

Variable Connectivity Index for Cycle-Containing Structures

Milan Randić,^{*,†} Dejan Plavšić,[§] and Nella Lerš[§]National Institute of Chemistry, 1001 Ljubljana, POB 3430, The Republic of Slovenia, and
The Rudjer Bošković Institute, P.O. Box 180, HR-10002 Zagreb, The Republic of Croatia

Received August 12, 2000

In the early applications of the connectivity index the index was empirically modified for some properties of cyclic structures by subtracting 1/2 from the computed value based solely on valence of vertices in the molecular graph. In this article we looked into the origin of this heuristic adjustment of the connectivity indices for cyclic structures. We have examined the relative role of carbon atoms in cycle-containing structures by differentiating carbon atoms making up a ring and carbon atoms in exocyclic bonds. We found in the case of the boiling points of cycloalkanes and alkylcycloalkanes that contributions of “cyclic” and “acyclic” atoms to molecular additivities differ somewhat.

1. INTRODUCTION

The connectivity index, which recently celebrated its 25th anniversary,¹ was *designed* as a bond additive molecular descriptor, in which bonds of different types assume different weights.² In the case of alkanes, excluding ethane, there are three terminal bond types, CH₃–CH₂, CH₃–CH, and CH₃–C, and six interior bond types: CH₂–CH₂, CH₂–CH, CH₂–C, CH–CH, CH–C, and C–C. To these bonds weights were assigned $(mn)^{-1/2}$ according to the rule of the inverse square root of connectivities, where m and n denote the numbers of C atoms attached to the carbon atoms making up a bond. The rule is not a result of some “subjectively selected mathematical function” as has been recently misrepresented,³ but it is a solution to a set of inequalities for bond type obtained by ordering isomers according to the relative magnitudes of their boiling points, as outlined in the first paper² on the connectivity index.

The inverse square root function applies to bond types irrespective of their location in a structure. Thus, e.g., the bond type CH₂–CH₂ makes the same contribution whether carbon atoms are part of a ring or part of an exocyclic structure. Kier et al.^{4,5} already in the early applications of the connectivity index found that bonds making up rings need to be differentiated from “acyclic” bonds. They suggested an empirical “ring correction” for selected molecular properties by decreasing the computed connectivity index of a cycle-containing molecule by 1/2 for each ring. The correlations based on so empirically corrected connectivity index values produced a better agreement with experimental data. A better understanding of this empirical correction seems desirable, and hence in this paper we investigate by means of the variable connectivity index^{6,7} the relative role of carbon atoms making up a ring and exocyclic carbon atoms in cycloalkanes and alkylcycloalkanes.

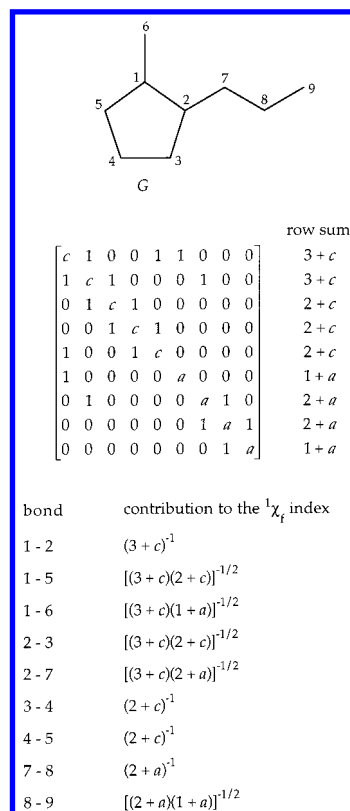


Figure 1. The construction of the variable connectivity index ${}^1\chi^f$ for 1-methyl-2-propylcyclopentane represented by the labeled hydrogen-suppressed graph G . The augmented adjacency matrix and the row sums used for calculating bond contributions are shown.

