A NEW APPROACH TO THE PREDICTION OF GAS CHROMATOGRAPHIC RETENTION INDICES FROM PHYSICO-CHEMICAL CONSTANTS*

Igor G. ZENKEVICH and Lyudmila M. KUZNETSOVA
Department of Organic Chemistry, Leningrad University, 199034 Leningrad, U.S.S.R.

Received September 27, 1990
Accepted January 18, 1991

A general approach is proposed to the calculation of gas chromatographic retention indices (RI) of organic compounds on standard polydimethylsiloxane stationary phases based on their principal physico-chemical constants such as the boiling temperature, molar refraction or molecular weight. A combination of logical criteria was established for comparing functions of the above parameters for the identification of substances whose RI values are only determined by their boiling temperatures. It is demonstrated that within homologous series, the dependence of the RI value on the boiling temperature or any additive molecular parameter (molecular weight, molar refraction, number of carbon atoms in the molecule, etc.) is nonlinear. If this dependence is taken into account, the RI value of any organic compound can be predicted with a precision comparable to the standard deviation of the statistically processed values determined in nonequivalent conditions.

Modern methods of identication of organic compounds based on the results of their GC-MS analysis rely most frequently on the sole intepretation of the mass spectra. Chromatographic retention parameters are used relatively seldom (cf., e.g. refs[1-5]) although they simplify the identification essentially, increase its reliability and make the results less ambiguous. The main reason for the limited use of chromatographic data is the absence of wealthy data banks of retention indices. While existing collections of mass spectra encompass more than 120 thousand substances[6], the most wealthy reference books on retention indices (RI)[7-9] contain no more than two thousand compounds. In these circumstances, the lack of experimental RI data can be reasonable well alleviated by calculating the values based on the readily available and well-systemized banks of physico-chemical constants of organic substances.

The RI values calculated from retention times of the analytes and reference n-alkanes according to Eq. (*1*) should not be considered as a purely analytical characteristics but rather as one of the parameters of the organic compounds, which provides more information than the conventional physico-chemical parameters.

---

* Presented at the 46th Congress of the Czechoslovak Chemical Society, Pardubice, July 4, 1990.

$$I_x = I_n + (I_{n+k} - I_n) \frac{f(t_{R,x}) - f(t_{R,n})}{f(t_{R,n+k}) - f(t_{R,n})}, \qquad (1)$$

where $t_{R,n}$, $t_{R,n}$ and $t_{R,n+k}$ are retention times of the analyte and of reference n-alkanes containing n and $n + k$ carbon atoms in their molecules, respectively, and RI $I_n = 100n$ and $I_{n+k} = 100(n + k)$. The values are obtained in different analytical conditions, with different precisions, and using different function $f(t_R)$. $f(t_R) = \log (t_R - t_m)$ are Kovats indices (isothermal analysis conditions); $f(t_R) = t_R$ are linear retention indices (conditions of linear temperature programming); and $f(t_R) = t_R + q \log . (t_R - t_m)$ are linear-logarithmic retention indices (generalized) (no limitations are imposed on the temperature conditions of the analysis)[10,11]; t_m is the retention time of the unsorbed gas.

