match no. 5 pp. 197 - 218 1979

CHEMICAL GRAPHS. XXXVI[1]. CORRELATIONS BETWEEN OCTANE NUMBERS AND TOPOLOGICAL INDICES OF ALKANES

Alexandru T. Balaban[a] and Ioan Moţoc[b]

[a]Department of Organic Chemistry, The Polytechnic Bucharest, Roumania

[b]Chemistry Research Centre, Timişoara, Roumania

Herrn Prof.Dr. O.E. POLANSKY zu seinem 60.Geburtstag mit allen guten Wünschen gewidmet.

(received: May 1979)

Abstract. Correlations between octane numbers of alkanes with 4-8 carbon atoms and topological indices were investigated. Statistical evaluation indicates that the newly proposed centric indices give best correlations, both in uniparametric and biparametric equations (the second parameter being either the molecular weight of the alkane or its number of carbon atoms). Intercorrelations between 14 topological indices were investigated on the basis of data for alkane isomers ; with the exception of I_{pc} all indices were found to be strongly intercorrelated.

1. Introduction

In a previous part of this series[2], topological indices were reviewed and five new topological indices were introduced. We shall briefly summarize below the notations and definitions of all topological indices which will be used in the present paper. For details, ref. 2 or the original references should be consulted.

The centric index[2] B is defined for acyclic hydrocarbons by using the hydrogen-depleted graph of the carbon skeleton. On finding the centre or bicentre of such a tree (graph with no cycles) the vertices of degree one are stepwise removed. The numbers of these steps removed at each step give rise to a pruning (lopping) sequence. This sequence v_1, v_2, ... which indicates the numbers of vertices with degree one removed at each step, is used for the definition and calculation of the

centric index B :

$$B = \sum_i v_i^2$$

If the degrees d_i of vertices in the hydrogen - depleted graph are taken into account, the sum for all vertices

$$M_1 = \sum_{i=1}^{n} d_i^2$$

defines[3] another topological index introduced by Gutman, Rusčić, Trinajstić and Wilcox Jr.

A related index N_2 was introduced by Platt[4] and by Gordon and Scantlebury[5]. Platt defined the "first-neighbour sum f " by summing for all edges the numbers of edges adjacent to an edge. Gordon and Scantlebury defined N_2 as the number of distinct ways a C - C - C fragment may be superimposed on the carbon graph. It can be shown that $f = 2N_2$. If we denote by V_{d_i} the number of vertices of degree d_i (d_i = 1, 2, 3 or 4), then it can be shown[2] that for a graph with n vertices

$$M_1 = 16V_4 + 9V_3 + 4V_2 + V_1 = 6V_4 + 2V_3 + 4n + 6 = \sum_{i=1}^{n} d_i^2$$

$$N_2 = 6V_4 + 3V_3 + V_2 = 3V_4 + V_3 + n - 2 = \frac{1}{2} \sum_{i=1}^{n} d_i (d_i - 1)$$

The above three indices increase with increasing branching, and also with the number of carbon atoms. In order to conserve only the variation with branching, these indices may normalized so as to be zero for any n - alkane (chain - graph). Thus the normalized centric index was defined[2] as :

$$C = \frac{1}{2} (B - 2n + U) = \frac{1}{2} \left[B - B \text{ (chain-graph)} \right]$$

where $U = \left[1 - (-1)^n \right] / 2$.

The normalized quadratic index related to M_1 and N_2 is defined[2] as :

$$Q = 4V_4 + \frac{3}{2}V_3 + 1 - \frac{1}{2} V_1 = 3V_4 + V_3 = \frac{1}{2} \sum_{i=1}^{n} iV_i - 2n + 3$$

The relationship between Q, N_2 and M_1 is proved by the

expressions [2] :

$$Q = N_2 - n + 2 = \frac{M_1}{2} - 2n + 3$$

$$M_1 = 2(N_2 + n - 1)$$

In order to approximate through topological indices the "topological shape" of the molecule, bunormalized topological indices denoted by C' and Q' were introduced [2]. By definition, these indices have value zero for the normal (linear) alkanes (chain-graphs), and value one for the most branched graphs (star-graphs) formed by $n - 1$ vertices of degree one, and one vertex of degree $n - 1$. The binormalization is carried out by dividing C and Q through the respective values for star-graphs :

$$C' = \frac{C}{C(\text{star-graph})} = \frac{B - 2n + U}{(n-2)^2 - 2 + U}$$

$$Q' = \frac{Q}{Q(\text{star-graph})} = \frac{3V_4 + V_3}{2(n-2)(n-3)}$$

The Wiener index [6] (or the path number) w is the sum of the number of bonds between all pairs of vertices :

