Semi-empirical topological method for the prediction of the chromatographic retention of *cis*- and *trans*-alkene isomers and alkanes

Vilma E.F. Heinzen, Miriam F. Soares, Rosendo A. Yunes[*]

Departamento de Química, Universidade Federal de Santa Catarina, Campus Universitário, Trindade, Florianópolis (S.C.) 88040-900, Brazil

Abstract

A new index is proposed for the prediction of the chromatographic retention of the *cis*- and *trans*-*n*-alkene isomers and alkanes. This index is based on the hypothesis that the chromatographic retention of the molecule is due to the interaction of each carbon atom with the stationary phase, and consequently the index is reduced by its neighbours' steric effects. The topological values are obtained by a numerical approximation considering the general behaviour of the chromatographic retention of the compounds. The simple linear regressions between the chromatographic retention and the index proposed for all branched alkanes and also isomers of the studied straight-chain C_5 and C_{14} alkenes (1-ene, *cis*- and *trans*-2-, 3-, 4-, 5-, 6- and 7-enes) is very good (the correlation coefficient is $r=0.9999$), and the elution sequence is correct for most of them. The models have a high predictive ability, as established by cross-validation values (r_{cv}^2). Thus, this new method, different from those already existent, can be used as complementary tool for the elucidation of the molecular structure, or prediction of the chromatographic retention of the *cis*- and *trans*-alkene isomers and branched alkanes. It could be extended with success, in the future, to the other types of compounds.   © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Retention prediction; Retention indices; Topological indices; Structure–retention relationships; Alkenes; Alkanes

1. Introduction

For the purpose of prediction, quantitative structure–retention relationships between retention data of gas chromatography (QSRRs) and different molecular descriptors have been studied by different authors [1].

Several methods of quantification of the descriptors of the molecular structure are generally used.

Efforts to describe the molecular structure by descriptors expressed in numerical values as the topological indices are interesting. Such methods use the chemical graph theory whereby the chemical structure formula that shows how bonds connect the atoms in the molecules is expressed as a mathematical graph. Each molecular graph may be represented in different ways: a matrix, a polynomial, a numeric index, etc. The most popular topological index is the molecular connectivity index in the Randic and Zagreb [2,3] approaches involving the degree D_i of the vertex i in the hydrogen-suppressed molecular

---

*Corresponding author. Tel.: +55-482-31-9000; fax: +55-482-31-9711.

graph. D_i equals the number of bonds. A generalised concept of the connectivity index has been developed by Kier and workers [4].

In the analysis of hydrocarbons (alkanes and alkenes), the molecular connectivity indices correlate well with the retention indices in the stationary phase with low polarity [1]. These indices are related to the number of atoms and how they are connected in a molecule. However, a severe limitation is that they do not distinguish the carbon atoms corresponding to *cis-* and *trans-*stereo-isomers, giving them the same values. Certainly, these types of unsaturation give different electrical properties to the involved carbon atoms.

On the other hand, chromatographically-determined empirical parameters related to molecular structure fragments' contribution to retention have been described [1]. But these parameters are usually very specific to the chromatographic system used for their measurement and for the solutes.

Several researchers have studied the gas chromatographic behaviour of alkenes with the objective of differentiating their *cis-* and *trans-*isomers, providing different correlations between the chromatographic retention index (I) and the molecular structure of these compounds, using QSRRs [5–26].

Soják and co-workers [9–11] presented a detailed analytical study on the behaviour of alkenes that considered structural increments such as the position of the double bonds and the differences between spatial isomers and column temperature. Similar considerations were used by Rang et al. [12]. Chrétien and co-workers [13–15] applied the Darc-Pelco method in their topological analysis of alkenes. Buydens et al. [16] used the complete overlap differential (CNDO/2) method to correlate the retention index and quantum-chemical parameters. Garcia-Raso et al. [17] used calculations of molecular orbital (MO) to establish which electronic parameters are required to determine the chromatographic retention index of alkenes. Bermejo and Guillen [18] studied the relationships between the retention indices and parameters related to electronic polarizability, such as molar refraction, refractive index, Van der Waals volume and molar volume of alkenes. Rohrbaugh and Jurs [19] studied the relationships between I and different structure-based descriptors of

alkenes. Voelkel [20] obtained good results when the multilinear relationships between the retention indices and the connectivity indices, dipolar moment and polarizability parameters were taken into consideration. Zhang and Hu [21] used the solubility parameters, molar volumes and number of carbon atoms to predict I for alkenes. Recently, Dimov and Osman [22] studied the influence of the conformation on the accuracy of quantitative-structure retention relationship calculations in the gas chromatography of alkenes.

However, no topological index alone is able to differentiate the *cis-* and *trans-*isomeric structures of alkenes. This fact leads us to develop a new index, organizing the topological and empirical approach on the following basic concepts:

1. In the connectivity index, only the carbon or heavy atoms in hydrogen suppressed graphs are taken into consideration.
2. The chromatographic process of separation results from the forces that operate between solute molecules and the molecule of the stationary phase.
3. The retention of the solute molecules in the case of alkanes and alkenes is due to the number and interaction of each specific carbon atom with the stationary phase. The interaction of the carbon atoms with the stationary phase is determined by its electrical properties and by the steric hindrance to this interaction by other carbon atoms attached to it. Thus, it is necessary to distinguish, for example, between a *cis-*double bond carbon and a *trans-*double bond carbon atom, because they are electrical and different sterically.
4. The values attributed to the carbon atoms were based on the results of the experimental chromatographic behaviour of the molecules that measure the electrical and steric characteristics of the carbons. These values are obtained by a numerical approximation considering the general behaviour of the chromatographic retention of the compounds.
5. We carried out this work firstly with hydrocarbons (alkanes and alkenes), with the objective of making a new index that is able to predict the chromatographic retention of *cis-* and *trans-*alkene isomers.