Processing of Published RI *Data*

Polydimethylsiloxane elastomers from various suppliers (for the determination of I_x, most frequently in the 500–3 500 range) and polyethylene glycols with molecular weights of 4 000 to 40 000 (I_x 500 to 2 500) are currently used as the standard stationary phases for the determination of retention indices in the identification of organic compounds. No standard conditions, however, have been universally adopted as regards the chromatographic columns and temperature conditions. As a consequence, the published indices for a given substance differ considerably; this must be taken into account when using data from the literature for identification purposes. The principal causes of these differences are the temperature dependence of the RI value ($\beta = dI/dT \neq 0$), sorption effects on the interfaces[12], as well as the effect of the reference compound-to-analyte amount ratio[13], which are difficult to make allowance for. In these circumstances, it is more appropriate to characterize organic compounds by their average RI values calculated from data from various sources, disregarding the differences in temperature and other experimental conditions of their determination, than to use single RI values determined in specific, strictly predetermined conditions. This approach has been applied in ref.[14] to the collection of RI data of toxic compounds and their metabolites. Table I demonstrates this processing of published data as well as of data obtained by the present authors for styrene under different conditions. In this case the average value is $\langle I \rangle = 879$ and its standard deviation is $s_I = 5$. This statistical value can change slightly if new data are included in the collection. The average standard deviation of the average RI values emerging from the processing of RI data for more than 5 thousand substances on the most universal polydimethylsiloxane stationary phases is approximately 8, the majority of the s_I values lying within the range of 5 to 15. The best reproducibility is observed for compounds with the lowest coefficients of temperature dependence of RI, which are those having no rings or active hydrogen atoms (alkanes, alkenes, esters, ethers, etc.); their β values lie between $-0{\cdot}1$ and $+0{\cdot}1$. All classes of cyclic compounds, with higher β coefficients (arenes, cycloalkanes, and their derivatives with $\beta > 0{\cdot}2$) exhibit considerably higher s_I values; for instance, $\langle I \rangle = 657 \pm 9$ for benzene, $1\,167 \pm 14$ for naphthalene, $1\,752 \pm 31$ for anthracene, etc. The low reproducibility of the RI value for more polar compounds containing active hydrogen atoms and thus giving asymmetric peaks in the chromatograms is mainly due to the experimental error of their determination (e.g. $\langle I \rangle = 658 \pm 11$ for butanol, 754 ± 30 for ethylene glycol, 965 ± 13 for phenol).

The methods of calculation of RI data are based on collections of statistically processed data, and thus the most objective criterion for their precision is the agreement between the calculated and experimental values; it should be such that $|I_x - \langle I \rangle| < 2s_I$, which corresponds to the confidence level $\alpha = 0{\cdot}95$.

General Principles of Calculation of RI *Values from Physico-Chemical Parameters*

The existing procedures for the calculation of RI from physico-chemical constants suffer from substantial limitations. The equations used are most frequently linear and they only differ in the number and nature of the independent variables x_i:

TABLE I

Results of statistical processing of the retention indices of styrene on polydimethylsiloxane stationary phases

I_x	Determination temperature, °C[a]	Index type: $f(t_R)$ in Eq. (*1*)	Source of information
870	$T_0 = 40, r = 2$	$t_R + q \log(t_R - t_m)$	authors' data
875	65	$\log(t_R - t_m)$	ref.[15]
878	$T_0 = 50, r = 3$	$t_R + q \log(t_R - t_m)$	authors' data
878	100	$t_R + q \log(t_R - t_m)$	authors' data
879	80	$\log(t_R - t_m)$	ref.[7]
881	$T_0 = 70, r = 3$	$t_R + q \log(t_R - t_m)$	authors' data
884	100	$\log(t_R - t_m)$	ref.[16]
885	100	$\log(t_R - t_m)$	ref.[17]
900[b]	[c]	t_R	ref.[14]

$\langle I \rangle = 879 \pm 5$ (number of measurements $N = 8$)

[a] T_0 is the starting temperature (°C) and r is the temperature increase rate (K min^{-1}) in the temperature programming mode; [b] this value was eliminated based on statistical processing with respect to the 3σ criterion ($900 > 879 + 3{\cdot}5$); [c] data are absent.

$$I_x = \sum a_i x_i + b\,. \qquad (2)$$

No theoretically grounded recommendations exist for the choice of the parameters x_i. They certainly must include those variables that are associated with the distribution coefficients and with the contribution of the induction and dispersion interactions in the "sorbate–stationary phase" system, viz. the boiling temperature T_b (refs[18,19]) or its reciprocal value[20], and molar refraction R_m (refs[18,19]). When the debatable and ambiguous concept of increasing the precision of calculation by increasing the the number of independent variables in Eq. (2) is pursued, the sets of x_i are extended greatly by including the molecular weight M_r (refs[19–22]), molar volume V_m (refs[19,22]), vapour pressure functions[23], dipole moments[24,25] and occasionally even more specific parameters, e.g. the basicity constants for pyridines[25]. Cases exist where arbitrarily chosen combined parameters such as $M_r T_b$ (ref.[21]) or R_m/V_m (ref.[22]) were used, only documenting the empirical approach to the formulation of the correlation equation of type (2). Structural and topological characteristics of organic substances are also used in the calculations using equations of this type; these include, e.g., the total number of carbon atoms in a molecule[19], the number of methyl groups[26,27], Wiener's number and molecular indices of connectivity[28,29], and some quantum chemical data[23] (cf. reviews[30,31]). In such cases, combination of physico-chemical, structural and topological parameters of organic compounds in a single equation lacks physical meaning. As a result, in spite of the great variation in the parameters used, the relationships obtained are only valid for limited sets of substances, usually isomers or homologues over narrow molecular weight or boiling temperature ranges. This is responsible for the low practical value of these methods of calculation. The precision of the results is very low: even for alkanes, one of the simplest classes of substances, the errors attain as much as 31 index units[30].

