Article

Not peer-reviewed version

Application of Gas Chromatographic Retention Indices to GC and GC–MS Identification with Variable Limits for Deviations Between Their Experimental and Reference Values

Igor G. Zenkevich *

Posted Date: 21 October 2025

doi: 10.20944/preprints202510.1599.v1

Keywords: gas chromatography; retention indices; reference values; correction of experimental data; identification; databases

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2025 doi:10.20944/preprints202510.1599.v1

Article

Application of Gas Chromatographic Retention Indices to GC and GC–MS Identification with Variable Limits for Deviations Between Their Experimental and Reference Values

Igor G. Zenkevich

Institute for Chemistry, St. Petersburg State University, 198504 St. Petersburg, Russia; izenkevich@yandex.ru; Tel.: +7-(812)-428-4045

Abstract

The potential of the new algorithm for comparing experimental and reference values of gas chromatographic retention indices (RI) is discussed. This algorithm is designed to eliminate significant elements of uncertainty typical of numerous contemporary recommendations, primarily the fixed limiting values of permissible deviations, $\Delta RI = (RI_{ref} - RI_{exp})$. The algorithm proposed implies the calculation of deviations ΔRI for selected most reliably identified constituents of multicomponent mixtures with known reference RI values, followed by calculation of coefficients of regression equations $\Delta RI = (RI_{ref} - RI_{exp}) = aRI_{exp} + b$ for total sets of analytes. These equations allow recalculating the experimentally determined RIs into the corrected values $RI_{corr} = RI_{exp} + \Delta RI$. Such algorithm makes it possible to use reference RI values for semi-standard nonpolar polydimethylsiloxane phases (with 5% phenyl groups and others) for the comparison with data determined with standard nonpolar polydimethylsiloxanes and *vice versa*. It is applicable both to statistically processed reference data and to results of single measurements.

Keywords: gas chromatography; retention indices; reference values; correction of experimental data; identification; databases

---

1. Introduction

At present, gas chromatography–mass spectrometry (GC-MS) appeared to be the most effective and widely applicable hyphenated technique for analysis of organic compounds in complex mixtures. This is due to the availability not only of modern equipment, but also of detailed and well-systemized informational support. It includes data bases of standard mass spectra (electron ionization, 70 eV) and gas chromatographic retention indices (RI [1]) on standard nonpolar (polydimethylsiloxanes) and polar (polyethylene glycols) stationary phases. The example of combination of these parameters is NIST mass spectral database [2], which has been supplemented with RIs since 2005. The last version of this database (2023) contains mass spectra of 347,000 compounds and RIs of 153,000 compounds.

The database application efficiency is determined not by the number of objects included, but by the algorithms for comparing experimental and reference (library) data, both mass spectra and RIs. Many of previously proposed mass spectrometric algorithms (some of them were considered in the monograph [3]) are currently of only historical interest because the most widely used is the algorithm proposed by employees of the Finnigan Co. in 1978 [4] and its subsequent modifications. Its essence is as follows: Each mass spectrum under comparison can be represented as a vector in an N-dimensional space, where N is the number of signals (in other words, the maximal m/z value). The numerical expression of mutual correspondence of these vectors (other terms are Similarity, Match Factor (MF), etc.) is the square of cosine of the angle (Θ) between them, $0 \le \cos^2\Theta \le 1$. For greater clarity, the normalization condition often transforms into $0 \le \mathrm{MF} \le 100$, or $0 \le \mathrm{MF} \le 1000$ (the maximal MF value corresponds to the maximal similarity of mass spectra, and the minimal MF value, to their complete dissimilarity). The normalization $0 \le \mathrm{MF} \le 100$ is sometimes called a percentage match, which is incorrect.

In contrast to multidimensional mass spectra, GC retention indices are unidimensional analytical parameters (one number). It would seem that assessing the degree of coincidence of unidimensional values should be much simpler compared to multidimensional values. However, this is not the case, and this problem remains unsolved to date. The experimental RI values cannot exactly match with reference data, because the former values are influenced by experimental errors, and the latter values, by interlaboratory irreproducibility. The reason of restricted interlaboratory reproducibility of RIs is their dependence on the conditions of GC analyses, primarily on temperature, even at fixing the stationary phase. In turn, the temperature depends on the geometry of the chromatographic column and the amount of the stationary phase in it. In addition, the RI values depend on the ratio of peak areas of the analytes and reference compounds [5]. Hence, if we postulate any RI values as the reference information, $\mathrm{RI}_{\mathrm{ref}}$, and compare the experimental $\mathrm{RI}_{\mathrm{exp}}$ values with them, the differences ΔRI between them turn out to be inconstant. Theoretically, we can imagine the existence of limiting values for these differences, $\Delta\mathrm{RI}_{\mathrm{lim}}$. Combining all these premises, we can formulate the simplest condition of GC identification using retention indices:

$$\Delta\mathrm{RI} = |\mathrm{RI}_{\mathrm{exp}} - \mathrm{RI}_{\mathrm{ref}}| \le \Delta\mathrm{RI}_{\mathrm{lim}} \qquad (1)$$

If $\Delta\mathrm{RI} > \Delta\mathrm{RI}_{\mathrm{lim}}$, the identification must be considered impossible even at the “ideal” mutual correspondence of the mass spectra.