2. VARIABLE CONNECTIVITY INDEX

The simple connectivity index,^{1,2,4,5} ${}^1\chi$, can be obtained, according to Balaban,⁸ from the row sums of the adjacency matrix R_i , R_j using the inverse square root algorithm making the contribution $(R_i R_j)^{-1/2}$ for bond i , j . The variable connectivity index,^{6,7} ${}^1\chi^f$, represents a generalization of the ${}^1\chi$ index. It is also obtained by means of the inverse square root algorithm and using the row sums, but to each row sum

* Corresponding author fax: (515)292-8629. Corresponding address: 3225 Kingman Rd., Ames, IA 50014.

[†] National Institute of Chemistry.

[§] The Rudjer Bošković Institute.

Table 1. Explicit Expressions for the Variable Connectivity Index ${}^1\chi^f$ of Six Cycloalkanes and 36 Alkylcycloalkanes

no.	compound ^a	variable connectivity index ${}^1\chi^f$
1	c3	$3(2+c)^{-1}$
2	c4	$4(2+c)^{-1}$
3	M-c3	$(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
4	c5	$5(2+c)^{-1}$
5	1,1MM-c3	$(2+c)^{-1} + 2[(2+c)(4+c)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
6	1,2MM-c3	$(3+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + 2[(3+c)(1+a)]^{-1/2}$
7	E-c3	$(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
8	M-c4	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
9	c6	$6(2+c)^{-1}$
10	1,1,2MMM-c3	$[(2+c)(3+c)]^{-1/2} + [(2+c)(4+c)]^{-1/2} + [(3+c)(4+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
11	E-c4	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
12	M-c5	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
13	c7	$7(2+c)^{-1}$
14	1,1MM-c5	$3(2+c)^{-1} + 2[(2+c)(4+c)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
15	1,2MM-c5	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + 2[(3+c)(1+a)]^{-1/2}$
16	E-c5	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
17	M-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
18	c8	$8(2+c)^{-1}$
19	1,1,2MMM-c5	$2(2+c)^{-1} + [(2+c)(3+c)]^{-1/2} + [(2+c)(4+c)]^{-1/2} + [(3+c)(4+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
20	1,1,3MMM-c5	$(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + 2[(2+c)(4+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
21	P-c5	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + (2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
22	iP-c5	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
23	1,1,MM-c6	$4(2+c)^{-1} + 2[(2+c)(4+c)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
24	1,2MM-c6	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + 2[(3+c)(1+a)]^{-1/2}$
25	1,3MM-c6	$2(2+c)^{-1} + 4[(2+c)(3+c)]^{-1/2} + 2[(3+c)(1+a)]^{-1/2}$
26	1,4MM-c6	$2(2+c)^{-1} + 4[(2+c)(3+c)]^{-1/2} + 2[(3+c)(1+a)]^{-1/2}$
27	E-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
28	M-c7	$5(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
29	B-c5	$3(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + 2(2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
30	1,2MP-c5	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + [(3+c)(1+a)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + (2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
31	1,2EE-c5	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + 2[(3+c)(2+a)]^{-1/2} + 2[(2+a)(1+a)]^{-1/2}$
32	1,1,3MMM-c6	$2(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + 2[(2+c)(4+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
33	1,2,4MMM-c6	$(2+c)^{-1} + 4[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + 3[(3+c)(1+a)]^{-1/2}$
34	1,3,5MMM-c6	$6[(2+c)(3+c)]^{-1/2} + 3[(3+c)(1+a)]^{-1/2}$
35	1,4ME-c6	$2(2+c)^{-1} + 4[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + 2[(2+a)(1+a)]^{-1/2}$
36	P-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + (2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
37	iP-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
38	1,4MiP-c6	$2(2+c)^{-1} + 4[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
39	B-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + 2(2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
40	sB-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + [(3+a)(1+a)]^{-1/2} + [(3+a)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
41	iB-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
42	tB-c6	$4(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(4+a)]^{-1/2} + 3[(4+a)(1+a)]^{-1/2}$

^a The abbreviations M, E, P, iP, B, iB, sB, and tB denote methyl, ethyl, propyl, isopropyl, butyl, isobutyl, secondary butyl, and tertiary butyl, respectively, while c3, c4, c5, etc. stand for cyclopropane, cyclobutane, cyclopentane ring, respectively.

a variable is added which leads to bond weights $[(R_k + x)(R_l + y)]^{-1/2}$, where x and y describe the kind of atoms k and l , respectively. In hydrocarbons all the weights assume the form $[(R_k + x)(R_l + x)]^{-1/2}$, unless one wants to differentiate carbon atoms in a different environment. In this way one modifies the actual valence of carbon atoms which instead of being 1, 2, 3, and 4 for primary, secondary, tertiary, and quaternary carbon atom is $1+x$, $2+x$, $3+x$, and $4+x$, respectively.