In the present paper, the following approaches are proposed to the calculation of retention indices of organic substances on the weakly polar dimethylsiloxane stationary phases:

1) A new principle of evaluation of the RI value of any organic compound, which is not associated with the application of a predetermined correlation equation but instead, involves a preliminary check of a series of logical criteria for comparing chromatographic equivalents of the M_r, T_b and R_m values.

2) A new form of the simplest three-parameter correlation equation for the calculation of the RI values of homologues over well-investigated homologous series.

Logical Criteria for the Evaluation of the RI *Values*

It is well-known that the RI value of any organic compound on a weakly polar stationary phase depends primarily on its boiling temperature. The character of the functional dependence $I(T_b)$, however, is not clear, and its approximation by a linear function of type (2) is very crude. In the n-alkane series this dependence is rather logarithmic. Hence, in the evaluation of RI it is more appropriate to use the chromatographic equivalent of the boiling temperature which was proposed as early as 1971 and named "standard retention index"[32], viz.

$$I_T = 100n + 100\frac{\log T_{b,x} - \log T_{b,n}}{\log T_{b,n+1} - \log T_{b,n}}, \qquad (3)$$

where $T_{b,x}$, $T_{b,n}$ and $T_{b,n+1}$ are boiling temperatures of analyte and of reference n-alkanes containing n and $n + 1$ carbon atoms, respectively.

The most important feature of the I_T quantity, which has not been studied in sufficient detail, is the fact that for some substances, isoalkanes for instance[32], it is nearly equal to the experimental RI value on a polydimethylsiloxane stationary phase. In other cases it is recommended that the indices be calculated by using the simplest relation of the type $I_x = I_T + \Delta I$ where ΔI is a tabulated increment for the given homologous series. Actually, the approximate equality of I_T to I_x is met not only for some hydrocarbons but also for some sole compounds from other classes, e.g. C_3H_7Cl ($I_x = 532 \pm 4$; $I_T = 533$), C_6H_5F ($I_x = 665 \pm 8$; $I_T = 665$), $C_3H_7CO_2CH_3$ ($I_x = 708 \pm 6$; $I_T = 715$), $(C_2H_5)_3N$ ($I_x = 676 \pm 3$; $I_T = 671$), $CH_2{=}CHSi(CH_3).Cl_2$ ($I_x = 680$; $I_T = 676$), etc. For all such compounds on the above stationary phases the RI values are virtually independent of any specific "sorbate–stationary phase" interaction and are determined by their boiling temperatures solely. In these cases the problem of the calculation of RI, for some of the substances at least, can be treated from another point of view, viz. as one of the identification of those compounds for which the condition $I_x \approx I_T$ is met. For tackling this problem, it is convenient to compare the I_T values with the chromatographic equivalents of other characteristics of the substances, introduced in an analogous manner. These include particularly the molecular weight and molar refraction, which reflect dispersion and induction interactions with the sorbate, important on nonpolar stationary phases.