The values $\Delta\mathrm{RI}_{\mathrm{lim}}$ depend not only on the nature of the stationary phase and specific conditions of the analysis, but also on the chemical nature of analytes and the features of the database used. In accordance with the long-established practice, the correctness of GC identification is most often illustrated by direct comparison (for visual perception) of the experimental and reference values; see, e.g., publications [6–15], but the number of examples can be increased many times.

Attempts have been made to create combined criteria for joint mass spectrometric and gas chromatographic identification. For example, Smith with coauthors [16] proposed the criterion $F = U \cdot \mathrm{MF}$ as follows:

$$\begin{aligned} U &= 1, \text{ if } \Delta\mathrm{RI} \le 5, \\ &1 - 0.05|\Delta\mathrm{RI}|, \text{ if } 5 < |\Delta\mathrm{RI}| \le 20 = \Delta\mathrm{RI}_{\mathrm{lim}} \\ &0, \text{ if } |\Delta\mathrm{RI}| > 20, \end{aligned} \qquad (2)$$

where MF is the mass-spectrometric match factor and U is the measure of correspondence of GC RIs. In other words, small differences ΔRI (≤ 5 index units) do not affect the results of mass-spectrometric identification, while large differences (> 20 i.u.) mean elimination of this option from further consideration despite the best correspondence of mass spectra. Such condition (cited in publication [17]) seems to be useful as very preliminary estimates, but it has obvious disadvantages, namely:

1. The choice of fixed value $\Delta\mathrm{RI}_{\mathrm{lim}} = 20$ i.u. seems to be illogical because the differences between experimental and reference data can take both larger and smaller values depending on the chemical nature of analytes and types of stationary phases under comparison;
2. Condition (2) implies the symmetrical distribution of ΔRI, in other words, the equally probable deviations of experimental and reference data both at $\Delta\mathrm{RI} < 0$ and $\Delta\mathrm{RI} > 0$, but in the general case it is not obvious.

The differences in the real interlaboratory RI distribution can be best illustrated by histograms for 2,6-dimethyloctane (Figure 1a) and 1,2,3,4-tetrahydronaphthalene (tetralin), (Figure 1b). Both histograms are unsymmetrical, which means the theoretical incorrectness of conventional statistical data processing. Moreover, the histogram for tetralin allows us to identify the two-modality RI distribution, which is quite common, but the explanations for this can be quite complex [17]. However, because other algorithms of data processing are too complicated, we have to calculate arithmetic averages together with their standard deviations, which gives 933 ± 3 for 2,6-dimethyloctane (a) and 1152 ± 15 for tetralin (b). The first compound illustrates the good interlaboratory RI reproducibility, but the second case is more typical of numerous organic compounds.