2. Methods

2.1. Calculation of semi-empirical topological indices for alkanes and alkenes

The calculation of the values to be attributed to the carbon atoms (vertexes of the molecular graphs) was based on the following ideas:

1. It is known that the relationship between the chromatographic retention and the number of carbon atoms for linear alkanes is logarithmic. The Kováts index, when applying the logarithm to the retention times or other parameters of retention, linearizes this relationship.
2. However, the branched alkanes do not present this linear relationship with the Kováts index, on account of the fact that the retention of the tertiary and quaternary carbon atoms is decreased by the steric effects of their neighbouring groups.
3. As a first approach, approximate values are given, based on experimental results, to the carbon atoms to separate them into primary, secondary, tertiary and quaternary. It is evident that the contribution to the retention, considering the steric effects, decreases in the indicated order. Thus, the retention index of 2,3-dimethylbutane is 568.1 [27]. If all the carbon atoms have a value of 100 as indicated by Kováts the total value should be 600. It is evident that the secondary and tertiary carbon atoms have a value less than 100.
4. From the experimental results, it is not possible to determine a constant value for each of the different carbon atoms, due to the complex variations produced by the structural changes. This fact is observed, for example, with the quaternary carbon atom in 2,2-dimethylpentane ($I=626.2$) and 3,3-dimethylpentane ($I=660.2$). Evidently, the values for the secondary and quaternary carbon atoms are not constants. In this way, approximate numeric values were attributed: 100 for the carbon atom in the methyl group in agreement with Kováts, 90 for the secondary carbon atoms, 80 for the tertiary and 70 for the quaternary, dividing by 100 to adapt them to the common topological values. It should be observed that the logarithm of these values gives a linear relation with the number of carbon atoms bonded to the

methyl group and the secondary, tertiary and quaternary carbon atoms respectively. Thus, it is necessary to observe two facts: (1) When the correlation between the values of the new method and the number of carbon atoms in the molecule of *n*-alkanes is made, as might be expected, a straight line with smaller angular coefficient than that using the Kováts retention indexes is obtained. (2) However, when very branched hydrocarbons (alkanes) are correlated, an unacceptable linearization is observed.

5. The determination of the degree or value of the steric effects present in the carbon depends on the size of the substituent group and not only on its characteristic of being primary, secondary, tertiary or quaternary. As the tendency of the deviations are logarithmic, it is also necessary to add the value of the logarithm of adjacent carbon atom(s) according to whether it has a primary, secondary, tertiary or quaternary characteristic. Thus, the new semi-empirical topological index (I_{ET}) is expressed as:

$$I_{ET} = \sum C_i + \delta_i$$

$$\delta_i = \log C_1 + \log C_2 + \log C_3 + \log C_4$$

where C_i is the value attributed to each carbon atom i in the molecule and δ_i is the sum of the logarithm of the value of each adjacent carbon atom (C_1, C_2, C_3 and C_4).

6. The approach to the determination of the retention index of alkanes is primarily based on the supposition that the retention of a carbon atom of a molecule is caused by the dispersion forces of its interaction with the stationary phase that is reduced by neighbouring steric effects. For alkenes, the electrostatic forces are also involved. The interaction of these forces with the other structural factors makes it very difficult to predict its effect, based on theoretical considerations, in order to give different values to the *cis*- and *trans*-carbon atoms. For this reason, we prefer to take these values from experience. The C_i values were given based on the experimental I of alkenes with regard to whether they were *cis*- or *trans*- and depending on the position of the double bond in the chain.

2.2. Data set

The retention indices of alkenes were taken from the literature [14]. They were determined at 100°C by high-performance capillary gas chromatography on a stainless-steel capillary column (200 m×0.25 mm I.D.) coated with squalane. The reproducibility of the measurements, calculated as the standard deviation, was a 0.18 Kováts index with C_6–C_{10} n-alkenes and a 0.10 Kováts index with higher n-alkenes.

The experimental retention GC indices for 157 alkanes, measured using squalane at 100°C, were taken from Ref. [27].

2.3. Regression analysis

All calculations of regression analysis were carried out using the *Origin* program. To test the quality of the regression equation, the correlation coefficient (r), the coefficient of determination (r^2), the standard deviation (SD) and the test of null hypothesis (F-test) were utilized as statistical parameters. The validity of the model is tested in a cross-validation-like procedure, with computation of r_{cv}^2 [28].

3. Results and discussion

The molecule of 2-methylbutane is taken as an example of the calculation of the new semi-empirical topological index.
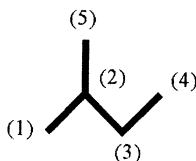
$$C(1) = 1.0 + (\log 0.8) = 1.0 - 0.0969 = 0.9031$$

$$C(2) = 0.8 + (\log 0.9 + \log 1.0 + \log 1.0)$$
$$= 0.8 - 0.0458 = 0.7542$$

$$C(3) = 0.9 + (\log 0.8 + \log 1.0) = 0.9 - 0.0969$$
$$= 0.8031$$

$$C(4) = 1.0 + (\log 0.9) = 1.0 - 0.0458 = 0.9542$$

$$C(5) = 1.0 + (\log 0.8) = 1.0 - 0.0969 = 0.9031$$

$$I_{ET} = 4.3177$$

Table 1
Values of C_i and log C_i for primary, secondary, tertiary and quaternary carbon atoms in alkanes and carbon atoms in the carbonic chain of *cis*- and *trans*-alkene isomers