The concept of the chromatographic equivalent of molecular weights coincides with that of dispersion retention indices introduced by Evans and coworkers[33] in 1986, viz.

$$I_M = 100n + 100\frac{M_{r,x} - M_{r,n}}{M_{r,n+1} - M_{r,n}} = 7{\cdot}14\,(M_{r,x} - 2)\,. \qquad (4)$$

The I_M values can also be employed for the evaluation of RI for compounds from different classes according to the scheme $I_x = I_M + \Delta I$ by using tabulated corrections ΔI (ref.[33]). The chromatographic equivalent of molar refraction is introduced for the first time as

$$I_R = 100n + 100\frac{R_{m,x} - R_{m,n}}{R_{m,n+1} - R_{m,n}} = 21{\cdot}645\,(R_{m,x} - 2{\cdot}2)\,, \qquad (5)$$

where $R_m = V_m(n^2 - 1)/(n^2 + 2)$, n and V_m are the refractive index and molar volume of the substance, respectively, 2·2 is the double hydrogen refraction, and $R_{m,n+1} - R_{m,n} = 4{\cdot}62$ is the refraction increment of the CH_2 homologous difference. Both experimental data at the same temperature and the known additive schemes can be used for the calculation of the R_m and I_R values.

Thus, according to the above algorithm, three parameters, viz. I_M, I_R and I_T, are sufficient for the identification of cases, where $I_X \approx I_T$ and thus for the evaluation of gas chromatographic RI values without additional calculations, and they can be attributed to any organic compound whose M_r, T_b, n_D^T and ϱ_4^T values are known. Analysis of the collection of RI data for approximately 200 substances of more than 100 homologous series revealed that in the evaluation of RI, the interrelation between I_M, I_R and I_T should be taken into account as follows.

A. The inequality $I_M > I_T$, I_R holds true, i.e from among the three chromatographic equivalents considered, I_M is the highest. (This applies to approximately 12% substances included in the collection). In this case the RI value can be directly equated with the I_T value, the average error being ±12 index units.

B. Among the three parameters considered, any pair exhibits approximate equality, i.e. $I_M \approx I_T$, $I_R \approx I_T$, $I_M \approx I_R$; in the limiting case, $I_M \approx I_R \approx I_T$ (about 23% substances in the collection). In this treatment, two indices are regarded as approximately equal if their difference does not exceed the standard deviation of the statistically processed experimental data, i.e. 8 – 10 i.u. In this case, also, the RI value is best approximated by I_T, the average error being ±11 i.u.

Groups *A* and *B* involve a large fraction of all types of hydrocarbons (unsaturated in particular), nearly all their halo derivatives, silanes, derivatives with trivalent nitrogen(amines, diamines and hydrazines), esters, ethers, mercaptans, some sulfides, etc. It should be emphasized that the problem of the feasibility of evaluating indices without extra calculations based on the relationship $I_x \approx I_T$ is addressed separately in each individual case, no additional information being required.

C. The inequality $I_R < I_M < I_T$ is most frequently encountered (65% of all substances in the collection). In this case the experimental RI values usually lie within very wide ranges of the inequality $I_R < I_x < I_T$. However, even in this case it is suitable to single out a group of substances for which the relation $I_x \approx I_T$ is valid. For this purpose it is necessary to estimate the dispersion of the three I_i parameters in terms of the corresponding standard deviation,

$$s_I = [\sum(\langle I \rangle - I_i)^2/2]^{1/2}\,, \qquad (6)$$

where $\langle I \rangle = (I_M + I_R + I_T)/3$. If $s_I < 100$ (19% substances of the collection), then, as in the cases *A* and *B*, the best approximation for RI is I_T, although in this case the average error increases to ±19 (*C1* variant of evaluation). Attempts to improve the results by replacing the proposed rule by a three-parameter correlation equation of the type

$$I_x = aI_T + bI_R + c \tag{7}$$

with the coefficients $a = 0{\cdot}882$, $b = 0{\cdot}122$ and $c = 14{\cdot}3$ determined by the least squares method lead to increase in precision which is of no practical significance (average error ±18).

This variant of evaluation of RI applies to aliphatic ketones, unsaturated aldehydes, furans, oxiranes, phenol ethers, sulfides, thiophenes, alkyl hydrazones, etc.

If $s_I > 100$ (33% of the collection), the relation $I_x \approx I_T$ no longer holds true. In this case, virtually identical results can be obtained by using a correlation equation of type (*7*) with coefficients $a = 0{\cdot}542$, $b = 0{\cdot}538$ and $c = 48{\cdot}7$, or a similar equation involving the s_I value, viz.

$$I_x = a'I_T + b's_I + c' \tag{8}$$

with $a' = 1{\cdot}085$, $b' = -0{\cdot}979$ and $c' = 33{\cdot}9$. In this case the average errors of the indices attain ±20 and ±18 i.u. for Eqs (*7*) and (*8*), respectively (*C2* evaluation variant).