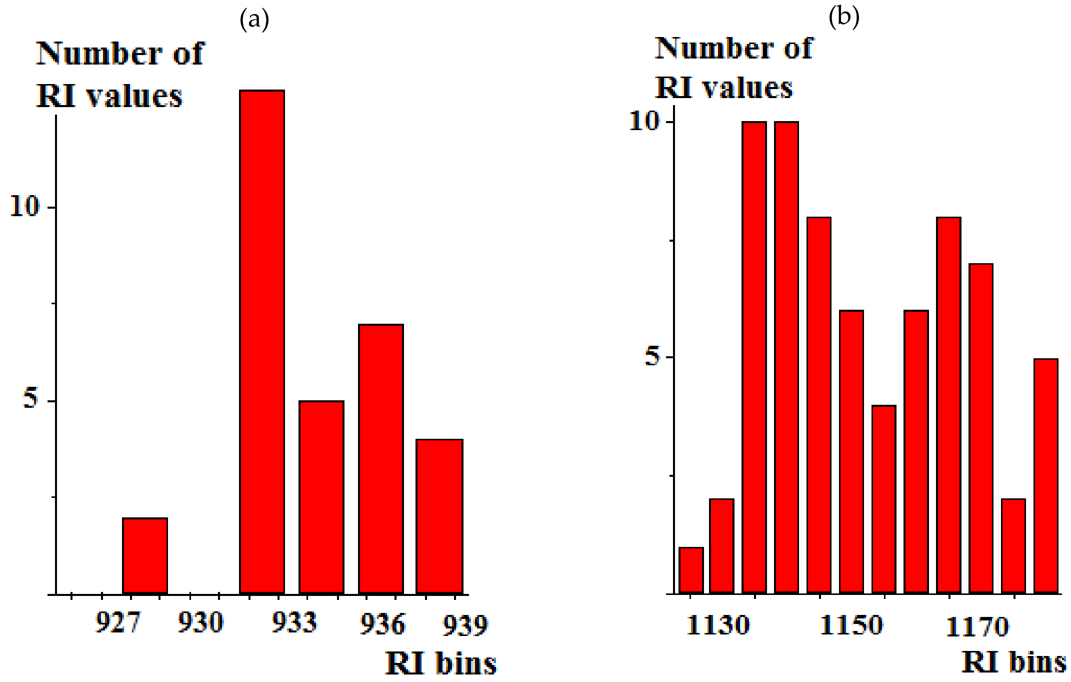

Figure 1. Histograms illustrating the distribution of reference RI values for (a) 2,6-dimethyloctane and (b) 1,2,3,4-tetrahydronaphthalene (tetralin) on standard nonpolar polydimethylsiloxane stationary phases [2]. The bin sizes are (a) 2 i.u. (index units) and (b) 5 i.u. The average RI values are (a) 933 ± 3 and (b) 1152 ± 15.

Such spread of RI values depends on the chemical nature of analytes and determines the difficulties and uncertainties of their application [18,19]. It explains the use of arithmetic averages in combination with the corresponding evaluations of permissible deviations. The author's preferences are commonly accepted standard deviations, <RI> $\pm s_{RI}$ [20]. According to the relationships of statistical processing, the intervals <RI> ± s_{RI} include about 68% of values of the initial data set; the intervals <RI> ± $2s_{RI}$, about 95%; and the intervals <RI> ± $3s_{RI}$, about 99.7%. In addition, the reference data presented without deviations are known, e.g., reference data for constituents of essential oils [21]. In the NIST database [2] and in some secondary data summaries based on [2] (see, e.g., [22]), the RI spread historically is customarily characterized by MAD values (medians of absolute deviations). This means that the intervals <X> ± MAD contain only 50% of values of the initial data set. Furthermore, the number of single RI measurements (measured in one laboratory) in the last versions of the NIST database increased noticeably; such data have no measures of possible deviations at all.

The mentioned features of interlaboratory distributions of RI values explain the difficulties of selecting the criteria for comparing their experimental and reference values. Nevertheless, there is an actual need for creating the corresponding algorithm. The algorithm must have the following properties:

- it should not include any limiting values of possible deviations between experimental and reference RI values;
- it should be applicable to asymmetrically distributed reference RI values;
- it should be applicable to both statistically processed reference data and the results of single measurements;
- it should be applicable to comparing the RI values determined using columns with semi-standard stationary phases (terminology used in database [2]) with reference data for standard nonpolar stationary phases and *vice versa*;

- it should be applicable to RI values measured for both capillary and packed chromatographic columns at various temperatures without any artificial restrictions.

This paper discusses the possibilities of creating an algorithm that meets all the above-listed conditions.

2. Results and Discussion

It is appropriate to start the discussion of an algorithm that meets *a priori* requirements listed above not with a theoretical consideration, but with a specific example. At the same time, this example illustrates the symbolism used.

2.1. Optimization of Comparing the Experimental and Reference Values of GC Retention Indices

To start the discussion, let us select the example of GC retention indices of 20 compounds of the same chemical class (alkylarenes) listed in Table 1, but measured under very specific conditions, namely, packed column with semi-standard phase Apiezon L (15% on Celite C-22) at 100°C [23]. According to contemporary concepts, such conditions are completely outdated because too high content of the stationary phase leads to too high temperature of chromatographic separation.

Table 1. Comparison of the retention indices published in [23] with RI values from NIST database [2] on standard nonpolar polydimethylsiloxane stationary phases.

Compound	RI_{exp} [23]	RI_{ref} [2]	$\Delta_{ref-exp}$	Complete set of reference data		Reduced set of reference data	
				RI_{corr}	$\Delta_{corr-ref}$	RI_{corr}*	$\Delta_{corr-ref}$
Benzene	**679**	**654 ± 7**	**-25**	**650**	**-4**	**650**	**-4**
Toluene	**790**	**757 ± 6**	**-33**	**761**	**+4**	**760**	**+3**
Ethylbenzene	879	850 ± 6	-29	850	0	849	-1
m-Xylene	893	860 ± 6	-35	866	+6	865	+5
p-Xylene	893	860 ± 6	-33	864	+4	863	+3
o-Xylene	919	881 ± 6	-37	890	+9	889	+8
Isopropylbenzene	934	919 ± 7	-15	905	-14	904	-15
Propylbenzene	966	945 ± 5	-21	937	-8	936	-9
1-Methyl-4-ethylbenzene	983	953 ± 5	-30	954	+1	953	0
tert-Butylbenzene	998	986 ± 7	-12	969	-17	968	-18
1-Methyl-2-ethylbenzene	999	969 ± 5	-30	970	+1	969	0
1,3,5-Trimethylbenzene	**1002**	**962 ± 6**	**-40**	**974**	**+12**	**972**	**+10**
sec-Butylbenzene	1019	1000 ± 5	-19	991	-9	989	-11
1,2,4-Trimethylbenzene	1027	983 ± 5	-44	999	+16	997	+14
1,3-Diethylbenzene	1054	1040 ± 5	-14	1026	-14	1023	-17
1,2,3-Trimethylbenzene	1058	1010 ± 6	-48	1030	+20	1027	+17
Butylbenzene	**1068**	**1047 ± 6**	**-21**	**1040**	**-7**	**1037**	**-10**
1-Methyl-2-propylbenzene	1076	1058 ± 5	-18	1048	-10	1045	-13
1,2,4,5-Tetramethylbenzene	**1139**	**1107 ± 5**	**-32**	**1111**	**+4**	**1108**	**+1**
1,2,3,5-Tetramethylbenzene	1152	1110 ± 6	-42	1124	+14	1121	+11
Average standard deviation of reference RI values, s_{RI}		5.8					
Average difference $\Delta_{ref-exp}$:			-29 ± 10				
Average difference $\Delta_{corr-ref}$:					9 ± 6 (0 ± 11)**		8 ± 6* (-1 ± 10)**