$$w = \sum_i i g_i$$

where g_i denotes the number of pairs of vertices whose distance is i, i. e. half the number of times a distance i occurs in the distance matrix $\underline{D}$; Rouvray's index [7] I is the sum of entries in the distance matrix $\underline{D} = [d_{ij}]$:

$$I = \sum_{i,j} d_{ij} = 2w$$

The Hosoya index [8] denoted by $Z = \sum_k p(G,k)$ for a graph G is the number of ways in which k edges may be chosen so that no two of them are adjacent ; for trees, Z is the sum of the absolute values of coefficients p(G,k) in the characteristic polynomial

$$\sum_k (-1)^k \, p(G,k) \, x^{n-2k} = (-1)^n \det \left| \underline{A} - x\underline{E} \right|$$

where $\underline{A}$ is the adjacency matrix and $\underline{E}$ is the unit matrix.

The largest eigenvalue of the characteristic polynomial was also considered [9] to be a measure of branching.

Bonchev and Trinajstić [10] introduced several informational topological indices. The information content I_X of a sequence of numbers X_i, where $\sum_i X_i = X$, is

$$I_X = X \log_2 X - \sum_i X_i \log_2 X_i$$

and the mean information content $\bar{I}_X$ is

$$\bar{I}_X = \frac{I_X}{X} = - \sum_i p_i \log_2 p_i$$

where $p_i = X_i/X$.

Thus, the information content and the mean information content on polynomial coefficients, I_{pc} and $\bar{I}_{pc}$, respectively, are obtained from the coefficients of the characteristic polynomial of graph G :

$$I_{pc} = Z \log_2 Z - \sum_k p(G,k) \log_2 p(G,k)$$

$$\bar{I}_{pc} = - \sum_k \frac{p(G,k)}{Z} \log_2 \frac{p(G,k)}{Z} = I_{pc} / t$$

where t is the number of terms in the characteristic polynomial.

The information content and the mean information content on realized distances (I_D^W and $\bar{I}_D^W$, respectively) as well as on the distribution of distances (I_D^E and $\bar{I}_D^E$, respectively) are connected to the distance matrix $\underline{D}$ of the graph G, and to the Wiener index w :

$$I_D^W = w \log_2 w - \sum_i g_i \, i \log_2 i$$

$$\bar{I}_D^W = - \sum_i g_i \frac{i}{w} \log_2 \frac{i}{w} = I_D^W / w$$

$$I_D^E = \frac{n(n-1)}{2} \log_2 \frac{n(n-1)}{2} - \sum_i g_i \log_2 g_i$$

$$\bar{I}_D^E = - \sum_i \frac{2g_i}{n(n-1)} \log_2 \frac{2g_i}{n(n-1)} = \frac{2}{n(n-1)} I_D^E$$

Finally, Randić's index denoted by χ or M_2 is [11] :

$$\chi = M_2 = \sum_q (v_i v_j)^{-1/2}$$

where v_i and v_j denote the degrees of two adjacent vertices i and j in the hydrogen-depleted graph G, and the summation

is extended over all q edges of the graph G.

An intercomparison between all above topological indices shews [2,10,12] that the highest discriminating ability among isomers is associated with indices Z, χ and I_D^E .

2. Correlations between octane numbers and topological indices

The correlational ability of some topological indices mentioned in the previous section and denoted in general by T_j , was investigated in monoparametric (linear) equations of type (1) :

$$\hat{Y} = a + b\, T_j \qquad (1)$$

where $\hat{Y}_i$ estimates the measured values Y_i.

A property of alkanes which is known to depend strongly on the degree of branching is their "anti-knock" ability : the more branched the alkane, the less prone it is to self--ignition on sudden compression in air. Conventionally, the anti-knock ability is determined in terms of octane numbers (ON), with n-heptane having ON = 0.0 and 2,2,4-trimethylpentane (isoöctane) having ON = 100.0 as standards. No correlation with topological indices has been described so far. The ON of an alkane or of a gasoline sample is determined under prescribed conditions with a test internal combustion engine and is equal to the ratio (percent) of isoöctane in that dual mixture with n-heptane which affords the same engine power as the given sample. Two different types of ON exist [13], namely Research ON (denoted by RON or Y_1) and Motor ON (denoted by MON or Y_2). The correlations were investigated primarily on heptane and octane isomers, mainly for two reasons : (a) practically all their isomers have been tested experimentally for ON determination (unlike the higher alkanes C_9 , C_{10} , etc.), and (b) in these two cases of C_7 and C_8 , a fair number of isomers with various degrees of branching exist allowing meaningful correlations (unlike the lower alkanes C_4 , C_5 , C_6 which have small numbers of isomers.

Experimental data for ON's [13] and structures for all alkane isomers are presented in Table 1, and calculated topo-

TABLE 1. Values of octane numbers for alkanes[13] used in correlations.