Fragment	Double bond position	Value	C_i	Log C_i
–CH_3	–	100	1.0	–
–CH_2–	–	90	0.9	−0.04576
–CH<	–	80	0.8	−0.09691
>C<	–	70	0.7	−0.1549
CH_2=; –CH=	1C	89.75	0.8975	−0.04696
–CH=*trans*	2C	89.5	0.895	−0.04818
cis		91.0	0.910	−0.04096
–CH=*trans*[a]	3C	87.5	0.875	−0.05799
cis[a]		88.5	0.885	−0.05306
–CH=*trans*[a]	4C	86.5	0.865	−0.06298
cis[a]		87.0	0.870	−0.06048
–CH=*trans*	5C	86.5	0.865	−0.06298
cis		85.5	0.855	−0.06803
–CH=*trans*	6C	86.0	0.860	−0.06550
cis		85.0	0.850	−0.07058
–CH=*trans*	7C	85.75	0.8575	−0.06676
cis		84.5	0.845	−0.07314

[a] Above 10 carbon atoms in the carbonic chain, the values for *cis*- and *trans*-alkene isomers should be inverted.
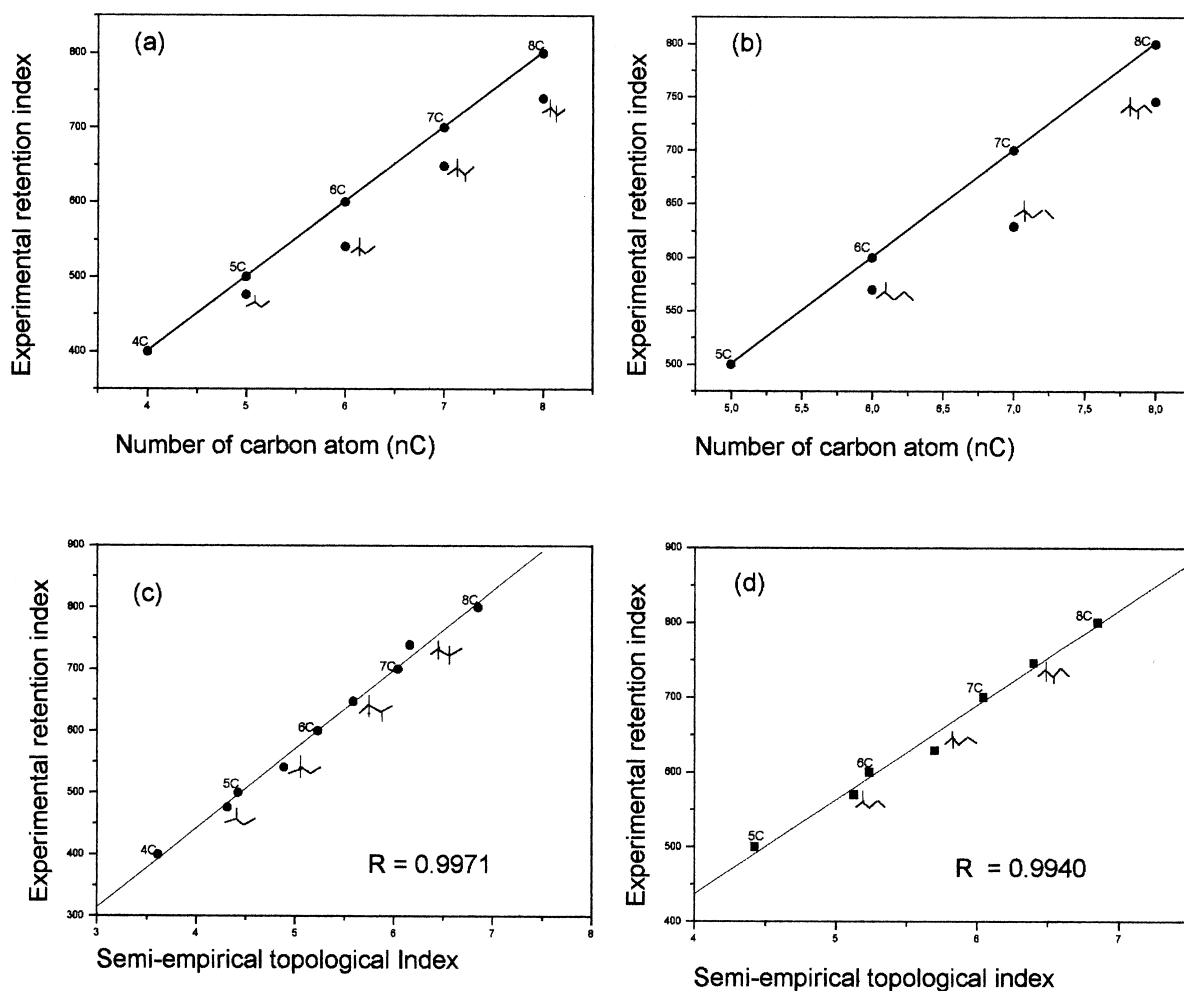
Fig. 1. Simple linear regression between experimental retention indices (using squalane as stationary phase [23]) for linear and branched alkanes (nC) and: (a), (b) number of carbon atoms; (c), (d) semi-empirical topological index, I_{ET}.

The values of C_i and log C_i for primary, secondary, tertiary and quaternary carbon atoms in *cis*- and *trans*-alkene isomers are in Table 1.

Fig. 1(a and b) shows that the relationship between the Kováts index and the number of carbon atoms for branched alkanes is not linear.