Footnotes: The names and numerical data for compounds included in the reduced data set are marked in bold; *) RI_{corr} and $\Delta_{corr-exp}$ values are calculated using the reduced data set; **) average values ± standard deviations considered taking into account the signs of the initial data are indicated in parentheses. The parameters of the equation $\Delta_{ref-exp} = RI_{ref} - RI_{exp} = aRI_{exp} + b$ for the complete data set are $a = 0.002 \pm 0.024$, $b = -30.5 \pm 23.3$, $R = -0.02$, and $S_0 = 10.8$, and for the reduced data set, $a = -0.005 \pm 0.021$, $b = -25.3 \pm 20.6$, $R = -0.14$, and $S_0 = 8.4$.

As a result of using such obsolete analytical conditions and semi-standard phase, all the experimental RIs exceed contemporary reference RI values [2] by approximately 30 i.u. (differences vary from 12 to 48 i.u.). The values of $\Delta_{ref-exp} = RI_{ref} - RI_{exp}$ for all the compounds are presented in Table 1. The comparison of these values requires taking into account not only their random variations, but

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 October 2025 doi:10.20944/preprints202510.1599.v1

also systematic differences caused by nonequivalence of determination conditions. This means that we must take into account the possible dependence of amendments $\Delta_{ref\text{-}exp} = RI_{ref} - RI_{exp}$ on the values of RI_{exp} themselves:

$$\mathbf{\Delta_{ref\text{-}exp} = (RI_{ref} - RI_{exp}) = \mathit{a}RI_{exp} + \mathit{b}} \tag{3}$$

The plot of the dependence (3) (swarm of dots) is shown in Figure 2(a).

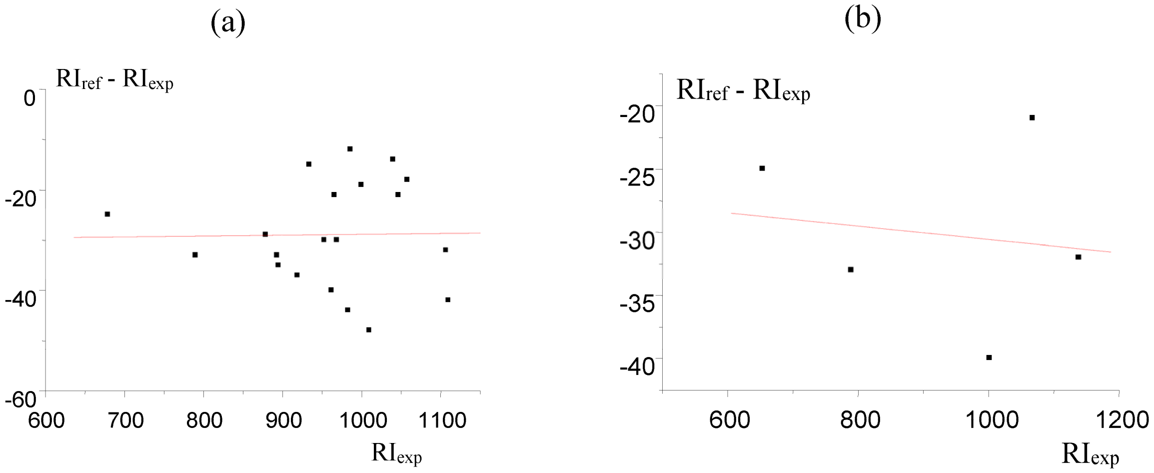

Figure 2. (a) Linear approximation of the differences $\Delta RI_{ref\text{-}exp} = (RI_{ref} - RI_{exp})$ *vs.* RI_{exp} plots for the complete data set from [23]; reference data are taken from NIST database [2]. The parameters of the regression equation are indicated in the footnotes to Table 1; (b) the same data for the reduced data set for five alkylarenes selected from their total list [23].