No.	C_nH_{2n+2}	RON (Y_1)	MON (Y_2)	No.	C_nH_{2n+2}	RON (Y_1)	MON (Y_2)	No.	C_nH_{2n+2}	RON (Y_1)	MON (Y_2)
1.	n-C_4	93.6	90.1	13.	3-Me-C_6	52.0	55.0	25.	2,2-Me_2C_6	72.5	77.4
2*	i-C_4	-	97.6	14.	3-Et-C_5	65.0	65.0	26.	2,3-Me_2C_6	71.3	78.9
3.	n-C_5	61.7	61.9	15.	2,2-Me_2-C_5	92.8	95.6	27.	2,4-Me_2C_6	65.2	69.9
4.	2-Me-C_4	92.3	90.3	16.	2,3-Me_2-C_5	91.1	88.5	28.	2,5-Me_2C_6	55.5	55.7
5.	2,2-Me_2-C_3	85.5	80.2	17.	2,4-Me_2-C_5	83.1	83.8	29.	3,3-Me_2-C_6	75.5	83.4
6.	n-C_6	24.8	26.0	18.	3,3-Me_2-C_5	80.8	86.6	30.	3,4-Me_2-C_6	76.3	81.7
7.	2-Me-C_5	73.4	73.5	19*	2,2,3-Me_3-C_4	-	-	31.	2-Me-3-Et-C_5	87.3	88.1
8.	3-Me-C_5	74.5	74.3	20*	n-C_8	-	-	32.	3-Me-3-Et-C_5	80.8	88.7
9.	2,2-Me_2-C_4	91.8	93.4	21.	2-Me-C_7	21.7	23.8	33.	2,2,4-Me_3-C_5	100.0	100.0
10*	2,3-Me_2C_4	-	94.4	22.	3-Me-C_7	26.8	35.0	34*	2,2,3-Me_3-C_5	-	99.9
11.	n-C_7	0.0	0.0	23.	4-Me-C_7	26.7	39.0	35*	2,3,3-Me_3-C_5	-	99.4
12.	2-Me-C_6	42.4	46.4	24.	3-Et-C_6	33.5	52.4	36*	2,2,3,3-Me_4-C_4	-	-
								37*	2,3,4-Me3-C5	-	95.9

* These compounds were not used in correlations.

TABLE 2. Values of topological indices T_j, of n and MW for alkanes with the same numbering as in Table 1.

No.	B	C	M_1	Q	C'	Q'	N_2	I_D^E	w	Z	I_D^W
1.	8	0	10	0	0.000	0.000	2	8.7549	10	5	24.4644
2.	10	1	12	1	1.000	1.000	3	6.0000	9	4	22.5293
3.	9	0	14	0	0.000	0.000	3	18.4644	20	8	62.9288
4.	13	2	16	1	0.500	0.333	4	15.2193	18	7	57.5489
5.	17	4	20	3	1.000	1.000	6	9.7095	16	5	52.0000
6.	12	0	18	0	0.000	0.000	4	32.2388	35	13	129.6506
7.	14	1	20	1	0.143	0.167	5	28.6292	32	11	119.7353
8.	14	1	20	1	0.143	0.167	5	27.3841	31	12	116.5605
9.	24	6	24	3	0.857	0.500	7	22.5873	28	9	106.3413
10.	20	4	22	2	0.571	0.333	6	23.4839	29	10	109.8619
11.	13	0	22	0	0.000	0.000	5	50.3644	56	21	233.4633
12.	17	2	24	1	0.167	0.100	6	46.4642	52	18	218.1841
13.	17	2	24	1	0.167	0.100	6	44.8546	50	19	210.5087
14.	19	3	24	1	0.250	0.100	6	40.9544	48	20	203.5488
15.	21	4	28	3	0.333	0.300	8	39.9740	46	14	195.0642
16.	21	4	26	2	0.333	0.200	7	39.5676	46	17	195.5544
17.	21	4	26	2	0.333	0.200	7	41.0774	47	15	203.0586
18.	21	4	28	3	0.333	0.300	8	37.2191	44	16	187.6856

TABLE 2 (continued)