The simple linear correlation obtained between the values of this new semi-empirical topological index and the values of the experimental retention index for linear and branched alkanes is good (see Fig. 1(c and d) and Tables 3 and 4). This demonstrates that in

Table 2
Statistical characteristics of relationship between the retention indices of 157 alkanes and the Wiener numbers, $^3W_{CH}$, the connectivity indices, ^1X and semi-empirical topological indices, I_{ET}

Parameter	$^3W_{CH}$	^1X	I_{ET}
a	119.5±21.8	186.9±2.2	116.8±1.3
b	−202.9±57.0	69.81±9.31	−19.05±0.37
c	0.2947±0.0174	–	–
r	0.9862	0.9902	0.9901
s	31.0	26.0	26.2
F	2728	7827	7800

Table 3
Values of experimental retention indices, I_{EXP}, (squalane 100°C) [14], calculated retention indices, I_{CALC} (with I_{ET}), $\Delta I = I_{EXP} - I_{CALC}$ and values of semi-empirical topological index, I_{ET} calculated for different alkenes with C_5–C_{14} carbon atoms in the linear chain

No.	Compound	I_{EXP}	I_{CALC}	ΔI	I_{ET}
1	1-Pentene	482.6	483.0	−0.4	4.271
2	1-Hexene	583.8	582.3	1.5	5.0795
3	1-Heptene	683.1	681.6	1.5	5.888
4	1-Octene	782.6	780.9	1.7	6.6965
5	1-Nonene	882.5	880.2	2.3	7.505
6	1-Decene	982.5	979.6	2.9	8.3135
7	1-Undecene	1082.4	1078.9	3.5	9.122
8	1-Dodecene	1183	1178.2	4.8	9.9305
9	1-Tridecene	1283.1	1277.5	5.6	10.739
10	1-Tetradecene	1383.2	1376.8	6.4	11.5475
11	*trans*-2-Pentene	499.5	499.5	−0.0	4.4057
12	*cis*-2-Pentene	505.4	506.7	−1.3	4.4645
13	*trans*-2-Hexene	596.5	598.8	−2.3	5.2142
14	*cis*-2-Hexene	604.9	606.1	−1.2	5.273
15	*trans*-2-Heptene	698.7	698.2	0.5	6.0227
16	*cis*-2-Heptene	704.7	705.4	−0.7	6.0815
17	*trans*-2-Octene	797.5	797.5	0.0	6.8312
18	*cis*-2-Octene	803.2	804.7	−1.5	6.89
19	*trans*-2-Nonene	896.4	896.8	−0.4	7.6397
20	*cis*-2-Nonene	901.9	904.0	−2.1	7.6985
21	*trans*-2-Decene	996.7	996.1	0.6	8.4482
22	*cis*-2-Decene	1001.7	1003.3	−1.6	8.507
23	*trans*-2-Undecene	1096.6	1095.4	1.2	9.2567
24	*cis*-2-Undecene	1101.5	1102.7	−1.2	9.3155
25	*trans*-2-Dodecene	1196.9	1194.8	2.1	10.0652
26	*cis*-2-Dodecene	1201.7	1202.0	−0.3	10.124
27	*trans*-2-Tridecene	1297	1294.1	2.9	10.8737
28	*cis*-2-Tridecene	1301.6	1301.3	0.3	10.9325
29	*trans*-2-Tetradecene	1396.9	1393.4	3.5	11.6822
30	*trans*-3-Hexene	591.1	589.1	2.0	5.1348
31	*cis*-3-Hexene	593.3	594.0	−0.7	5.1746
32	*trans*-3-Heptene	687.4	688.4	−1.0	5.9432
33	*cis*-3-Heptene	692	693.3	−1.3	5.9831
34	*trans*-3-Octene	788.2	787.7	0.5	6.7517
35	*cis*-3-Octene	789.8	792.6	−2.8	6.7916
36	*trans*-3-Nonene	886.4	887.0	−0.6	7.5602
37	*cis*-3-Nonene	887.5	891.9	−4.4	7.6001
38	*trans*-3-Decene	985.8	986.3	−0.5	8.3687
39	*cis*-3-Decene	985.8	991.2	−5.4	8.4086
40	*trans*-3-Undecene	1085.4	1090.6	−5.2	9.2171
41	*cis*-3-Undecene	1085.3	1085.7	−0.4	9.1772
42	*trans*-3-Dodecene	1185.1	1189.9	−4.8	10.0256
43	*cis*-3-Dodecene	1185.1	1185.0	0.1	9.9857
44	*trans*-3-Tridecene	1284.9	1289.2	−4.3	10.8341
45	*cis*-3-Tridecene	1284.4	1284.3	0.1	10.7942
46	*trans*-3-Tetradecene	1384.6	1388.5	−3.9	11.6426
47	*cis*-3-Tetradecene	1384.1	1383.6	0.5	11.6027
48	*trans*-4-Octene	784.1	782.8	1.3	6.712
49	*cis*-4-Octene	788.2	785.3	2.9	6.732
50	*trans*-4-Nonene	884.2	882.1	2.1	7.5205

Table 3. (*Continued*)

No.	Compound	I_{EXP}	I_{CALC}	ΔI	I_{ET}
51	*cis*-4-Nonene	885.4	884.6	0.8	7.5405
52	*trans*-4-Decene	982.5	981.5	1.0	8.329
53	*cis*-4-Decene	982.8	983.9	−1.1	8.349
54	*trans*-4-Undecene	1081.1	1083.2	−2.1	9.1575
55	*cis*-4-Undecene	1080.5	1080.8	−0.3	9.1375
56	*trans*-4-Dodecene	1180.6	1182.6	−2.0	9.966
57	*cis*-4-Dodecene	1179.6	1180.1	−0.5	9.946
58	*trans*-4-Tridecene	1279.9	1281.9	−2.0	10.7745
59	*cis*-4-Tridecene	1278.6	1279.4	−0.8	10.7545
60	*trans*-4-Tetradecene	1379.3	1381.2	−1.9	11.583
61	*cis*-4-Tetradecene	1377.7	1378.7	−1.0	11.563
62	*trans*-5-Decene	984.1	981.5	2.6	8.329
63	*cis*-5-Decene	981.6	976.5	5.1	8.2886
64	*trans*-5-Undecene	1081.8	1080.8	1.0	9.1375
65	*cis*-5-Undecene	1078.2	1075.8	2.4	9.0971
66	*trans*-5-Dodecene	1180.6	1180.1	0.5	9.946
67	*cis*-5-Dodecene	1175.6	1175.1	0.5	9.9056
68	*trans*-5-Tridecene	1279.5	1279.4	0.1	10.7545
69	*cis*-5-Tridecene	1273.8	1274.5	−0.7	10.7141
70	*trans*-5-Tetradecene	1378.4	1378.7	−0.3	11.563
71	*cis*-5-Tetradecene	1372	1373.8	−1.8	11.5226
72	*trans*-6-Dodecene	1179.6	1177.6	2.0	9.9256
73	*cis*-6-Dodecene	1175	1172.7	2.3	9.8856
74	*trans*-6-Tridecene	1277.4	1276.9	0.5	10.7341
75	*cis*-6-Tridecene	1271.2	1272.0	−0.8	10.6941
76	*trans*-6-Tetradecene	1375.7	1376.2	−0.5	11.5426
77	*cis*-6-Tetradecene	1368.6	1371.3	−2.7	11.5026
78	*trans*-7-Tetradecene	1374.5	1375.0	−0.5	11.5328
79	*cis*-7-Tetradecene	1366.7	1368.8	−2.1	11.4822