The most polar compounds containing –OH or –COOH functional groups with active hydrogen atoms are characterized by a low reproducibility of experimental RI data, and thus their prediction will involve the highest errors. It is desirable to include them in a separate group of substances which invariably belong to type *C* with respect to the interrelation of the parameters I_M, I_R and I_T and are characterized by values of $s_I > 100$ (13% of the collection). Their indices can only be evaluated by means of correlation equations (*7*) and (*8*) with different sets of coefficients, viz. $a = 0{\cdot}708$, $b = 0{\cdot}472$ and $c = -88{\cdot}0$, and $a' = 1{\cdot}185$, $b' = -0{\cdot}881$, $c' = -99{\cdot}1$ (*C3* evaluation variant); the errors are 20–30 i.u.

The comparison of the I_M, I_R and I_T practically implies verification of a series of logical criteria for their comparison, which in many cases eliminates the necessity for further calculation. For approximately one-half (54%) of organic compounds from the collection investigated, the I_T values provide the best approximation of the indices. In other cases the experimental RI data are best evaluated from the parameters I_T and I_R contained in correlation equations with similar coefficients ($a \approx b \approx 0{\cdot}54$ in *C2* variant and $a \approx 0{\cdot}71$, $b \approx 0{\cdot}47$ in *C3* variant). These values characterize compounds whose chromatographic retention is determined not only by their boiling temperatures but also, to a considerable extent, by induction interactions with the stationary phase (only polar compounds belong to type *C*). Analysis of coefficients in Eq. (7) for the *C1* variant, where the validity of $I_x \approx I_T$ can be assumed, shows that for these substances the effect of the second factor is much lower ($a \approx 0{\cdot}88 > b \approx 0{\cdot}12$). It is noteworthy that in this method, the dispersion retention indices I_M (ref.[33]) are not directly used in the calculations; they only occur in the stage of logical operations or are indirectly involved in the values of s_I in calculations from equations of type (*8*). This fact can be indicative of a low contribution of the corresponding mechanism of "sorbate–stationary phase" interactions.

The precision of the evaluation of RI with the assistance of the logical criteria proposed is relatively close to the maximum attainable precision because the standard deviations of the statistically processed experimental data are on the same order (± 5 to ± 15), and are much higher for substances having active hydrogen atoms. Moreover, for many organic compounds the boiling temperatures reported by different authors differ by $1 - 3$ K and more. This is manifested by additional errors of the RI values being no less than $5 - 15$ i.u.

Table II gives a comparison of RI values calculated by the proposed method (I_x) and the average experimental RI values for some organic substances on polydimethylsiloxane stationary phases.

Calculation of RI *Values within Homologous Series*

The establishment of the main source in equations of type (*2*), i.e. evaluation of the dependence of RI values of organic compounds on their boiling temperatures in a linear form, makes it possible to consider the feasibility of calculating the indices of unknown representatives of well-investigated homologous series from a new point of view. Verification of a set of the simplest correlation equations showed that a three-parameter equation of the form

$$\log I_x = a \log T_b + bA + c \qquad (9)$$

is the most suitable for the calculation of RI from the data for other homologues over relatively wide ranges of molecular weights and boiling temperatures. In this equation, T_b is the boiling temperature (in K) of the compound under investigation and A is any additive property of the homologues such as the molecular weight M_r, molar refraction R_m, total number of carbon atoms in the molecule n_C, number of carbon atoms in alkyl substituents, etc. The coefficients in Eq. (*9*) are calculated from data for a set of homologues of the given series by the least squares method.