No.	I_{pc}	$\bar{I}_D^E$	$\bar{I}_D^W$	$\bar{I}_{pc}$	%	n	MW
1.	6.8547	1.4591	2.4464	2.2849	1.9142	4	58
2.	3.2451	1.0000	2.5033	1.6226	1.7321	4	58
3.	11.2451	1.8464	3.1464	3.7984	2.4142	5	72
4.	9.6515	1.5219	3.1972	3.2172	2.2701	5	72
5.	7.6096	0.9710	3.2500	1.8048	2.0000	5	72
6.	20.9863	2.1493	3.7043	5.2466	2.9142	6	86
7.	14.8345	1.9086	3.7417	4.9448	2.7701	6	86
8.	13.8003	1.8256	3.7600	4.9501	2.8081	6	86
9.	12.1648	1.5058	3.7979	4.0549	2.5607	6	86
10.	13.6097	1.5656	3.7883	4.5366	2.6427	6	86
11.	35.5096	2.3983	4.1690	8.8774	3.4142	7	100
12.	29.0195	2.2126	4.1959	7.2549	3.2701	7	100
13.	31.9166	2.1359	4.2102	7.9792	3.3081	7	100
14.	34.3995	1.9502	4.2406	6.0472	3.3461	7	100
15.	18.1417	1.9035	4.2405	6.9943	3.0607	7	100
16.	27.9771	1.8842	4.2512	6.3645	3.1807	7	100
17.	19.0936	1.9561	4.2304	8.5999	3.1259	7	100
18.	26.8187	1.7723	4.2656	6.7098	3.1213	7	100

TABLE 2 (continued)

No.	B	C	M_1	Q	C'	Q'	N_2	I_D^E	w	Z	I_D^W
19.	29	8	30	4	0.667	0.400	9	32.6898	42	13	179.9479
20.	16	0	26	0	0.000	0.000	6	73.0802	84	34	400.5244
21.	18	1	28	1	0.059	0.067	7	68.9384	79	29	362.3758
22.	18	1	28	1	0.059	0.067	7	67.0382	76	31	349.9745
23.	18	1	28	1	0.059	0.067	7	66.1835	75	30	345.9030
24.	22	3	28	1	0.176	0.067	7	62.0418	72	32	333.7311
25.	24	4	32	3	0.235	0.200	9	62.0606	71	23	328.0287
26.	24	4	30	2	0.235	0.133	8	61.3099	70	27	324.5463
27.	24	4	30	2	0.235	0.133	8	61.8350	71	26	328.8834
28.	24	4	30	2	0.235	0.133	8	63.3448	74	25	341.2866
29.	24	4	32	3	0.235	0.200	9	58.7736	67	25	311.5341
30.	24	4	30	2	0.235	0.133	8	58.9544	68	29	316.2987
31.	26	5	30	2	0.294	0.133	8	55.3448	67	28	312.3888
32.	26	5	32	3	0.294	o.200	9	53.1409	64	28	299.2o59
33.	30	7	34	4	0.412	0.267	10	54.6154	66	19	3o7.1555
34.	30	7	34	4	0.412	0.267	10	52.9802	63	22	294.5294
35.	30	7	34	4	0.412	0.267	10	51.2058	62	23	290.3661
36.	30	7	32	3	0.412	0.200	9	54.4251	64	24	303.4147
37.	40	12	38	6	0.706	0.400	12	43.4056	58	17	272.9688

TABLE 2 (continued)

No.	I_{pc}	$\bar{I}_D^E$	$\bar{I}_D^W$	$\bar{I}_{pc}$	χ	n	MW
19.	17.0862	1.5567	4.2895	5.6954	2.9434	7	100
20.	**61.4996**	**2.6100**	**4.7681**	**12.2999**	3.9142	8	114
21.	48.2755	2.4621	4.5870	12.0689	3.7701	8	114
22.	56.6256	2.3992	4.6049	11.3251	3.8o81	8	114
23.	50.2522	2.3637	4.6120	12.5631	3.8081	8	114
24.	58.5162	2.2158	4.6352	11.7032	3.8461	8	114
25.	36.6160	2.2165	4.6221	9.1540	3.5607	8	114
26.	45.1150	2.1894	4.6364	11.2788	3.6807	8	114
27.	42.8488	2.2084	4.6322	10.7112	3.6639	8	114
28.	40.3392	2.2623	4.6120	10.0848	3.6259	8	114
29.	41.8158	2.0991	4.6498	10.4540	3.6213	8	114
30.	53.4727	2.1055	4.6515	10.6945	3.7187	8	114
31.	47.1972	1.9766	4.6625	11.7993	3.7187	8	114
32.	52.2834	**1.8979**	**4.6751**	10.4567	3.6820	8	114
33.	23.0054	1.9506	4.6539	7.6685	3.4165	8	114
34.	35.6474	1.8922	4.6751	8.9119	3.4814	8	114
35.	38.3367	1.8288	4.6833	9.5842	3.5o40	8	114
36.	39.3681	1.9438	4.6679	9.8420	3.5534	8	114
37.	21.3061	1.5502	4.2895	7.1020	3.2500	8	114

logical indices are displayed in Table 2. The numbers (No.) refer to the same isomer in Tables 1 and 2.