To look into the relative role of carbon atoms of a cyclic part and carbon atoms of an acyclic part of an alkylcycloalkane molecule, it is necessary to differentiate them in some way. A possibility to differentiate atoms of these two groups is to associate atoms making up the ring with a parameter c (cyclic) and atoms in the acyclic part with a parameter a (acyclic). Using the parameters a and c one can tell three bond types apart in cycloalkanes and alkylcycloalkanes, C_c - C_c , C_c - C_a , and C_a - C_a , whose corresponding weights are as

follows: $[(R_i + c)(R_j + c)]^{-1/2}$, $[(R_i + c)(R_j + a)]^{-1/2}$, and $[(R_i + a)(R_j + a)]^{-1/2}$, respectively.

In Figure 1 the construction of the variable connectivity index for 1-methyl-2-propylcyclopentane is shown. The five bonds forming the cyclopentane part of the compound make the following contributions: $(3+c)^{-1}$, $2[(2+c)(3+c)]^{-1/2}$, and $2(2+c)^{-1}$. The two exocyclic bonds involving carbon atoms of the cyclopentane ring give $[(3+c)(1+a)]^{-1/2}$ and $[(3+c)(2+a)]^{-1/2}$. Finally, the remaining two bonds of the propyl group contribute $(2+a)^{-1}$ and $[(2+a)(1+a)]^{-1/2}$. In Table 1 we give the explicit expressions for ${}^1\chi^f$ of six cycloalkanes and 36 alkylcycloalkanes studied in this paper as a function of the variables a and c . Among these compounds only 1,3-dimethylcyclohexane and 1,4-dimethylcyclohexane have the same count of the corresponding bond types and hence identical expressions for ${}^1\chi^f$. For cycloalkanes and a certain number of families of structurally related alkylcycloalkanes one can write down the general

Table 2. General Expressions for the ${}^1\chi^f$ Index of Cycloalkanes and Several Families of Structurally Related Alkylcycloalkanes

compounds	expression
cn	$n(2+c)^{-1}$
M-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(1+a)]^{-1/2}$
1,1-MM-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(4+c)]^{-1/2} + 2[(4+c)(1+a)]^{-1/2}$
1,2-MM-cn	$(n-3)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + (3+c)^{-1} + 2[(3+c)(1+a)]^{-1/2}$
E-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
P-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + (2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
iP-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
B-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + 2(2+a)^{-1} + [(2+a)(1+a)]^{-1/2}$
iB-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(2+a)]^{-1/2} + [(2+a)(3+a)]^{-1/2} + 2[(3+a)(1+a)]^{-1/2}$
sB-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(3+a)]^{-1/2} + [(3+a)(1+a)]^{-1/2} + [(3+a)(2+a)]^{-1/2} + [(2+a)(1+a)]^{-1/2}$
tB-cn	$(n-2)(2+c)^{-1} + 2[(2+c)(3+c)]^{-1/2} + [(3+c)(4+a)]^{-1/2} + 3[(4+a)(1+a)]^{-1/2}$

Table 3. Values for the ${}^1\chi$ (under the Heading (0, 0)) and ${}^1\chi^f$ (under the Heading (-0.45, -0.25)) Indices of Six Cycloalkanes and 36 Alkylcycloalkanes

