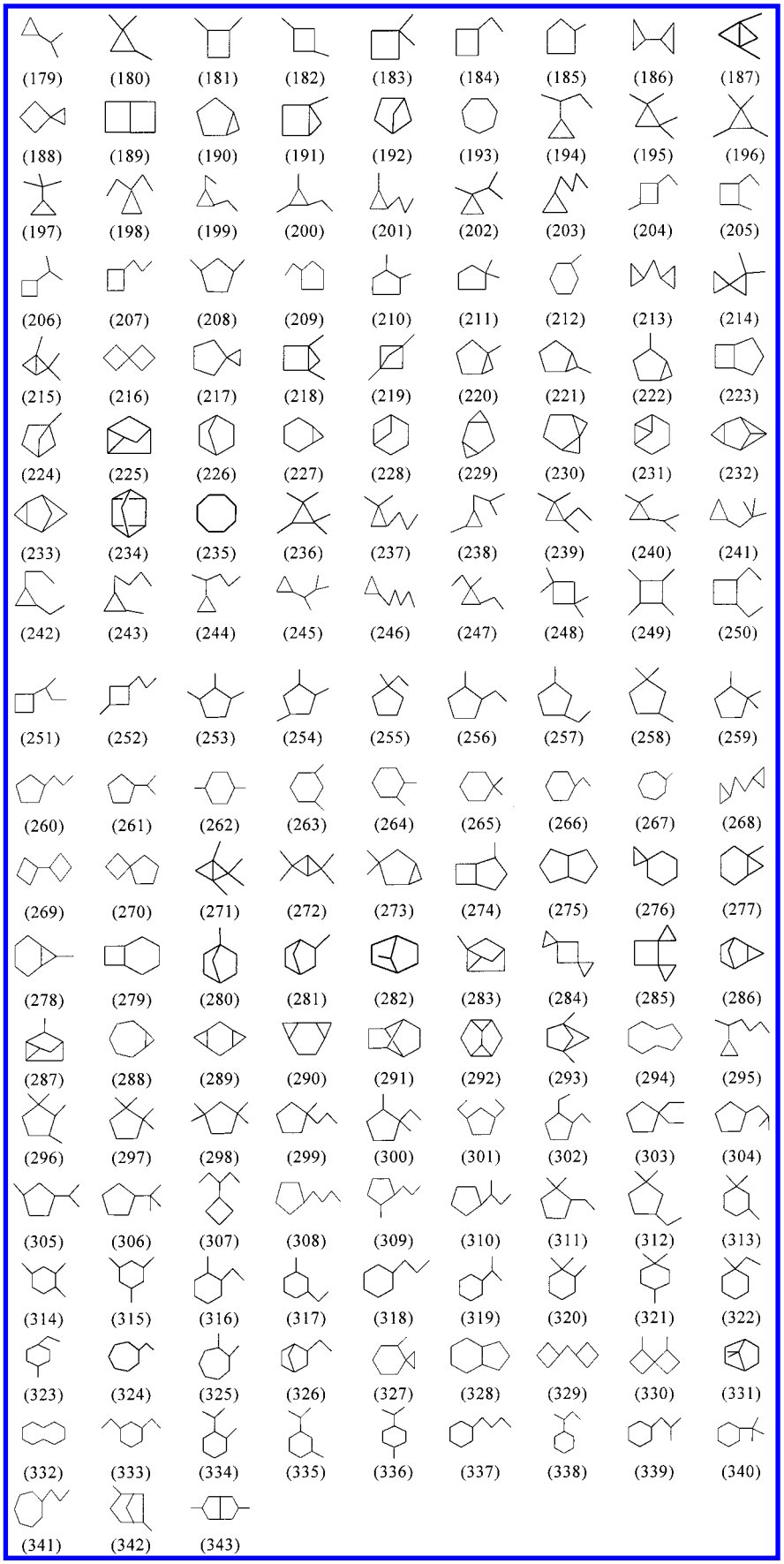


Figure 3. The molecular structures of 160 paraffins and 183 cycloalkanes in Table 3.

According to the regression models 6a and 7a, we obtained the corresponding eqs 6b and 7b, respectively:

$$\mathrm{Bp} = -182.821198 + 77.491737N^{2/3} - 4.904214\mathrm{VDI} + 0.731171\mathrm{OEI} + 6.742943\mathrm{RDI} \quad (6b)$$

$$F = 2650.74,\ r = 0.9844,\ \mathrm{rms} = 9.12\ ^\circ\mathrm{C},\ n = 343$$

$$\ln(577 - \mathrm{Bp}) = 6.729609 - 0.154107N^{2/3} + 2.285632 \times 10^{-2}\mathrm{VDI} - 7.921410 \times 10^{-3}\mathrm{OEI} - 1.821962 \times 10^{-2}\mathrm{RDI} \quad (7b)$$

$$F = 6455.09,\ r = 0.9935,\ \mathrm{rms} = 6.44\ ^\circ\mathrm{C},\ n = 343$$

where rms is the root-mean-square error, $\mathrm{rms} = \sqrt{\sum_{i=1}^{M}(\mathrm{Bp_{exp}} - \mathrm{Bp_{calc}})^2/(M-1)}$, in which M is the number of examined compounds.