Thus, we come to the following conclusions. Firstly, the $\Delta_{ref\text{-}exp}$ values show significant scatter. Hence, secondly, there are no reasons to approximate these data by polynomials of higher orders. Thirdly, the standard deviation of coefficient "*a*" of dependence (3) exceeds in absolute value the value of the coefficient itself, which means weak dependence of amendments $\Delta_{ref\text{-}exp}$ on RI_{exp} values. Finally, the determination of the parameters of equation (3) (even at the small value of coefficient "*a*") allows us to convert the experimental data published in [23] to the corrected RI_{corr} values specially for comparing with information from the contemporary database [2]:

$$\mathbf{RI_{corr} = RI_{exp} + \Delta_{ref\text{-}exp}} \tag{4}$$

All the RI_{corr} values are listed in Table 1 also. As we can see, the average absolute values of differences $\Delta_{corr\text{-}ref}$ (9 ± 6) appeared to be three times smaller than the average absolute value of differences $\Delta_{ref\text{-}exp}$ (29 ± 10). In should be noted that the average corrected values ± standard deviations considered taking the signs of the initial data into account differ from zero statistically insignificantly (0 ± 11). Their distribution can be additionally illustrated by the corresponding histogram (Figure 3). It is slightly asymmetric, but this is typical of GC retention indices of many organic compounds:

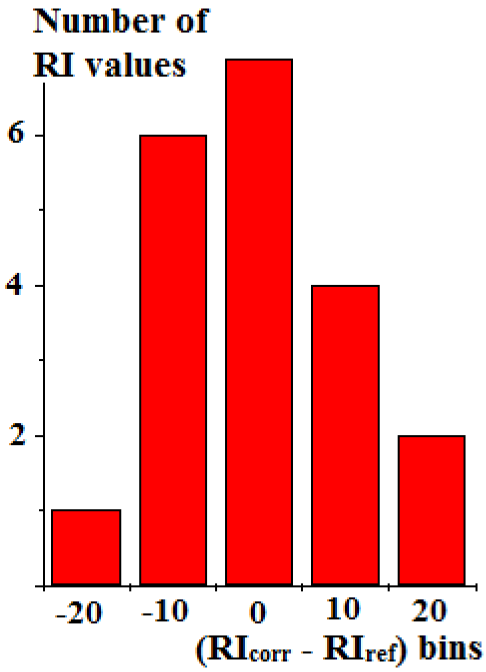

Figure 3. Histogram of the distribution of retention indices of alkylarenes [23] recalculated into $\mathrm{RI_{corr}}$ values. The bin size is 10 i.u. The average RI value of the data considered with their signs is 0 ± 11.

However, the correction of the experimental RI values for providing the possibility of their optimal comparison with the available reference data is only the "half" of the problem. The second part is the evaluation of the permissible deviations of $\mathrm{RI_{corr}}$ values from $\mathrm{RI_{ref}}$ for accepting or rejecting the gas chromatographic identification. In the least-squares method, the formula exists for so-called "corridor of errors" for evaluating the possible deviations of points from the regression equation $y = ax + b$ [24]:

$$\boldsymbol{S(x)} = s_{\mathrm{y}}\sqrt{\frac{(1-R\uparrow 2)}{N-2} \times \left[1 + \frac{(a-<x>)\uparrow 2}{sx\uparrow 2}\right]} \tag{5}$$

where $s_x^2 = (<x^2> - <x>^2/N)/N$, $s_y^2 = (<y^2> - <y>^2/N)/N$, $R = (<xy> - <x><y>/N)/(Ns_xs_y)$

However, this relationship seems too complicated for routine calculations; simpler evaluations are preferable. If the dependence $\Delta\mathrm{RI_{corr}} = f(\mathrm{RI_{exp}})$ is expressed rather weakly (like it is in the example considered), we can accept $\Delta\mathrm{RI_{corr}} \approx$ const. The first recommendation implies using the evaluations of the standard deviations of $\Delta\mathrm{RI_{corr}}$, namely, $s_{\Delta\mathrm{RIcorr}}$. Most of $\Delta\mathrm{RI_{corr}}$ values do not exceed $\Delta\mathrm{RI_{corr}} + s_{\Delta\mathrm{RIcorr}}$, while the values exceeding $\Delta\mathrm{RI_{corr}} + 2s_{\Delta\mathrm{RIcorr}}$ can be excluded from consideration. The second way to evaluate the possible deviations is based on such parameter as "sum of residuals", S_0. For numerous examples (including those considered above), the following inequality is correct:

$$\Delta\mathbf{RI_{corr}} + s_{\Delta\mathrm{RI}} < S_0 < \Delta\mathbf{RI_{corr}} + 2s_{\Delta\mathrm{RI}} \tag{6}$$

This means that the $\mathrm{RI_{corr}}$ data for GC identification should be chosen within intervals $\mathbf{RI_{ref}} \pm 2S_0$:

$$\mathbf{RI_{ref}} - 2S_0 < \mathbf{RI_{corr}} < \mathbf{RI_{ref}} + 2S_0 \tag{7}$$

The $\Delta_{\mathrm{corr\text{-}ref}}$ values in Table 1 indicate that only one of 20 compounds (1,2,3-trimethylbenzene) does not meet this criterion. It is a statistically acceptable result (95% of correct answers).