It appears, as known, that the anti-knock ability of alkanes increases with (i) increasing branching and (ii) decreasing molecular weight or number n of carbon atoms. The former feature deserves more detailed comments : the linear (normal) alkanes have the lowest ON's among all their isomers ; with increasing numbers of tertiary and/or quaternary carbon atoms, i. e. of branches, the ON increases ; the ON's of isomeric systems increase if on a longer chain a shorter branch moves farther from the endpoints of the chain.

Equations presented in the subsequent tables express more quantitatively the dependence of ON versus branching as reflected in topological indices ; in addition, these equations possess predictive capacity. Regression coefficients were determined by the method of least squares using a program devised for this purpose on a FELIX C-256 computer. The statistical parameters for estimating the accuracy of correlations are :

r (correlation coefficient)
s (standard deviation)
F (Fisher statistics)
EV (explained variance)

Tables 3 and 4 list for octane isomers the regression coefficients a and b in equations (1) as well as the above statistical parameters for Y_1 and Y_2 , respectively. Tables 5 and 6 list the same coefficients and parameters in correlations with Y_1 and Y_2, respectively, for heptane isomers.

It is seen from Tables 3 - 6 that from all topological indices listed in these tables, the centric indices B, C and C' give systematically the best regressional equations. Next best are the Wiener index w and the informational indices I_D^W , I_D^E and $\overline{I}_D^E$. Only in one case, as seen in Table 4, I_D^W gives better results for Y_2 in the C_8H_{18} series. Indices I_{pe} and $\overline{I}_{pe}$ give systematically the worst regressional equations.

In order to include in a single regression both the C_7H_{16} and the C_8H_{18} series, we used equations (2) :

$$\hat{Y}_i = a + b\,T_j + c\,\delta \quad , \text{ with } i = 1 \text{ or } 2 \qquad (2)$$

TABLE 3. Equations $\hat{Y}_1(RON)$ *versus* topological indices T_j for C_8H_{18} isomers.

T_j	a	b	r	s	F	EV
B	-100.031	6.932	0.945	8.443	41.652	0.883
C	10.884	13.864	0.945	8.443	41.652	0.883
M_1	-296.616	11.860	0.877	12.371	16.731	0.749
Q	11.744	23.720	0.877	12.371	16.731	0.749
Q'	11.794	355.431	0.876	12.451	16.454	0.746
C'	10.883	235.836	0.945	8.441	41.675	0.746
N_2	-130.575	23.72o	0.877	12.371	16.731	o.749
I_D^E	350.951	-4.756	0.897	11.400	20.593	0.787
I_D^W	473.651	-1.259	0.905	10.981	22.583	0.802
I_{pc}	135.392	-1.621	0.593	20.773	2.707	0.292
$\bar{I}_D^E$	353.204	-134.001	0.897	11.425	20.478	0.786
$\bar{I}_{pc}$	205.661	-13.436	0.678	18.965	4.247	0.410
w	435.755	-5.295	0.899	11.269	21.531	0.791
z	199.064	-5.099	0.696	18.508	4.785	0.438
χ	690.500	-170.769	0.766	16.571	7.112	0.550

TABLE 4. Equations $\hat{Y}_2$(MON) *versus* topological indices T_j for C_8H_{18} isomers.

T_j	a	b	r	s	F	EV
B	-63.429	5.571	0.940	8.443	49.057	0.875
C	25.717	11.143	0.940	8.443	49.057	0.875
M_1	-234.750	10.010	0.882	11.623	22.820	0.762
Q	25.528	20.020	0.882	11.623	22.820	0.762
C'	25.765	189.289	0.939	8.463	48.806	0.874
Q'	25.597	299.781	0.881	11.678	22.544	0.760
N_2	-94.596	20.020	0.882	11.623	22.820	0.762
I_D^E	326.238	-4.257	0.940	8.401	49.616	0.876
I_D^W	434,535	-1.123	0.959	7.022	73.819	0.913
I_{pc}	143.936	-1.597	0.587	19.988	3.414	0.298
$\bar{I}_D^E$	325.798	-118.887	0.941	8.375	49.966	0.877
$\bar{I}_{pc}$	214.093	-13.409	0.920	17.580	6.316	0.457
w	398.611	-4.697	0.702	7.167	71.414	0.910
Z	201.154	-4.868	0.957	17.520	6.254	0.460
χ	631.963	-152.964	0.778	15.499	9.987	0.578

TABLE 5. Equations Y_1(RON) *versus* topological indices T_j for C_7H_{16} isomers.