no.	compound	(0, 0)	(-0.45, -0.25)
1	c3	1.50000	1.93548
2	c4	2.00000	2.58065
3	M-c3	1.89385	2.37425
4	c5	2.50000	3.22581
5	1,1MM-c3	2.20711	2.72347
6	1,2MM-c3	2.30453	2.84435
7	E-c3	2.43185	2.99740
8	M-c4	2.39385	3.01941
9	c6	3.00000	3.87097
10	1,1,2MMM-c3	2.62783	3.21047
11	E-c4	2.93185	3.64257
12	M-c5	2.89385	3.66458
13	c7	3.50000	4.51613
14	1,1MM-c5	3.20711	4.01380
15	1,2MM-c5	3.30453	4.13467
16	E-c5	3.43185	4.28773
17	M-c6	3.39385	4.30974
18	c8	4.00000	5.16129
19	1,1,2MMM-c5	3.62783	4.50079
20	1,1,3MMM-c5	3.60095	4.45257
21	P-c5	3.93185	4.85916
22	iP-c5	3.80453	4.71172
23	1,1,MM-c6	3.70711	4.65896
24	1,2MM-c6	3.80453	4.77983
25	1,3MM-c6	3.78769	4.74851
26	1,4MM-c6	3.78769	4.74851
27	E-c6	3.93185	4.93289
28	M-c7	3.89385	4.95490
29	B-c5	4.43185	5.43058
30	1,2MP-c5	4.34254	5.32925
31	1,2EE-c5	4.38054	5.38098
32	1,1,3MMM-c6	4.10095	5.09773
33	1,2,4MMM-c6	4.19838	5.21860
34	1,3,5MMM-c6	4.18154	5.18728
35	1,4ME-c6	4.32570	5.37166
36	P-c6	4.43185	5.50432
37	iP-c6	4.30453	5.35688
38	1,4MiP-c6	4.69838	5.79565
39	B-c6	4.93185	6.07575
40	sB-c6	4.84254	5.98929
41	iB-c6	4.78769	5.90848
42	tB-c6	4.60517	5.69887

expressions for the ${}^1\chi^f$ index. Several such expressions are given in Table 2. In Table 3 under the heading (0, 0) the values for ${}^1\chi$ of all 42 considered compounds are listed. Note (0, 0) is the ordered pair (c , a) whose $c = 0$ and $a = 0$. The ${}^1\chi^f$ index for the pair (0, 0) is reduced to the ordinary connectivity index ${}^1\chi$. Using ${}^1\chi$ one can develop the high-grade quadratic regression model of the boiling points of the cycloalkanes and alkylcycloalkanes. All experimental boiling points data (see Table 7) were taken from ref 9. The stepwise regression equation, using ${}^1\chi$ and $({}^1\chi)^2$, is listed in

Table 4. Linear and the Quadratic Regression Models of the Boiling Points of the Studied Cycloalkanes and Alkylcycloalkanes and the Corresponding Statistical Parameters n (Number of Compounds), r (Multiple Correlation Coefficient), s (Standard Error of Estimate), and F (Fisher Ratio)

no.	${}^1\chi$	$({}^1\chi)^2$	const.	n	r	s	F
1	60.0177		-106.3326	42	0.9938	6.012	3205.0
2	96.2602	-5.4331	-162.3132	42	0.9972	4.094	3480.0
3	60.0177	-5.4331	-106.3326	42	0.9972	4.094	3480.0
	${}^1\chi^f$	$({}^1\chi^f)^2$	const.	n	r	s	F
4	49.3455		-111.9828	42	0.9965	4.5147	5714.0
5	72.8662	-2.8402	-157.1589	42	0.9985	3.0289	6372.3
6	49.3455	-2.8402	-111.1653	42	0.9985	3.0289	6372.3

Table 4. The entries shown in bold define the regression equation belonging to the orthogonalized descriptors.¹⁰⁻¹³