However, the illustration of data processing mentioned above looks like an artificial example because it implies recalculating RI values of all analytes without exception (at the hypothetical condition that we know all of them *a priori*). In real analyses of complex mixtures, we do not know all analytes before analysis (to identify them is the final aim), but usually we can identify only several constituents using their specific mass spectrometric signs or preliminary chemical information.

Hence, the application of the algorithm considered to real multicomponent mixtures requires modification.

2.2. Application of the Algorithm of Comparing the Experimental and Reference GC Retention Indices to Multicomponent Mixtures

This is one of the most common analytical tasks. Most often, complex multicomponent mixtures under analysis contain some simple constituents that can be unambiguously identified without special processing of GC data. These components are, for example, impurities in commonly used solvents and reagents, plasticizers (like, e.g., phthalates), and so on. Hence, we can select just these compounds for comparing their RI_{exp} and RI_{ref} values. The number of such "reference points" may not be so large, but it is important that they should be evenly distributed in different parts of chromatograms (at the beginning, in the middle, and in the final part). Of course, if the samples under analysis contain small number of constituents (e.g., 1–2), the preliminary analysis of artificial mixtures of similar composition becomes necessary.

Continuing the consideration of the above-mentioned example (Table 1, Figures 2 and 3), we can select from 20 alkylarenes only five, namely, benzene, toluene (the first segment of the chromatogram), 1,3,5-trimethylbenzene (middle position), butylbenzene, and 1,2,4,5-tetramethylbenzene (the last part). All the selected compounds are marked in Table 1 in bold. After that, we should repeat the same mathematical operations as were done for the complete data set (equations (3), (4), (6), and (7)) for this reduced data set. Obviously, we obtain different values of the coefficients of the linear regression equation (3) than for the full data set (see footnotes to Table 1).

The plot of the dependence (3) for the reduced data set is shown in Figure 2(b). It illustrates slight variations of the angular coefficient "*a*" in equation (3); the dependence becomes slightly ascending instead of slightly descending. The values of $\Delta RI_{corr}^* = (RI_{corr} - RI_{ref})$ together with $\Delta_{corr\text{-}ref}^*$ are marked with asterisks in Table 1 for comparison with the initial values of $\Delta RI_{corr} = (RI_{corr} - RI_{ref})$ and $\Delta_{corr\text{-}ref}$. However, the average value of $\Delta_{corr\text{-}ref}^*$ (8 ± 6) appeared to be very close to $\Delta_{corr\text{-}ref}$ (9 ± 6). The same is true for the sum of residuals (S_0): 10.8 and 8.4, respectively. This means that RI_{corr}^* does not correspond to $RI_{ref} \pm 2S_0$ intervals for only one compound of 20 (namely, for *tert*-butylbenzene), which is statistically acceptable (95% of correct results).

The above reasoning can be illustrated by considering the RI values for 32 essential oil constituents, published by Engewald and co-authors [25] for a column with a standard nonpolar polydimethylsiloxane stationary phase, DB-1. The results of this data processing in comparison with the reference RI values from database [2] for identical phases are presented in Table 2. The reduced data set includes six compounds easily identifiable by mass spectra, namely, α-pinene, limonene, linalool, camphor, neral, and geranial (marked in bold). Similarly to Table 1, the parameters of all equations $(RI_{ref} - RI_{exp}) = aRI_{exp} + b$ for both complete and reduced data sets are indicated in the footnotes to Table 2. The average value of the difference $RI_{ref} - RI_{exp}$ for the complete data set is 13 i.u., and for the corrected reduced data set it is twice less, 6 i.u.

Table 2. Comparison of the retention indices of essential oils constituents [24] with RI values from the NIST database [2] on standard nonpolar polydimethylsiloxane stationary phases.

Compound	RI_{exp}	RI_{ref} [2]	$\Delta_{ref\text{-}exp}$	Complete set of reference data		Reduced set of reference data	
				RI_{corr}	$\Delta_{corr\text{-}ref}$	RI_{corr}^*	$\Delta_{corr\text{-}ref}^*$
α-Pinene**	**950**	**933 ± 4**	**-17**	**936**	**+3**	**930**	**-3**
Camphene	969	946 ± 5	-23	956	+10	950	+4
Myrcene	985	983 ± 3	-2	970	-15	967	-18
3-Carene	1022	1006 ± 5	-16	1010	+4	1005	-1
Limonene	**1038**	**1017 ± 3**	**-21**	**1026**	**+9**	**1022**	**+5**
Linalool	**1089**	**1086 ± 3**	**-3**	**1077**	**-9**	**1076**	**-10**
Isofenchol	1114	-	-	-	-	-	-
Fenchone	1120	1105 ± 6	-15	1109	+4	1108	+3