T_j	a	b	r	s	F	EV
B	-136.265	10.649	0.988	4.907	100.086	0.972
C	2.170	21.297	0.988	4.907	100.076	0.972
M_1	-264.277	12.977	0.876	15.167	8.237	0.728
Q	21.222	25.955	0.876	15.167	8.238	0.728
C'	2.083	256.020	0.988	4.891	100.754	0.972
Q'	21.222	259.555	0.876	15.167	8.237	0.728
N_2	-108.554	25.955	0.876	15.167	8.237	0.728
I_D^E	352.823	-6.800	0.934	11.251	17.013	0.851
I_D^W	480.156	-2.024	0.943	10.434	20.186	0.871
I_{pc}	160.741	-3.494	0.717	21.900	2.650	0.434
$\bar{I}_D^E$	352.841	-142.818	0.934	11.253	17.006	0.850
$\bar{I}_{pc}$	189.170	-17.104	0.563	25.966	1.163	0.204
W	438.835	-7.721	0.955	9.264	77.473	0.898
Z	245.737	-10.419	0.812	18.347	6.382	0.603
χ	760.907	-216.054	0.857	12.201	6.910	0.690

TABLE 6. Equations $Y_2(MON)$ *versus* topological indices T_j for C_7H_{16} isomers

T_j	a	b	r	s	F	EV
B	-135.242	1o.714	0.988	4.845	103.918	0.973
C	4.043	21.428	0.988	4.845	103.918	0.973
M_1	-268.525	13.235	0.888	14.526	9.339	0.754
Q	22.637	26.470	0.888	14.526	9.339	0.754
C'	3.952	257.614	0.988	4.814	105.276	0.973
Q'	22.637	264.697	0.888	14.526	9.339	0.754
N_2	-109.711	26.470	0.888	14.526	9.339	0.754
I_D^E	360.252	-6.922	0.945	10.333	20.898	0.875
I_D^W	490.359	-2.063	0.956	9.293	26.429	0.899
I_{pc}	162.489	-3.476	0.710	22.274	2.535	0.421
$\bar{I}_D^E$	360.277	-145.377	0.945	10.334	20.891	0.875
$\bar{I}_{pc}$	169.739	-17.827	0.584	25.661	1.294	0.231
w	447.306	-7.849	0.966	8.207	84.882	0.921
Z	248.692	-10.460	0.811	18.516	6.167	0.599
χ	766.922	-217.220	0.857	16.308	6.893	0.689

where δ is the indicator variable defined by relations (3) :

$$\delta = 1 \text{ for } C_8H_{18}$$
$$\delta = 0 \text{ for } C_7H_{16} \qquad (3)$$

The coefficients and statistical parameters of equations (2) are presented in Table 7 for the centric and quadratic indices B and M_1 , and for their normalized counterparts, C and Q, respectively. It may be observed from Table 7 that the normalization takes into account most, but not all, of the influence exerted by the molecular weight (or number of carbon atoms) of the alkane on the octane numbers ; this is evidenced by the lower value of the coefficient c of the indicator variable for indices C and Q than for indices B and M_1 in the case of Research ON's (Y_1) and the much lower values of c for indices C and Q in the case of Motor ON's (Y_2). By comparison between the two series Y_1 and Y_2 , the variations of coefficients a and b are insignificant.

In conclusion, the present Section demonstrates that good correlations can be found between octane numbers and selected topological indices for each series of isomeric alkanes, and that fair correlations are obtained by combining such series into a single one with the aid of indicator variables.

3. Biparametric correlations : octane numbers versus topological indices and number of carbon atoms or molecular weight of alkanes

Since in the preceding Section the introduction of the indicator variable δ was an ad-hoc solution, valid only for the C_7H_{16} plus C_8H_{18} series, we investigated a more general correlation which should take into account both factors (i) and (ii), cf. previous section : the degree of branching or shape of the alkane expressed by a topological index, and the size of the alkane expressed by the number of its carbon atoms, n, or by its molecular weight, MW. A biparametric equation (4) was employed :

TABLE 7. Equations $\hat{Y}_1$ or $\hat{Y}_2$ *versus* topological indices T_j and indicator variable δ for the combined series C_7H_{16} and C_8H_{18}

$\hat{Y}$	T_j	a	b	c	r	$r(T_j)$	s	F	EV
$\hat{Y}_1$	B	-86.532	7.996	-38.222	0.945	0.751	8.974	46.847	0.880
$\hat{Y}_1$	C	17.422	15.992	-14.333	0.945	0.910	8.974	46.848	0.880
$\hat{Y}_1$	M_1	-247.887	12.328	-62.848	0.876	0.510	13.171	18.740	0.741
$\hat{Y}_1$	Q	23.333	24.656	-13.535	0.876	0.843	13.171	18.710	0.742
$\hat{Y}_2$	B	-76.174	7.564	-32.311	0.930	0.781	9.666	36.090	0.849
$\hat{Y}_2$	C	25.694	13.898	-5.502	0.902	0.896	11.323	24.763	0.793
$\hat{Y}_2$	M_1	-230.212	11.717	-55.879	0.866	0.562	13.104	17.055	0.723
$\hat{Y}_2$	Q	27.569	23.434	-9.010	0.866	0.850	13.104	17.055	0.723

$$\hat{Y} = a + b\,A_k + c\,T_j \text{ , with } i,k = 1 \text{ or } 2. \qquad (4)$$

The size of the alkane was expressed either by A_1 - its number n of carbon atoms ($A_1 = n$) or by A_2 - its molecular weight ($A_2 = 14n + 2$). A series of 29 alkanes with $n = 4 \div 8$ was tested.