3. SEARCH FOR OPTIMAL VALUES OF PARAMETERS c AND a

To find optimal values of the parameters c and a for the quadratic regression model of the boiling points of the cycloalkanes and alkylcycloalkanes we initially varied c and a in the intervals $[-1, 1]$ and $[-0.75, 1.25]$, respectively, in the steps of 0.25. The intervals and the magnitude of the step were selected taking into account the following facts: (i) the value of c has to be greater than -2 due to the presence of the factor $(2+c)^{-1/2}$ in the expressions for ${}^1\chi^f$ of cycloalkanes and an alkylcycloalkane molecule which contains at least one methylene group in the ring; (ii) owing to the factor $(1+a)^{-1/2}$ in the expression for the variable connectivity index of an alkylcycloalkane molecule the value of a has to be greater than -1 ; and (iii) the quadratic regression model based on ${}^1\chi^f$ and the pair (0, 0) is of high quality, hence in the ca -plane (here c and a denote mutually perpendicular coordinate axes) the points corresponding to the pair (0, 0) and the pairs (c , a), which would improve the model's performance relative to its performance for the pair (0, 0), have to be close to each other. The comparison of the quadratic regression models based on ${}^1\chi^f$ and different pairs (c , a) is made on the basis of their standard error of estimate, s . In Table 5 we show how the value of s changes for various pairs (c , a). If one takes the difference, $c-a$, between the values of c and a in a pair (c , a) into account, then the set of pairs (c , a) in Table 5 can be partitioned into three disjoint subsets. The first subset consists of the pairs (c , a) whose $c-a = 0$. The values of s associated with these pairs are approximately 4 °C if $c \in [0, 1]$ and well over 4 °C if $c < 0$. The second subset is composed of the pairs (c , a)

Table 5. Values of the Standard Error of Estimate for Various Values of the Parameters a and c

$n = 42$	$c = -1.00$	$c = -0.75$	$c = -0.50$	$c = -0.25$	$c = 0.00$	$c = 0.25$	$c = 0.50$	$c = 0.75$	$c = 1.00$
$a = -0.75$	5.33	10.89	15.42						
$a = -0.50$	8.00	3.11	5.96	9.75	12.95	15.58	17.78		
$a = -0.25$	13.66	7.08	3.17	4.55	7.55	10.29	12.66		
$a = 0.00$		11.10	6.09	3.18	4.09	6.50	8.85		
$a = 0.25$			9.16	5.26	3.16	3.96	5.95		
$a = 0.50$	22.08		11.76	7.69	4.62	3.18	3.96	5.66	
$a = 0.75$					6.58	4.17	3.24	4.03	5.51
$a = 1.00$	24.84					5.75	3.88	3.35	4.13
$a = 1.25$							5.14	3.73	3.50

Table 6. Values of the Standard Error of Estimate for Various Values of the Parameters a and c

$n = 42$	$c = -0.80$	$c = -0.70$	$c = -0.65$	$c = -0.60$	$c = -0.55$	$c = -0.50$	$c = -0.45$	$c = -0.40$	$c = -0.35$	$c = -0.30$	$c = -0.25$	$c = -0.20$	$c = -0.15$	$c = -0.10$	$c = -0.05$	$c = 0.00$	$c = 0.05$	$c = 0.10$	$c = 0.30$
$a = -0.60$	3.75																		
$a = -0.55$	3.15	4.03																	
$a = -0.50$	3.45	3.23	3.71																
$a = -0.45$		3.10	3.12	3.50	4.09														
$a = -0.40$		3.55	3.12	3.07	3.35	3.87	4.51												
$a = -0.35$		4.29	3.57	3.14	3.04	3.26	3.71												
$a = -0.30$				3.59	3.16	3.03	3.20	3.59											
$a = -0.25$					3.59	3.17	3.03	3.16	3.51	3.99									
$a = -0.20$						3.58	3.18	3.03	3.14	3.45									
$a = -0.15$							3.56	3.18	3.04	3.13	3.41								
$a = -0.10$							4.08	3.54	3.19	3.05	3.12	3.38							
$a = -0.05$									3.52	3.18	3.05	3.12	3.36	3.72	4.15				
$a = 0.00$										3.49	3.18	3.06	3.12	3.35	3.68	4.09			
$a = 0.05$										3.90	3.46	3.18	3.07	3.13	3.34				
$a = 0.10$										4.38	3.85	3.43	3.17	3.07	3.14	3.34			
$a = 0.15$										4.89	4.29	3.79	3.41	3.17	3.08	3.15	3.35		
$a = 0.20$										5.41		4.20	3.73	3.38	3.16	3.09	3.16	3.35	
$a = 0.25$														3.68	3.36	3.16	3.10	3.18	
$a = 0.30$															3.63	3.34	3.16	3.12	
$a = 0.35$																3.59	3.32	3.16	
$a = 0.40$																		3.30	
$a = 0.45$																			3.25
$a = 0.50$																			3.18
$a = 0.55$																			3.18
$a = 0.60$																			3.27