Citronellal	1138	1134 ± 4	-4	1127	-7	1127	-7
(*Z*)-Verbenol	1144	1133 ± 5	-11	1133	0	1148	+15
Camphor	**1146**	**1123 ± 6**	**-23**	**1135**	**+12**	**1136**	**+13***
(*Z*)-Pinocarveol***	1146	1135 ± 7	-11	1135	0	1136	+1
		1126 ± 6	-20	1135	+9	1136	+10
(*E*)-Verbenol	1147	1133 ± 5	-14	1136	+3	1137	+4
Isopulegol	1148	1144 ± 5	-4	1137	-7	1138	-6
(*Z*)-Pinocamphone	1160	1140 ± 5	-20	1149	+9	1150	+10
Borneol	1171	1151 ± 18	-20	1161	+10	1162	+11
Menthol***	1174	1157 ± 2	-17	1164	+7	1165	+8
		1165 ± 6	-9	1164	-1	1165	0
Terpinen-4-ol	1178	1164 ± 5	-14	1168	+4	1169	+5
Carenol	1188	-	-	-	-	-	-
α-Terpineol	1189	1175 ± 5	-14	1179	+4	1181	+6
Neocarveol	1189	-	-	-	-	-	-
Myrtenol	1195	1181 ± 5	-14	1185	+4	1187	+6
Verbenone	1202	1184 ± 7	-18	1092	+8	1194	+10
Neral	**1223**	**1218 ± 5**	**-5**	**1213**	**-5**	**1216**	**-2**
Carvone	1230	1218 ± 6	-12	1220	+2	1227	+9
Geraniol	1238	1237 ± 4	-1	1229	-8	1232	-5
Linalyl acetate	1242	1241 ± 2	-1	1233	-8	1236	-5
Geranial	**1250**	**1249 ± 8**	**-1**	**1241**	**-8**	**1245**	**-4**
Safrol	1275	1269 ± 7	-6	1266	-3	1271	+2
Bornyl acetate	1280	1270 ± 5	-10	1271	+1	1276	+6
(*Z*)-Pinocarvyl acetate	1301	-	-	-	-	-	-
Isosafrol	1357	1327 ± 31	-30	1349	+22^{4*}	1357	+30^{4*}
		1358 ± 6^{5*}	+1	1349	**-9**	1357	**-1**
Average standard deviation of reference RI values, s_{RI}		6.3					
Average difference $\Delta_{ref\text{-}exp}$:			13.2				
Average difference $\Delta_{corr\text{-}ref}$:					7 ± 5		6 ± 4

Footnotes: Dash means no information in database [2]; *) RI$_{corr}$ and $\Delta_{corr\text{-}exp}$ values are calculated using the reduced data set; **) the names and numerical data for compounds included in the reduced data set are marked in bold; ***) compounds with two different RI values in database [2]; ^{4*}) isosafrole is the single compound for which the values of ΔRI$_{corr}$ for both complete and reduced data sets differ from the average values of RI$_{corr}$ – RI$_{ref}$ by more than two standard deviations (see comments in the text); ^{5*}) RI value for β-isosafrole from author's RI collection. The parameters of the equation $\Delta_{ref\text{-}exp} = RI_{ref} - RI_{exp} = aRI_{exp} + b$ for the complete data set are $a = 0.014 \pm 0.021$, $b = -26.8 \pm 24.2$, $R = 0.13$, and $S_0 = 11.6$, and for the reduced data set, $a = 0.049 \pm 0.035$, $b = -66.6 \pm 39.6$, $R = 0.57$, and $S_0 = 9.0$.

The single compound with RI$_{corr}$ values that do not meet condition (7) for both complete and reduced data sets is isosafrole, 5-(1-propenyl)-1,3-benzodioxole. This is most probably due to the unreliable reference RI value for this compound in database [2]. Indeed, the RI value for isosafrole in [25] is 1357, while the reference value is 1327 ± 31 [2]. Such a large standard deviation is explained by uniting the RI values for so-called α- and β-isosafroles (*cis* and *trans* isomers) together. The RI value for the most often determined β-isosafrole (*trans*-isomer) is 1358 ± 6 (value from author's RI collection). Keeping this fact in mind, we obtain for this compound $\Delta_{corr\text{-}ref}^{*} = -1$ instead of +30. After this correction, the RI value for isosafrole (the last eluted compound) can be included into reduced data set. It is an additional illustration of the efficiency of the suggested approach.

The same example of 32 essential oil components can be considered using another database of reference information for semi-standard stationary phases [21]. Data processing in this case is illustrated by Table 3 similar to Table 2. The parameters of the equations $(RI_{ref} - RI_{exp}) = aRI_{exp} + b$ are presented in the footnotes.

Table 3. Comparison of the retention indices of essential oils constituents [24] with the reference RI values from [21] for semi-standard non polar stationary phases.