Tables 8 and 9 contain the coefficients a, b, c and the statistical parameters of the correlational equations versus five topological indices and A_1 or A_2, respectively. Since A_1 is linearly dependent on A_2 (i.e., $A_2 = 14A_1 + 2$), the corresponding correlations from Tables 9 and 10 are of the same quality. Through the values of correlation coefficients r are only fairly satisfactory, the equations are significant at a confidence level above 95%. Among all five indices investigated, the best correlations are again presented by the centric indices B and C.

4. Intercorrelation of topological indices

In order to investigate which of the 14 topological indices mentioned in the present paper are intercorrelated, we used the topological indices for the 18 octane isomers from Table 1, and fitted them pairwise according to equation (5):

$$T_i = \alpha + \beta\,T_j \text{ , with } i,j = 1 \div 14,\ i \neq j \qquad (5)$$

The result is displayed as a correlation matrix in Table 10, giving the correlation coefficients, r.

It can be observed that 13 topological indices are strongly intercorrelated (in some cases, e.g. for B-C-C', M_1-Q-Q'--N_2, r = 1.00 because the respective indices are linearly related). Therefore it would be meaningless to develop multiparametric equations involving more than one of these 13 topological indices. On the other hand, I_{pc} is almost orthogonal to all the above 13 topological indices, presenting correlation coefficients in the range 0.2 - 0.6 to all other 13 topological indices.

TABLE 8. Biparametric equations $\hat{Y}_1$ or $\hat{Y}_2$ *versus* topological indices T_j and number n of carbon atoms for all alkanes with n=4 ÷ 8.

$\hat{Y}$	T_j	a	b	c	r	s	$r(T_j)$	F	EV
$\hat{Y}_1$	B	132.940	-25.375	5.718	0.85	13.94	0.32	21.69	0.70
$\hat{Y}_1$	C	134.119	-15.007	12.557	0.86	13.41	0.62	24.12	0.72
$\hat{Y}_1$	M_1	186.421	-53.817	1.077	0.82	15.12	0.02	17.17	0.65
$\hat{Y}_1$	Q	125.956	-13.509	20.152	0.82	15.12	0.61	17.18	0.62
$\hat{Y}_1$	C'	75.368	-3.743	64.473	0.61	21.02	0.58	4.88	0.32
$\hat{Y}_1$	w	-119.893	-3.296	50.815	0.65	20.18	0.40	5.99	0.37
$\hat{Y}_1$	Z	-5.501	-4.198	21.758	0.57	21.66	0.27	4.11	0.28
$\hat{Y}_1$	%	100.566	-148.615	63.360	0.75	17.33	-0.42	11.10	0.54
$\hat{Y}_2$	B	126.089	-24.875	6.007	0.85	13.30	0.41	21.02	0.69
$\hat{Y}_2$	C	116.979	-11.847	11.921	0.84	13.56	0.67	19.91	0.68
$\hat{Y}_2$	M_1	166.483	-48.605	9.548	0.80	15.14	0.12	14.31	0.60
$\hat{Y}_2$	Q	1o9.201	-10.413	19.099	0.80	15.14	0.66	14.31	0.60
$\hat{Y}_2$	C'	61.453	-1.171	60.716	0.57	20.56	0.56	3.95	0.27
$\hat{Y}_2$	w	-127.970	51.591	-3.184	0.62	19.53	-0.29	5.29	0.34
$\hat{Y}_2$	Z	-5.211	20.516	-3.603	0.49	21.78	-0.32	2.61	0.18
$\hat{Y}_2$	%	81.930	-143.66o	64.291	0.74	17.06	-0.23	11.24	0.52

TABLE 9. Biparametric equations $\hat{Y}_1$ or $\hat{Y}_2$ versus topological indices T_j and molecular weight MW for all alkanes C_4H_{10} - C_8H_{18}.