In comparison, eq 7b gets much better correlation result than eq 6b, which shows that the linear model does not reflect the exact relationship between Bp and the four variables. Many examinations have revealed that the plot of Bp vs $N^{2/3}$ observes a logarithm function. Therefore, we prefer the logarithm transformation mode of Bp as the dependent variable instead of Bp.[6] The high correlation coefficient and low rms error show that eq 7b is a good model for estimating or predicting the boiling points of alkanes. Furthermore it employed only four variables for such a large and combined sample and the computation of these variables is simple. So we may say that these four variables have expressed most of the molecular structure information and can be used for QSPR study. Before reaching the above conclusion, this paper contains the four variables which have been studied in the following several aspects:

(1) In general, the variables employed in one regression equation should be independent of each other, that is to say, the molecular structure information can be expressed by fewer variables instead of more variables with high linear correlation among them. So the interrelation of variables is an important mark to evaluate the rationality of the defined indices. In this paper, we studied the interrelations of these four variables (listed in Table 1) and found that $N^{2/3}$ is closely related to OEI. But for the isomers with the main carbon chain of the same length, the values of $N^{2/3}$ are equal, while the values of OEI vary in a wavy line with the variation of the substituted situation (for instance, *n*-methylnonane, when $n = 2, 3, 4, 5$, their N are all 10 and their OEI are 14.5433, 14.7968, 14.6309, 14.7665, respectively), which indicates that OEI can reflect more details about the molecular structure features than $N^{2/3}$. So OEI is an important molecular descriptor and cannot be replaced by $N^{2/3}$ in the Bp-structure study.

(2) To evaluate the importance of each variable in the QSPR study and the contribution of the number of variables to the result, we regressed one to four variables ($N^{2/3}$, VDI, OEI, RDI) against the boiling points of 343 hydrocarbons as to model 7a, respectively. The obtained results are listed in Table 2.

As shown in Table 2, the single-variable regression of $N^{2/3}$ has a high correlation coefficient ($r = 0.9761$), which proves that Bp is closely related to $N^{2/3}$. The poor correlation results obtained by regressions with VDI, RDI, and OEI without $N^{2/3}$ also indicate that $N^{2/3}$ is very important in Bp-structure studies. But only one variable, $N^{2/3}$, is not enough to deliver most of the molecular structure information. It should be in combination with other variables VDI, RDI, and OEI to express more details about the structure features. The four-variable regression yields the best result (see eq 7b), by which the calculated Bps($\mathrm{Bp_{calc}}$) and the differences ($\Delta = \mathrm{Bp_{exp}} - \mathrm{Bp_{calc}}$) between $\mathrm{Bp_{exp}}$ and $\mathrm{Bp_{calc}}$ are listed in Table 3.

Table 4. Comparison of the Results from the Recent Literatures with Ours on the Bp-Structure Studies for Paraffins and Cycloalkanes

compounds	NC[a]	variables	NV[b]	*r*	rms (°C)
paraffins	150	x'[c]	1	0.9945	5.15
		$\lambda_{11}, \cdots, \lambda_{44}$[d]	10	0.9954	4.82
		$N^{2/3}$, VDI, OEI[e]	3	0.9952	4.21
cycloalkanes	80	$\mu_0, \cdots \mu_7$[f]	6	0.9937	4.80
		$N^{2/3}$, VDI, OEI, RDI[e]	4	0.9955	4.78

[a] Number of compounds. [b] Number of variables. [c] From ref 11. [d] From ref 6. [e] From this work. [f] From ref 7.

The linear relationship between the experimental and calculated boiling points for 343 compounds is given in Figure 2.

(3) In comparison, in this paper are collected several previous literatures on the study of Bp-structure relationships for alkanes and made regression analysis against various samples with the newly defined indices. The results of each case were listed in Table 4.

4. CONCLUSION

People have been going into QSPR/QSAR studies with more or less different variables, but so far, there is not a wide-admitted criterion about the number of variables employed in a QSPR model. For different samples and properties, it is difficult to define a specific number of variables. However a common rule should be complied with that the variables employed not only can render most of the molecular structure information but also not seem to be burdensome.

In the Bp-structure regression for a large combined sample of acyclic, monocyclic, and polycyclic alkanes with $N^{2/3}$, OEI, VDI, and RDI, the correlation coefficient is good and the root-mean-square error is within the reasonable limits, which proves that the newly defined indices based on the essential elements of molecular graph are useful in QSPR study and worthy of further investigation.

ACKNOWLEDGMENT

This project was supported by Hunan Province Education Commission (HPEC).

REFERENCES AND NOTES

(1) Randic, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.
(2) Balaban, A. T. Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett.* **1982**, *89*, 399−404.
(3) Schultz, H. P. Topological Organic Chemistry. 1. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci*. **1989**, *29*, 227−228.
(4) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc*. **1947**, *69*, 17−20.
(5) Mihalic, Z.; Trinajstic, N. A Graph-Theoretical Approach to Structure−Property Relationships. *J. Chem. Educ.* **1992**, *69*, 701−712.

(6) Liu, S.; Cao, C.; Li, Z. Approach to Estimation and Prediction for Normal Boiling Point (NBP) of Alkanes Based on a Novel Molecular Distance-Edge (MDE) Vector, *λ*. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 387−394.
(7) Estrada, E. Spectral Moments of the Edge Adjacency Matrix in Molecular Graphs. 3. Molecules Containing Cycles. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 23−27.
(8) Cao, C.; Liu, S.; Li, Z. On Molecular Polarizability: 2. Relationship to the Boiling Point of Alkanes and Alcohols. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1105−1111.
(9) *CRC Handbook of Chemistry and Physics*; 58th ed.; Weast, R. C., Ed.; CRC Press: 1977−1978; c280−c532.
(10) Rücker, G.; Rücker, C. On Topological Indices, Boiling Points, and Cycloalkanes. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 788−802.
(11) Mihalic, Z. Comparative Study of Molecular Descriptors Derived from the Distance Matrix. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 28−37.

CI000467T