whose $c-a < 0$. The pairs with $c-a = -0.25$ and $c \in [-0.75, 0.25]$ are of considerable importance because the values of s associated with them are mutually very close and the lowest among them, 3.11 °C, corresponding to the pair $(-0.75, -0.5)$ is also the lowest in Table 5. The values of s associated with the remaining pairs (c, a) of this subset presented in Table 5 are significantly higher than 3.11 °C. The third subset contains the pairs (c, a) whose $c-a > 0$. All the values of s associated with these pairs are substantially higher than the value connected with the pair $(0, 0)$. These results indicate that the quantity $c-a$ primarily determines the value of s associated with a pair (c, a) whose $c \in [-0.75, 0.25]$ and $a \in [-0.5, 0.5]$. In harmony with this claim the values of c and a have much less influence on the value of s than their relative value and hence it seems likely that quite a few pairs (c, a) will give the quadratic regression models of the boiling points of the cycloalkanes and alkylcycloalkanes of "optimal" properties ("minimal" value of s).

The next step in pursuit of the optimal values of c and a was the study of the pairs (c, a) corresponding in the ca -plane to the points on the intersections of the lines $a = c + 0.05k$ ($k = 1, 2, \dots, 9$) with the lines $c = -0.8 + 0.05m$ ($m = 0, 1, 2, \dots, 22$). Note in this set of pairs (c, a) the values of c and a are varied in the five times smaller steps than in the aforementioned investigations (see Table 5). Hence this

set of pairs (c, a) rendered it possible to investigate in a systematic way the relationships between s , the parameters c and a , and the quantity $c-a$ as well as to determine the pair(s) (c, a) giving the model(s) "optimal" properties. In Table 6 the values of s associated with just a part of the 207 studied pairs (c, a) are shown. These data corroborate the aforesaid assumption that the quantity $c-a$ has the dominant influence on the value of s , and they show that the optimal value of the quantity $c-a$ is -0.2 . The same value was obtained from a plot of s against $c-a$ and the corresponding regression equation. The lowest value of s associated with the studied pairs (c, a) is 3.03 °C (see Table 6), and it is connected with the pairs whose $c-a = -0.2$ and $c \in [-0.5, -0.4]$. Note this value of s is significantly lower (25.9%) than the value 4.09 °C corresponding to the pair $(0, 0)$. If one takes the degree of accuracy of the experimental boiling points listed in Table 7 into account, then the set of pairs (c, a) giving the quadratic regression models of "optimal" properties consists of the pairs whose $c-a = -0.2$ and $c \in [-0.56, -0.32]$ ("optimal" pairs (c, a)). The end numbers of the interval were determined by changing the value of c in the steps of 0.01. These results confirm the aforesaid supposition about the existence of quite a few "optimal" quadratic regression models of the boiling points of the considered cycloalkanes and alkylcycloalkanes. As c and a in all "optimal" pairs are less than zero the values of the $^1\chi^f$

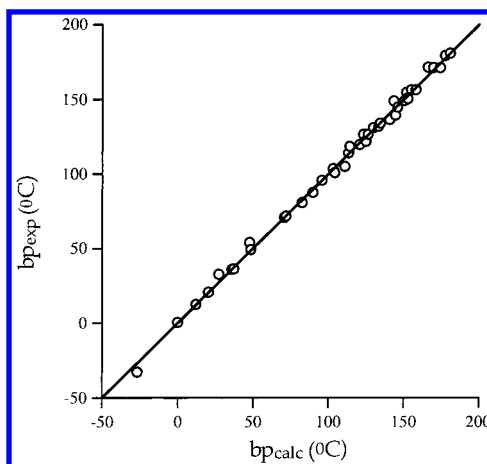
Table 7. Experimental Boiling Points (bp_{exp} ($^{\circ}C$)), the Calculated Boiling Points (bp_{calc}) Using Eq 5 from Table 4, and the Residuals