this case the fundamental factor that governs the chromatographic retention is the steric effect of the retention of each carbon atom in the molecule.

The simple linear correlation between the chromatographic retention indices, I (squalane, 100°C) [27] and the new semi-empirical topological index, I_{ET} for 157 branched alkanes (with methyl, ethyl, dimethyl, trimethyl and tetramethyl, diethyl groups, with the carbonic chain having C_1–C_{12} carbon atoms) is very good: the statistical parameters are in Table 2.

We also compared our results with those obtained from the paper of Trinajstic et al. [27]. The simple correlation with the same compounds using the three-dimensional (3D) Wiener number, $^3W_{CH}$ ($I_{CALC} = aW^c + b$), and the molecular connectivity index of first order, ^1X, ($I_{CALC} = a^1X + b$), was also good. The statistical characteristics of all three correlations are of the same good quality and the values of the statistical parameters are given in Table 2.

All of $^3W_{CH}$, ^1X and I_{ET} model the shape and the size of alkanes [27] and they are highly intercorrelated indices. Thus, these indices express similar constitutional information. The simple linear correlation between the chromatographic retention indices [19] and the new semi-empirical topological index, I_{ET} for 229 branched alkanes (with methyl, ethyl, propyl, butyl, dimethyl, trimethyl and tetramethyl groups, with the carbonic chain having C_4–C_{24} carbon atoms) is very good: the correlation coefficient, r is 0.9983. The correlation with the same compounds using the Wiener index and the molecular connectivity index of first order, ^1X of Kier and Hall is also good (0.9980 and 0.9985), the values of the correlation coefficient being of the same order.

The values attributed to the carbon atoms C_i for alkenes (Table 1) were given observing their chromatographic retention behaviour, for example:

1. The retention indices of the 1-alkenes (squalane, 100°C) [14] are generally lower than those of its corresponding alkanes, independent of the size of the chain (C_5–C_{14} n-alkenes were studied). The retention indices of the 1-alkenes are lower than those of cis-2, trans-2-alkenes, cis-3 and trans-3-alkenes but they are close to the values of cis-4, trans-4-alkenes.

2. The trans-2-alkenes are eluted before their corresponding cis-isomers. The trans-2-alkenes have a retention index which is slightly lower than the corresponding alkanes, but the cis-2-alkenes' retention index is higher.

3. The cis- and trans-3 and 4-isomers change their elution sequence gradually when the number of carbon atoms in the molecule is increased (for individual homologous series). With smaller al-kenes ($<C_{10}$ n-alkenes), the trans-3 and 4-alkenes are eluted first, but with larger alkenes ($>C_{10}$ n-alkenes) the cis-3 and 4-alkenes are eluted first.

The correlation between the Kováts retention indices [14] (squalane, 100°C) and the new semi-empirical topological index, I_{ET} for 79 alkenes with C_5–C_{14} carbon atoms in the molecule, having the double bond in the positions 1, 2, 3, 4, 5, 6 and 7, is of fairly good quality. The best equation is linear:

$$I_{CALC} = aI_{ET} + b \tag{1}$$

The following values of statistical parameters are obtained: $n = 79$, $b = -41.7054 \pm 1.1276$, $a = 122.8446 \pm 0.1247$, $r = 0.99996$, SD $= 2.3541$.

According to Muresan et al. [28], in a cross-validation procedure, the r_{cv}^2 value is 0.99991. The elution order of the compounds is correct and the majority of the differences between the experimental
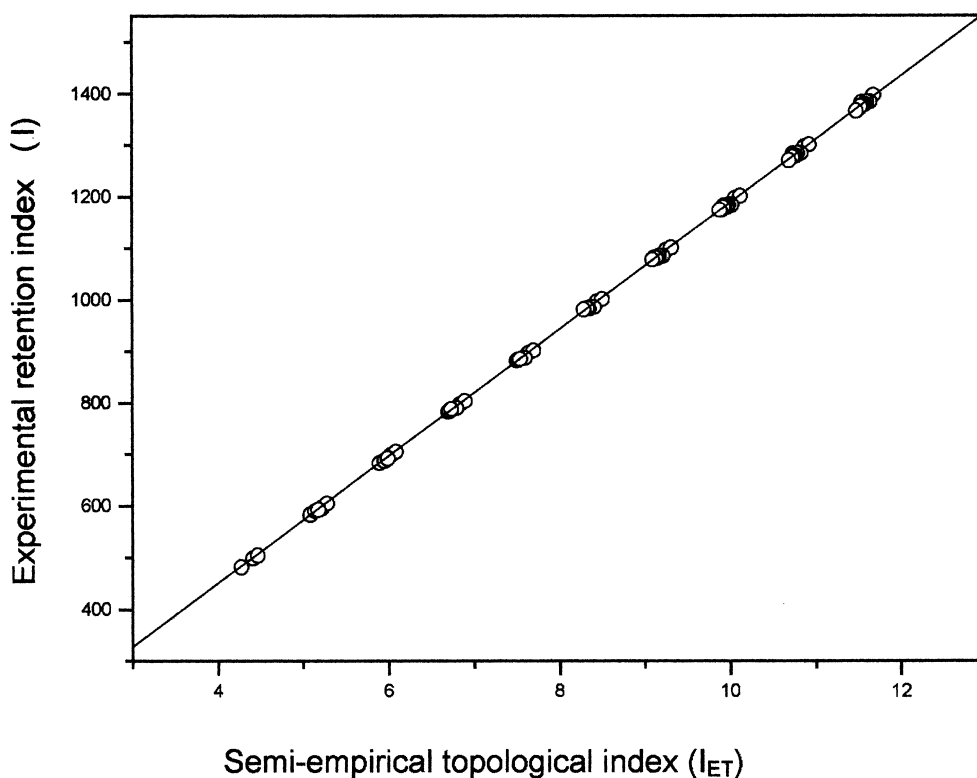