Any of the above additive parameters A gives virtually the same results, thus indicating that all equations of type (*9*) can be regarded as equivalent. The use of the M_r and n_C values is most convenient, molar refraction calculations requiring additional data. Verification of the applicability of Eq. (*9*) to the calculation of RI on the example of several homologous series (Table III) gave evidence that this

TABLE II

Comparison of calculated (I_x) and statistically processed ($\langle I \rangle$) retention indices of some compounds

Compound	M_r	T_b	ϱ g ml^{-1}	n_D^{20}	I_M	I_R	I_T	Method of calculation	I_x	$\langle I \rangle \pm s_I$
tert-Butyl mercaptan	90	64·2	0·800	1·422	628·6	571·2	586·8	*A*	587	589 ± 20
Dichloromethane	85	40·1	1·325	1·424	592·9	306·7	512·8	*A*	512	527 ± 12
Ethyl butyrate	116	121·6	0·879	1·395	814·3	637·2	788·6	*A*	789	780 ± 11
Phenyl acetylene	102	142·4	0·930	1·549	714·3	707·6	868·5	*B*	868	862 ± 12
Triethylsilane	116	107·7	0·730	1·412	814·3	808·2	736·1	*B*	736	744
1,1-Diisopropylhydrazine	116	133·8	0·817	1·431	814·3	747·8	835·6	*B*	836	835
Thiophene	84	84·1	1·064	1·529	585·7	479·4	652·9	*C1*	653; 648[a]	653 ± 10
2-Nonanone	142	195·3	0·818	1·421	1 000·0	905·2	1 097·3	*C1*	1 097; 1 092	1 080 ± 12
Benzonitrile	103	191·1	1·005	1·528	721·4	635·5	1 078·5	*C2*	974; 975	958 ± 6
Ethyl 4-methylbenzoate	164	238	1·025	1·508	1 157·1	984·6	1 313·6	*C2*	1 291; 1 299	1 288
Dimethylacetamide	87	165·5	0·937	1·436	607·1	477·9	964·0	*C2*	828; 834	840 ± 4
Tetrahydrofurfuryl alcohol	102	177	1·050	1·452	714·3	519·7	1 014·0	*C3*	883; 875	878 ± 8

[a] Values calculated by different procedures (see text).

TABLE III

Parameters of the equation $\log I_x = a \log T_b + bA + c$ for some homologous series (polydimethylsiloxane stationary phases)

Homologous series	n_C	A	a	b	c	Number of compounds	$\langle\lvert\Delta I\rvert\rangle$
Alkanes	4—10	n_C	2·0094	−5·206 . 10^{-3}	−2·2800	73	3
Cycloalkanes	3—11	n_C	1·8587	−2·7446 . 10^{-3}	−1·9006	46	10
Alkenes	4—12	n_C	1·8891	−1·5906 . 10^{-3}	−1·9933	53	4
Arenes	6—13	M_r	1·7952	4·5606 . 10^{-5}	−1·7622	50	5
Ethers	4—8	n_C	1·8627	7·4906 . 10^{-4}	−1·9402	8	6
Esters	2—10	M_r	2·2645	−4·3732 . 10^{-5}	−2·9847	50	10
Ketones	3—9	n_C	1·9290	7·3749 . 10^{-3}	−2·1696	19	10
Alkyl chlorides	2—8	n_C	2·0292	−4·060 . 10^{-3}	−2·3453	15	4
Alkyl bromides	1—5	n_C	1·8515	−2·3901 . 10^{-3}	−1·9012	12	4
Dichloroalkanes	1—5	n_C	1·6235	4·075 . 10^{-3}	−1·3437	12	7
Alkyl pyridines	5—8	M_r	1·6540	8·3101 . 10^{-4}	−1·4842	15	4

TABLE IV

Results of the calculation of retention indices of some alkanes from the equation $\log I_x = a \log T_b + bn_C + c$ ($a = 2{\cdot}0094$, $b = -0{\cdot}005206$, $c = -2{\cdot}2800$)