no.	compound	bp_{exp}	bp_{calc}	residual
1	c3	-32.8	-26.8	-6.0
2	c4	12.6	12.0	0.6
3	M-c3	0.7	-0.2	0.9
4	c5	49.3	48.3	1.0
5	1,1MM-c3	20.6	20.2	0.4
6	1,2MM-c3	32.6 ^a	27.1	5.4
7	E-c3	35.9	35.7	0.2
8	M-c4	36.3	37.0	-0.7
9	c6	80.7	82.3	-1.6
10	1,1,2MMM-c3	54.0	47.5	6.5
11	E-c4	70.8	70.6	0.2
12	M-c5	71.8	71.7	0.1
13	c7	118.4	114.0	4.4
14	1,1MM-c5	87.5	89.6	-2.1
15	1,2MM-c5	95.7 ^a	95.6	0.1
16	E-c5	103.5	103.1	0.4
17	M-c6	100.9	104.1	-3.2
18	c8	149.0	143.3	5.7
19	1,1,2MMM-c5	114.0	113.3	0.7
20	1,1,3MMM-c5	104.9	110.9	-6.0
21	P-c5	131.0	129.8	1.2
22	iP-c5	126.5	123.1	3.4
23	1,1,MM-c6	119.6	120.7	-1.1
24	1,2MM-c6	126.7 ^a	126.2	0.5
25	1,3MM-c6	122.3 ^a	124.8	-2.5
26	1,4MM-c6	121.9 ^a	124.8	-2.9
27	E-c6	131.9	133.2	-1.3
28	M-c7	134.0	134.2	-0.2
29	B-c5	156.6	154.8	1.8
30	1,2MP-c5	149.5 ^a	150.5	-1.0
31	1,2EE-c5	150.6 ^a	152.7	-2.1
32	1,1,3MMM-c6	136.6	140.5	-3.9
33	1,2,4MMM-c6	144.8 ^b	145.8	-1.0
34	1,3,5MMM-c6	139.5 ^c	144.4	-4.9
35	1,4ME-c6	150.8 ^a	152.3	-1.5
36	P-c6	156.7	157.9	-1.2
37	iP-c6	154.8	151.7	3.1
38	1,4MiP-c6	171.3 ^a	169.7	1.6
39	B-c6	180.9	180.7	-0.2
40	sB-c6	179.3	177.4	1.9
41	iB-c6	171.3	174.2	-2.9
42	tB-c6	171.5	165.9	5.6

^a Mean value of the boiling points of cis and trans isomers. ^b Mean value of the boiling points of (1 α , 2 α , 4 α), (1 α , 2 β , 4 α), (1 α , 2 α , 4 β), and (1 α , 2 β , 4 β) isomers. ^c Mean value of the boiling points of (1 α , 3 α , 5 α) and (1 α , 3 α , 5 β) isomers.

index based on an "optimal" pair (c , a) are always higher than the corresponding values of the ordinary connectivity index $^1\chi$. In Table 3 we give the values of the $^1\chi^f$ index based on an "optimal" pair (-0.45, -0.25) (hereafter called the $^1\chi^f$ index) for the cycloalkanes and alkylcycloalkanes. Since all "optimal" pairs (c , a) are equivalent with respect to s of the model, the pair (-0.45, -0.25) was chosen as final. The stepwise regression equation, using $^1\chi^f$ and $(^1\chi^f)^2$, is presented in Table 4. The entries shown in bold define regression equations belonging to the orthogonalized descriptors. The experimental boiling points, the computed boiling points using eq 5 from Table 4, and the residuals are listed in Table 7. Observe that all residuals are within two standard errors of the estimate with the exception of 1,1,2-trimethylcyclopropane which slightly deviates. In the case where the ordinary connectivity index $^1\chi$ is the predictor variable of the quadratic regression model, then there are three outliers. In Figure 2 the plot of bp_{exp} versus bp_{calc} is shown.