Fig. 2. Simple linear regression between experimental retention indices, I [14], for cis- and trans-alkene isomers (C_5–C_{14}) and semi-empirical topological index (I_{ET}).

Table 4

Experimental retention indices, I_{EXP} [27], 3D Wiener numbers, $^3W_{CH}$, connectivity indices, ^1X and semi-empirical topological indices, I_{ET} of different alkanes

No.	Compound	$^3W_{CH}$	^1X	$I_{EXP.}$	I_{ET}
1	Methane	0	0.000	100	1.0000
2	Ethane	1	1.000	200	2.0000
3	Propane	4	1.141	300	2.8084
4	Butane	10	1.914	400	3.6168
5	2,2-Dimethylpropane	16	2.000	412.6	4.0804
6	2-Methylbutane	18	2.270	475.4	4.3177
7	Pentane	20	2.414	500	4.4253
8	2,2-Dimethylbutane	28	2.561	537.6	4.8888
9	2,3-Dimethylbutane	29	2.643	568.1	5.0186
10	2-Methylpentane	32	2.770	569.8	5.1261
11	3-Methylpentane	31	2.808	584.6	5.1262
12	Hexane	35	2.914	600	5.2338
13	2,2-Dimethylpentane	46	3.061	626.3	5.6972
14	2,4-Dimethylpentane	48	3.126	630.1	5.8270
15	2,2,3-Trimethylbutane	42	2.943	641.1	5.5897
16	3,3-Dimethylpentane	44	3.121	660.2	5.6973
17	2-Methylhexane	52	3.270	666.8	5.9346
18	2,3-Dimethylpentane	46	3.181	672.5	5.8270
19	3-Methylhexane	50	3.308	676.5	5.9346
20	3-Ethylpentane	48	3.346	686.6	5.9346
21	2,2,4-Trimethylpentane	66	3.417	690.9	6.3981
22	Heptane	56	3.414	700	6.0423
23	2,2-Dimethylhexane	71	3.561	719.9	6.5057
24	2,2,3-Trimethylpentane	63	3.481	738.6	6.3981
25	2,3-Dimethylhexane	70	3.681	760.8	6.6354
26	2,3,3-Trimethylpentane	62	3.504	761.4	6.3981
27	3-Ethyl-2-methylpentane	67	3.719	762.4	6.6355
28	2-Methylheptane	79	3.770	765	6.7431
29	4-Methylheptane	75	3.808	767.4	6.7430
30	3,4-Dimethylhexane	68	3.719	771.6	6.6354
31	3-Methylheptane	76	3.808	772.6	6.7431
32	2,2,4,4-Tetramethylpentane	88	3.707	774.6	6.9692
33	3,3-Dimethylhexane	67	3.621	775.7	6.5057
34	2,2,4-Trimethylhexane	98	3.955	777.3	7.2066
35	2,2,5-Trimethylhexane	94	3.917	790.7	7.2065
36	Octane	84	3.914	800	6.8508
37	2,4,4-Trimethylhexane	92	3.977	809.7	7.3363
38	2,3,5-Trimethylhexane	96	4.037	813.2	7.3363
39	2,2-Dimethylheptane	104	4.061	816.2	7.3142
40	2,2,5,5-Tetramethylhexane	127	4.207	820.1	7.7776
41	2,4-Dimethylheptane	102	4.164	821.2	7.4439
42	2,2,3,4-Tetramethylpentane	86	3.854	821.9	7.0990
43	2,2,3-Trimethylhexane	92	3.981	823.3	7.2065
44	2,2-Dimethyl-3-ethylpentane	88	4.019	824.4	7.2066
45	4-Ethyl-2-methylhexane	98	4.202	824.9	7.4439
46	2,6-Dimethylheptane	108	4.126	827.5	7.4439
47	4,4-Dimethylheptane	96	4.121	828.6	7.3141
48	2,5-Dimethylheptane	104	4.164	833.7	7.4440
49	3,5-Dimethylheptane	100	4.202	834.4	7.4440
50	3,3-Dimethylheptane	98	4.121	837.5	7.3142
51	2,4-Dimethyl-3-ethylpentane	90	4.091	838.4	7.3363

Table 4. (*Continued*)