Alkane	T_b, °C	n_C	$\langle I\rangle \pm s_I$	I_x	$\lvert\Delta I\rvert$
iso-C_4	−11·8	4	365 ± 5	360	5
2,2-diMeC_3	9·5	5	412 ± 2	416	4
3-MeC_5	63·3	6	584 ± 2	584	0
2,3-diMeC_4	58·0	6	568 ± 3	565	3
n-C_7	98·4	7	700	704	4
2,4-diMeC_5	80·5	7	630 ± 2	638	8
3,3-diMeC_5	86·1	7	659 ± 3	658	1
3-EtC_5	93·5	7	687 ± 2	686	1
3,3-diMeC_6	112·0	8	744 ± 4	748	4
2,2,3-triMeC_5	109·8	8	738 ± 4	739	1
2,2,4-triMeC_5	99·2	8	691 ± 2	699	8
2-Me,3-EtC_5	115·6	8	762 ± 2	762	0
2,2,3,3-tetraMeC_4	106·5	8	725 ± 12	726	1
3-MeC_8	144·2	9	871 ± 1	868	3
2,6-diMeC_7	135·2	9	827 ± 1	831	4
2,2,3,4-tetraMeC_5	133·0	9	823 ± 4	822	1
2,3,3,4-tetraMeC_5	141·6	9	862 ± 5	857	5
3,3-diEtC_5	146·2	9	881 ± 5	877	4
					$\langle\lvert\Delta I\rvert\rangle = 3$

equation is obeyed irrespective of the chemical nature of the substances concerned; the average errors vary from 3 to 10 i.u., which is less than the average standard deviations of the starting data. In this case the coefficients a vary from 1·6 to 2·3, which bears out the basically nonlinear character of the dependence $I(T_b)$ and the inadmissibility of using linear equations of type (2).

Tables IV and V give specific examples of the use of Eq. (9) with different sets of coefficients for the calculation of RI of isoalkanes and alkylpyridines. The analysis of the maximum deviations $|\Delta I|$ indicated that they are not related to the structure of the substance and are most probably fortutious and caused by errors in the determination of either RI or the boiling temperatures. However, for the majority of series, the calculated RI values of higher homologues may be markedly underestimated as compared with the experimental data. This effect is certainly related with the starting data processing principle. When the indices are averaged, the temperature ranges within which the experimental data lie are displaced towards higher values for higher homologues. Moreover, the average $\langle I \rangle$ values themselves

TABLE V

Results of the calculation of retention indices of some alkyl pyridines from the equation $\log I_x = a \log T_b + b M_r + c$ ($a = 1{\cdot}6540$, $b = 8{\cdot}3101 \cdot 10^{-4}$, $c = -1{\cdot}4842$)

Pyridine	T_b, °C	M_r	$\langle I \rangle \pm s_I$	I_x	$\|\Delta I\|$
2-Me	129·6	93	799 ± 13	794	5
3-Me	144·0	93	849 ± 11	842	7
4-Me	145·4	93	847 ± 12	846	1
2,3-diMe	161·2	107	925 ± 17	924	1
2,4-diMe	158·4	107	919 ± 11	914	5
2,5-diMe	157·0	107	913 ± 11	909	4
2,6-diMe	144·0	107	873 ± 7	864	9
3,4-diMe	179·1	107	995 ± 6	987	8
3,5-diMe	172·5	107	970 ± 7	964	6
2-Et	148·6	107	878 ± 12	880	2
3-Et	166·0	107	938 ± 10	941	3
4-Et	167·7	107	951 ± 9	947	4
4-isoPr	181·5	121	1 026[a]	1 022	4
4-Pr	173·0	121	1 047 ± 17	1 050	3
2,3,6-triMe	177·0	121	1 002 ± 10	1 006	4
2,4,6-triMe	170·4	121	981 ± 7	981	0
				$\langle \|\Delta I\| \rangle =$	4

[a] The single value of RI of this compound from literature data (without statistical processing).

pertain to different temperature conditions and can be most effectively used in temperature programming conditions. Evaluation of the temperature dependence of RI is not included in any of the existing methods of their calculation, including the method proposed in this paper. Therefore, underestimated results should be expected to emerge from extrapolation of the indices of higher homologues from data of the simplest members of the series. In order to eliminate these errors, the coefficients of Eq. (9) should preferably be calculated from the data of homologues differing by no more than 6 – 8 carbon atoms.

The advantages of the proposed simple method for the evaluation and calculation of the RI value of nearly any organic compound which is characterized by its boiling temperature at atmospheric pressure are evident: they are based on readily available physico-chemical constants and require no complex calculations, the precision of the results comparing well with that provided by other, considerably more laborious methods[30,31].

The authors wish to express their gratitude to the Organizing Committee for having been given the opportunity to present the results of this work at the 46th Congress of the Czechoslovak Chemical Society.