The relation between c and a in an "optimal" pair (c , a), $c-a = -0.20$, shows that "cyclic" and "acyclic" carbon

**Figure 2.** A comparison of the calculated and experimental boiling points of the studied cycloalkanes and alkylcycloalkanes.

atoms in the cycloalkanes and alkylcycloalkanes play a somewhat different role and make somewhat different contributions to the connectivity index. Since c and a are negative and $c < a$, carbon atoms making up a ring contribute a relatively larger share to the partitioning of a molecular property into bond contributions of a bond additive model. It is interesting to observe that already at the beginning of the use of the connectivity index it was noticed that *different* properties may necessitate *different* corrections, but a full realization that different properties require different weights had to wait for many years till the development of the variable connectivity index.^{6,7} A possible reason for such belated recognition of the dependence of weights on molecular property may be due to the fact that often different molecular properties have been described either using several topological indices which are not simply related or by different linear combinations of connectivity indices, which are equally difficult to compare, particularly when different descriptors appear in different regression models.

CONCLUDING REMARKS

The relative role of "cyclic" and "acyclic" carbon atoms with regard to the boiling points of the cycloalkanes and alkylcycloalkanes has been investigated by the variable connectivity index based on the variables a and c . It has been found that atoms of these two groups play a somewhat different role and that "cyclic" carbon atoms make a relatively bigger contribution to the connectivity index. The set of pairs (c , a) which give the "optimal" quadratic regression models of the boiling points of the cycloalkanes and alkylcycloalkanes consists of the pairs whose $c-a = -0.2$ and $c \in [-0.56, -0.32]$. There is a considerable difference in the value of the standard error of estimate between an "optimal" quadratic regression model and the quadratic regression model based on the ordinary connectivity index $^1\chi$. An extensive study of other properties of cycloalkanes and alkylcycloalkanes, including solubility in water, appears desirable. Apparently the present methodology, that is, use of the variable connectivity index, appears appropriate for such study, which should, however, take into account a critical evaluation of the available experimental data recently reported by Rücker and Rücker.¹⁴

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Science and Technology of the Republic of Croatia. We are grateful to Dr. Milan Šoškić (Zagreb) for useful comments.

REFERENCES AND NOTES

- (1) Symposium: Molecular Connectivity. The First Quarter Century (Kier, L. B., Hall, L. H., organizers); 220th ACS Meeting, Washington, DC, August 20–24, 2000.
- (2) Randić, M. On the Characterization of Molecular Branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
- (3) Bonchev, D. *Overall Connectivity and Topological Complexity for QSAR/QSPR in: Topological Indices and Related Descriptors in QSAR and QSPR*; Delivers, J., Balaban, A. T., Eds.; Gordon and Breach Sci. Publ.: Amsterdam, 1999; pp 361–401.
- (4) Kier, L. B.; Hall, L. H.; Murray, W. J.; Randić, M. Molecular Connectivity I. Relationship to Nonspecific Local Anesthesia. *J. Pharm. Sci.* **1976**, *64*, 1971–1974.
- (5) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
- (6) Randić, M. Novel Graph Theoretical Approach to Heteroatoms in QSAR. *Chemometrics Intel. Lab. Syst.* **1991**, *10*, 213–227.
- (7) Randić, M. On Computation of Optimal Parameters for Multivariate Analysis of Structure–Property Relationship. *J. Comput. Chem.* **1991**, *31*, 311–370.
- (8) Balaban, A. T. Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett.* **1982**, *89*, 399–404.
- (9) Lide, D. R.; Ed. *CRC Handbook of Chemistry and Physics*, 79th ed., 1998–1999; CRC Press: Boca Raton, FL, 1998.
- (10) Randić, M. Orthogonal Molecular Descriptors. *New J. Chem.* **1991**, *15*, 517–525.
- (11) Randić, M. Resolution of Ambiguities in Structure–Property Studies by Use of Orthogonal Descriptors. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 311–370.
- (12) Randić, M. Fitting of Non Linear Regressions by Orthogonalized Power Series. *J. Comput. Chem.* **1993**, *14*, 363–370.
- (13) Randić, M. Curve Fitting Paradox. *Int. J. Quantum Chem.: Quantum Biol. Symp.* **1994**, *21*, 215–225.
- (14) Rücker, G.; Rücker, Ch. On Topological Indices, Boiling Points, and Cycloalkanes. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 788–802.

CI000118Z