No.	Compound	$^3W_{CH}$	^1X	$I_{EXP.}$	I_{ET}
52	2,3,3-Trimethylhexane	90	4.004	841.7	7.2065
53	3-Ethyl-2-methylhexane	96	4.219	844.4	7.4439
54	2,3,4-Trimethylhexane	92	4.091	849.7	7.3363
55	3,3,4-Trimethylhexane	88	4.042	855.1	7.2065
56	2,3-Dimethylheptane	102	4.181	855.5	7.4439
57	3-Ethyl-4-methylhexane	94	4.257	855.6	7.4439
58	2,2,3,3-Tetramethylpentane	82	3.811	855.8	6.9692
59	3-Ethyl-3-methylhexane	92	4.182	856	7.3141
60	3,4-Dimethylheptane	98	4.129	858	7.4438
61	4-Ethylheptane	102	4.346	858.2	7.5514
62	2,3,3,4-Tetramethylpentane	84	3.887	861.1	7.0990
63	4-Methyloctane	108	4.308	863.3	7.5515
64	2-Methyloctane	114	4.270	864.8	7.5516
65	3-Ethylheptane	104	4.346	867.4	7.5515
66	2,4,6-Trimethylheptane	135	4.520	870.1	8.1448
67	3-Methyloctane	110	4.308	870.8	7.5516
68	2,2,4,5-Tetramethylhexane	124	4.327	872.1	7.9074
69	2,2,6-Trimethylheptane	139	4.417	873	8.0150
70	2,2,3,5-Tetramethylhexane	123	4.337	873.3	7.9074
71	2,3-Dimethyl-3-ethylpentane	86	4.065	875	7.2066
72	2,2,4-Trimethylheptane	131	4.455	875.7	8.0150
73	2,2,5-Trimethylheptane	134	4.455	878.1	8.0150
74	3,3-Diethylpentane	88	4.243	880.2	7.3142
75	2,2-Dimethyl-4-ethylhexane	126	4.493	881.3	8.0150
76	2,2,4,4-Tetramethylhexane	119	4.268	888.6	7.7777
77	2,4,4-Trimethylheptane	127	4.477	889.4	8.0150
78	2,5-Dimethyl-3-ethylhexane	127	4.575	891.4	8.1448
79	2,5,5-Trimethylheptane	131	4.477	891.7	8.0150
80	Nonane	120	4.414	900	7.6593
81	2,2-Dimethyl-3-ethylhexane	122	4.519	902.1	8.0150
82	2,3,3,5-Tetramethylhexane	120	4.360	903.3	7.9074
83	3-Ethyl-2,2,4-trimethylpentane	115	4.392	903.9	7.9074
84	2,4,5-Trimethylheptane	130	4.575	906.7	8.1447
85	4-Ethyl-2-methylheptane	134	4.702	907.4	8.2523
86	3,3,5-Trimethylheptane	126	4.515	907.7	8.0150
87	2,2,3,4-Tetramethylhexane	118	4.392	908.8	7.9074
88	2,3,5-Trimethylheptane	131	4.575	912.9	8.1448
89	2,2,3-Trimethylheptane	130	4.481	914.4	8.0150
90	2,2-Dimethyloctane	146	4.561	914.9	8.1227
91	2,4-Dimethyl-3-isopropylpentane	117	4.464	915.1	8.0372
92	3-lsopropyl-2-methylhexane	124	4.591	915.5	8.1447
93	2,4-Dimethyloctane	142	4.664	915.8	8.2524
94	4,4-Dimethyloctane	134	4.621	918	8.2386
95	2,3,6-Trimethylheptane	136	4.537	919	8.1447
96	2,4-Dimethyl-4-ethylhexane	122	4.538	920.7	8.0150
97	2,2,3,4,4-Pentamethylpentane	111	4.155	921.7	7.6701
98	3,5-Dimethyloctane	143	4.6664	921.8	8.2524
99	2,5-Dimethyloctane	138	4.702	921.8	8.2523
100	2,3,4,5-Tetramethylhexane	121	4.464	923.1	8.0372
101	5-Ethyl-2-methylheptane	138	4.702	924.8	8.2523
102	4-lsopropylheptane	131	4.719	925	8.2523
103	2,7-Dimethyloctane	151	4.626	928.5	8.2524
104	2,2,3,3-Tetramethylhexane	115	4.311	928.8	7.7776

Table 4. (*Continued*)

No.	Compound	$^3W_{CH}$	^1X	$I_{EXP.}$	I_{ET}
105	3,6-Dimethyloctane	141	4.702	929	8.2524
106	2,4-Dimethyl-3-ethylhexane	122	4.629	929.8	8.1447
107	2,6-Dimethyloctane	146	4.664	931.5	8.2524
108	2,3,3-Trimethylheptane	127	4.504	931.7	8.0150
109	3,3-Dimethyloctane	138	4.621	932	8.1227
110	3,4,4-Trimethylheptane	122	4.542	932.2	8.0149
111	2,3,4-Trimethylheptane	128	4.591	933.4	8.1447
112	2,3,4,4-Tetramethylhexane	116	4.415	935	7.9074
113	4-Ethyl-3-methylheptane	129	4.757	935.7	8.2523
114	3,4-Dimethyloctane	137	4.719	936	8.2523
115	3,3,4-Trimethylheptane	123	4.542	936.6	8.0149
116	4-Ethyl-4-methylheptane	126	4.682	937.6	8.1225
117	3,3-Dimethyl-4-ethylhexane	118	4.580	937.8	8.0150
118	3-Ethyl-4-methylheptane	130	4.757	940.5	8.2523
119	3-Ethyl-2-methylheptane	134	4.719	941	8.2524
120	4,5-Dimethyloctane	135	4.719	943.1	8.2522
121	3,4,5-Trimethylheptane	125	4.629	945	8.1447
122	3,4-Diethylhexane	125	4.795	945.8	8.2524
123	2,3,3,4-Tetramethylhexane	115	4.425	949.1	7.9074
124	2,3-Dimethyl-4-ethylhexane	123	4.629	949.4	8.1448
125	4-Ethyloctane	141	4.846	951.5	8.3599
126	2,3-Dimethyloctane	143	4.681	952.1	8.2524
127	2-Ethyl-2-methylheptane	129	4.621	953	8.1807
128	2,2,3,3,4-Pentamethylpentane	108	4.193	953.4	7.6701
129	3,3-Diethylhexane	121	4.743	954.1	8.1226
130	5-Methylnonane	149	4.808	957.4	8.3600
131	4-Methylnonane	150	4.808	960	8.3600
132	2-Methylnonane	158	4.770	963.9	8.3601
133	3-Ethyloctane	145	4.846	964	8.3600
134	3,4-Dimethyl-3-ethylhexane	117	4.603	964.6	8.0150
135	3-Ethyl-2,2,3-trimethylpentane	110	4.371	965.7	7.7777
136	3-Ethyl-2,3,4-trimethylpentane	112	4.447	969.4	7.9074
137	3-Methylnonane	153	4.808	969.6	8.3601
138	3,3,4,4-Tetramethylhexane	111	4.371	983.7	7.7776
139	Decane	165	4.914	1000	8.4678
140	Undecane	220	5.414	1100	9.2763
141	6-Methylundecane	261	5.808	1151.8	9.9770
142	4-Methylundecane	265	5.808	1158.6	9.9770
143	2-Methylundecane	277	5.770	1164	9.9771
144	3-Methylundecane	270	5.808	1169.6	9.9771
145	Dodecane	286	5.914	1200	10.0848
146	5,7-Dimethylundecane	308	6.202	1190.4	10.6778
147	4,6-Dimethylundecane	310	6.202	1193	10.6778
148	3,5-Dimethylundecane	316	6.202	1207.2	10.6779
149	2,4-Dimethylundecane	326	6.164	1208.2	10.6779
150	2,5-Dimethylundecane	324	6.164	1210.4	10.6778
151	2,6-Dimethylundecane	324	6.164	1210.4	10.6778
152	2,7-Dimethylundecane	326	6.164	1215.8	10.6778
153	5,6-Dimethylundecane	306	6.219	1223.4	10.6777
154	4,5-Dimethylundecane	310	6.219	1230.4	10.6777
155	2,9-Dimethylundecane	336	6.164	1232.6	10.6779
156	3,4-Dimethylundecane	318	6.219	1247	10.6778
157	2,3-Dimethylundecane	330	6.181	1251.4	10.6779