REFERENCES

1. Jellum E., Helland P., Eldjarn L., Markwardt U., Warhöfer J.: J. Chromatogr. *112*, 573 (1975).
2. Smith D. H., Achendel M., Yeager W. J., Andersson P. J., Fitch W. L., Rindfleidch T. C.: Anal. Chem. *49*, 1623 (1977).
3. Tsuchiya Y., Boulanger J., Sumi K.: Chromatographia *10*, 154 (1977).
4. Gates S. C., Smisko M. J., Ashendel C. L., Young N. D., Holland J. F., Sweeley C. C.: Anal. Chem. *50*, 433 (1978).
5. Rasmussen P.: Anal. Chem. *55*, 1331 (1983).
6. *The Eight Peaks Index of Mass Spectra*, 3rd ed. The Royal Soc. of Chem., Nottingham 1983.
7. Bogoslovskii Yu. N., Anvaer B. I., Vigdergauz M. S.: *Khromatograficheskie postoyannye v gazovoi khromatografii*. Izdatelstvo Standartov, Moskva 1978.
8. Sprouse J., Varano A.: International Laboratory No. 6, 54 (1984).
9. Jennings W., Shibamoto T.: *Qualitative Analysis of Flavor and Fragrance Volatiles by Glass Capillary Gas Chromatography*, p. 472. Academic Press, New York 1980.
10. Zenkevich I. G.: Zh. Anal. Khim. *39*, 1297 (1984).
11. Zenkevich I. G., Ioffe B. V.: J. Chromatogr. *439*, 185 (1988).
12. Berezkin V. G.: Zh. Anal. Khim. *40*, 555 (1985).
13. Zenkevich I. G., Zibulskaya I. A.: Zh. Anal. Khim. *44*, 90 (1989).
14. *GC Retention Indices of Toxicologically Relevant Substances on SE-30 or OV-1.* Report II of the DFG Commission for Clinical-Toxicological Analysis. Special Issue of the TIAFT Bulletin. Verlag Chemie, Weinheim 1985.
15. Svob V., Deur-Šiftar D., Cramers C. A.: J. Chromatogr. *91*, 659 (1979).
16. Boneva S., Papazova D., Dimov N.: Godishnic Vissh. Khim.-Tekhnol. Inst. (Burgas) *18*, 143 (1983).
17. Engevald W., Topalova I., Petsev N., Dimitrov C.: Chromatographia *23*, 561 (1987).

18. Soják L., Krupčík J., Rijks J.: Chromatographia *7*, 26 (1974).
19. Dimov N., Milina R.: J. Chromatogr. *463*, 159 (1989).
20. Streuli C. A., Orloff M.: J. Chromatogr. *101*, 23 (1974).
21. Dimov N., Boneva S.: Zh. Anal. Khim. *43*, 96 (1988).
22. Chelghoum C., Haddad Y., Abdeddaim K., Guermonche M. H.: Analusis *14*, 49 (1986).
23. Dimov N., Mekenyan O.: J. Chromatogr. *471*, 227 (1989).
24. Papp O., Jozan M., Szasz G., Hermecz I.: J. Chromatogr. *403*, 11 (1987).
25. Premecz J. E., Ford M. E.: J. Chromatogr. *388*, 23 (1987).
26. Papazova D., Dimov N.: J. Chromatogr. *137*, 259 (1977).
27. Dimov N., Papazova D.: Chromatographia *12*, 443 (1979).
28. Nabivach V. M., Vasiliev V. E.: Izv. Vuzov. Khim. Khim. Tekhnol. (U.S.S.R.) *30*, (10) 72 (1987).
29. Khorasheh F., Gray M. R., Selucky M. L.: J. Chromatogr. *481*, 1 (1989).
30. Dimov N. P.: J. Chromatogr. *360*, 25 (1986).
31. Ecknig W.: Z. Chem. *25*, 240 (1985).
32. Robinson P. G., Odell A. L.: J. Chromatogr. *57*, 1 (1971).
33. Evans M. B., Haken J. K., Toth T.: J. Chromatogr. *351*, 155 (1986).

Translation revised by P. Adámek.