$\hat{Y}$	T_j	a	b	c	r	s	F	EV
$\hat{Y}_1$	B	147.613	-2.045	6.302	0.87	13.24	24.90	0.73
$\hat{Y}_1$	C	136.211	-1.071	12.555	0.86	13.39	24.16	0.72
$\hat{Y}_1$	M_1	194.007	-3.841	10.070	0.82	15.12	17.16	0.65
$\hat{Y}_1$	Q'	127.829	-0.964	20.141	0.82	15.12	17.16	0.65
$\hat{Y}_1$	C'	75.861	-0.267	64.465	0.61	21.00	4.88	0.32
$\hat{Y}_1$	w	-127.128	3.629	-3.296	0.65	20.18	5.67	0.37
$\hat{Y}_1$	Z	-8.604	1.554	-4.198	0.57	21.66	2.39	0.28
$\hat{Y}_2$	B	129.638	-1.777	6.007	0.85	13.30	21.02	0.69
$\hat{Y}_2$	C	118.668	-0.846	11.921	0.84	13.56	19.91	0.68
$\hat{Y}_2$	M_1	173.433	-3.472	9.549	0.80	15.14	14.32	0.60
$\hat{Y}_2$	Q'	110.686	-0.744	19.099	0.80	15.14	14.31	0.60
$\hat{Y}_2$	C'	61.618	-0.084	60.716	0.57	20.56	3.95	0.30
$\hat{Y}_2$	w	-135.339	3.685	-3.184	0.62	19.53	5.29	0.34
$\hat{Y}_2$	Z	-8.144	1.465	-3.603	0.49	21.78	2.61	0.18

TABLE 10. Intercorrelation of topological indices.

	B	C	M_1	Q	C'	Q'	N_2	I_D^E	w	Z	I_D^W	I_{pc}	$\bar{I}_D^E$	$\bar{I}_D^W$
B	1.00	1.00	0.90	0.90	1.00	0.90	0.90	-0.93	-0.91	-0.78	-0.91	-0.41	-0.93	0.86
C		1.00	0.90	0.90	1.00	0.90	0.90	-0.93	-0.91	-0.78	-0.91	-0.41	-0.93	0.86
M_1			1.00	1.00	0.90	1.00	1.00	-0.85	-0.85	-0.89	-0. 6	-0.43	-0.85	0.76
Q				1.00	0.90	1.00	1.00	-0.85	-0.85	-0.89	-0.86	-0.43	-0.85	0.76
C'					1.00	0.90	0.90	-0.93	-0.91	-0.78	-0.91	-0.41	-0.93	0.86
Q'						1.00	1.00	-0.84	-0.85	-0.89	-0.86	-0.43	-0.85	0.75
N_2							1.00	-0.85	-0.85	-0.89	-0.86	-0.43	-0.85	o.76
I_D^E								1.00	0.98	0.56	0.98	0.21	1.00	-0.97
w									1.00	0.58	1.00	0.21	0.98	-0.98
Z										1.00	0.59	0.64	0.59	-0.45
I_D^W											1.00	0.22	0.98	-0.98
I_{pc}												1.00	0.21	-0.13
$\bar{I}_D^E$													1.00	-0.97
$\bar{I}_D^W$														1.00

REFERENCES

1. Preceding part, A. T. BALABAN and V. BACIU, Math. Chem. 4, 131 (1978).
2. A. T. BALABAN, Theoret. Chim. Acta (submitted).
3. I. GUTMAN, R. RUSCIĆ, N. TRINAJSTIĆ and C. F. WILCOX JR., J. Chem. Phys. 62, 3399 (1975).
4. J. R. PLATT, J. Chem. Phys. 15, 419 (1947) ; J. Phys. Chem. 56, 328 (1952).
5. M. GORDON and G. R. SCANTLEBURY, Trans. Faraday Soc. 60, 605 (1964).
6. H. WIENER, J. Amer. Chem. Soc. 69, 17, 2636 (1947) ; J. Chem. Phys. 15, 766 (1947) ; J. Phys. Chem. 52, 425, 1082 (1948).
7. D. H. ROUVRAY and B. C. CRAFFORD, South African J. Sci. 72, 47 (1976) ; D. H. ROUVRAY, Amer. Sci. 61, 729 (1973) ; Math. Chem. 1, 125 (1975).
8. H. HOSOYA, Bull. Chem. Soc. Japan, 44, 2332 (1971) ; Internat. J. Quantum Chem. 6, 801 (1972) ; J. Chem. Docum. 12, 181 (1972) ; Chem. Letters (Japan), 65 (1972) ; Fibonacci Quart. 3, 255 (1973).
9. L. LOVASZ and J. PELIKAN, Period. Math. Hung. 3, 175 (1973).
10. D. BONCHEV and N. TRINAJSTIĆ, J. Chem. Phys. 67, 4517 (1977).
11. M. RANDIĆ, J. Amer. Chem. Soc. 97, 6609 (1975).
12. A. T. BALABAN, in "Steric Fit in QSAR" (written by A. CHIRIAC, A. T. BALABAN, I. MOTOC and Z. SIMON), Lecture Notes in Chemistry, Springer, Berlin, 1979 (in press).
13. "Science of Petroleum", vol. 5, part 3, p. 58 (1973).