and calculated values are smaller than 5. The *cis*- and *trans*- structures are differentiated with this index, which does not take place when other indices are used.

The values of semi-empirical topological indices, I_{ET}, experimental retention indices, I_{EXP}, and calculated retention indices, I_{CALC} and ΔI ($I_{EXP} - I_{CALC}$) are listed in Table 3. A plot of I_{EXP} vs. I_{CALC} is shown in Fig. 2.

4. Conclusion

The model proposed in this work has a high predictive ability as established by cross-validation values (r_{cv}^2), showing that the new method may be used as a complementary tool for the elucidation of the molecular structure or for the prediction of the chromatographic retention of the alkenes with *cis*- and *trans*-isomeric structures and branched alkanes, thus improving upon the well-known topological methods available up to the present. There is a real possibility that the method could be extended with success in the future to the other types of compounds.

References

[1] R. Kalisman, Quantitative Structure – Chromatographic Retention Relationships, Wiley, New York, 1987.
[2] M. Randic, J. Am. Chem. Soc. 97 (1975) 6609.
[3] I. Gutman, N. Trinajstic, Chem. Phys. Lett. 17 (1973) 535.
[4] L.B. Kier, L.H. Hall, Molecular Connectivity in Chemistry and Drug Research, Academic Press, New York, 1976.
[5] J. Takács, Zs. Tálas, I. Bernáth, Gy. Czakó, A. Fisher, J. Chromatogr. 67 (1972) 203.
[6] J. Sanz, J. Calderón, M.V. Dabrio, Anal. Chem. 75 (1979) 408.
[7] D. Papazova, N. Dimov, J. Chromatogr. 137 (1977) 259.
[8] N. Dimov, D. Papazova, Chromatographia 12 (1979) 720.
[9] L. Soják, J. Hrivnák, P. Majer, J. Janák, Anal. Chem. 45 (1973) 293.
[10] L. Soják, J. Hrivnák, I. Ostrovsky, J. Janák, J. Chromatogr. 91 (1974) 613.
[11] L. Soják, J. Krupcik, J. Janák, J. Chromatogr. 195 (1980) 43.
[12] S. Rang, K. Kuningas, A. Orav, O. Eisen, Chromatographia 10 (1977) 55.
[13] J.R. Chrétien, J.E. Dubois, Anal. Chem. 49 (1977) 747.
[14] J.E. Dubois, J.R. Chrétien, L. Soják, J.A. Rijks, J. Chromatogr. 194 (1980) 121.
[15] J.R. Chrétien, J.E. Dubois, R.F. Hirsch, R.J. Gaydosh, J. Chromatogr. 207 (1981) 115.
[16] L. Buydens, D.L. Massart, P. Geerlings, Anal. Chem. 55 (1983) 738.
[17] A. Garcia-Raso, F. Saura-Calixto, M.A. Raso, J. Chromatogr. 302 (1984) 107.
[18] J. Bermejo, M.D. Guillen, Int. J. Environ. Anal. Chem. 23 (1985) 77.
[19] R.H. Rohrbaugh, P.C. Jurs, Anal. Chem. 57 (1985) 2770.
[20] A. Voelkel, Chromatographia 25 (1988) 655.
[21] Z. Hu, H. Zhang, J. Chromatogr. A 653 (1993) 275.
[22] N. Dimov, A. Osman, J. Chromatogr. A 773 (1997) 368.
[23] N. Dimov, A. Osman, Anal. Chim. Acta 323 (1996) 15.
[24] S. Boneva, K. Vassilev, Chromatographia 47 (1998) 230.
[25] S. Rang, K. Kuningas, T. Strenze, A. Orav, O. Eisen, J. Chromatogr. 406 (1987) 75.
[26] F. Khorasheh, M.R. Gray, M. Selucky, J. Chromatogr. 481 (1989) 1.
[27] N. Bosnjak, Z. Mihalic, N. Trinajstic, J. Chromatogr. 540 (1991) 430.
[28] S. Muresan, C. Bologa, M. Mracec et al., J. Molec. Struct. (Theochem.) 342 (1995) 161.