Compound	**RI$_{exp}$**	**RI$_{ref}$ [21]**	**$\Delta_{ref-exp}$**	**Complete set of reference data**		**Reduced set of reference data**	
				RI$_{corr}$	**$\Delta_{corr-ref}$**	**RI$_{corr}$***	**$\Delta_{corr-ref}$***
α-Pinene**	**950**	**932**	**-18**	**933**	**+1**	933	+1
Camphene	969	946	-23	953	+7	954	+8
Myrcene	985	988	+3	970	-18***	972	-16***
3-Carene	1022	1008	-14	1010	+2	1012	+4
Limonene	**1038**	**1024**	**-14**	**1028**	**+4**	1029	+5
Linalool	**1089**	**1095**	**+6**	**1083**	**-12**	1085	-10
Isofenchol	1114	1114	0	1110	-4	1112	-2
Fenchone	1120	1118	-2	1116	-2	1119	+1
Citronellal	1138	1148	+10	1136	-12	1139	-9
(*Z*)-Verbenol	1144	1137	-7	1142	+5	1145	+8
Camphor	**1146**	**1141**	**-5**	**1144**	**+3**	1147	+6
(*Z*)-Pinocarveol	1146	1135	-11	1144	+9	1147	+12
(*E*)-Verbenol	1147	1140	-7	1145	+5	1148	+8
Isopulegol	1148	1145	-3	1146	+1	1149	+4
(*Z*)-Pinocamphone	1160	1172	+12	1159	-13	1163	-9
Borneol	1171	1165	-6	1171	+6	1175	+10
Menthol	1174	1167	-7	1174	+7	1178	+11
Terpinen-4-ol	1178	1174	-4	1179	+5	1182	+8
Carenol	1188	-	-	-	-	-	-
α-Terpineol	1189	1189	0	1191	+2	1194	+5
Neocarveol	1189	-	-	-	-	-	-
Myrtenol	1195	1194	-1	1197	+3	1201	+6
Verbenone	1202	1204	+2	1205	+1	1208	+4
Neral	**1223**	**1235**	**+12**	**1227**	**-8**	1231	-4
Carvone	1230	1239	+9	1235	-4	1239	0
Geraniol	1238	1249	+11	1244	-5	1248	-1
Linalyl acetate	1242	1254	+12	1248	-6	1252	-2
Geranial	**1250**	**1264**	**+14**	**1257**	**-7**	1261	-3
Safrol	1275	1285	+10	1284	-1	1288	+3
Bornyl acetate	1280	1284	+4	1289	+5	1293	+11
(*Z*)-Pinocarvyl acetate	1301	1311	+10	1312	+1	1316	+5
Isosafrol	**1357**	**1373**	**+16**	**1372**	**-1**	1377	+4
Average difference $\Delta_{ref-exp}$:			8.4				
Average difference $\Delta_{corr-ref}$:					5 ± 4		6 ± 4

Footnotes: Dash means no information in [21]; *) RI$_{corr}$ and $\Delta_{corr-exp}$ values are calculated using the reduced data set; **) the names and numerical data for compounds included in the reduced data set are marked in bold; ***) myrcene is the single compound for which ΔRI_{corr} values for both complete and reduced data sets differ from the average values of $RI_{corr} - RI_{ref}$ by more than two standard deviations. The parameters of the equation $RI_{ref} - RI_{exp} = aRI_{exp} + b$ for the complete data set are $a = 0.080 \pm 0.014$, $b = -93.4 \pm 16.5$, $R = 0.72$, and $S_0 = 7.8$, and for the reduced data set, $a = 0.091 \pm 0.019$, $b = -103.0 \pm 22.4$, $R = 0.90$, and $S_0 = 6.5$.

The statistical characteristics of the data sets in Tables 2 and 3 are close to each other. For the second of them, the average value of the difference $\Delta_{ref-exp}$ for the complete initial data set is 8 i.u., the average value of $\Delta_{corr-ref}$ for the corrected values is 5± 4 i.u., and for the reduced data set it is 6 ± 4 i.u. However, this example illustrates the possibility of identifying compounds characterized by RIs on

doi:10.20944/preprints202510.1599.v1

standard nonpolar phases using reference data for semi-standard phases. The algorithms of data processing are the same in both cases.

Among 32 compounds in Table 3, only one, myrcene, does not meet criterion (7). The values of RI_{corr} and RI_{corr}* for myrcene differ from reference RIs [21] by more than $2S_0$; the cause of this difference remains unclear.

Similarly to Figure 3 (above), it seems reasonable to characterize the distributions of $\Delta_{ref\text{-}exp}$ and $\Delta_{corr\text{-}ref}$ for the reduced data set by the corresponding histograms. Figure 4(a) illustrates the spread of the initial differences ($RI_{ref} – RI_{exp}$); it is rather asymmetric, and most of the values are located within the range from –15 to +15 i.u. The distribution of the ($RI_{corr} – RI_{ref}$) values (b) corresponds to the data of Table 3 for comparing the RI data measured on standard nonpolar phases with the reference data for semi-standard phases and, moreover, calculated for the reduced data set. It is only slightly narrower (from –10 to +15 i.u.), but it becomes much more symmetric.

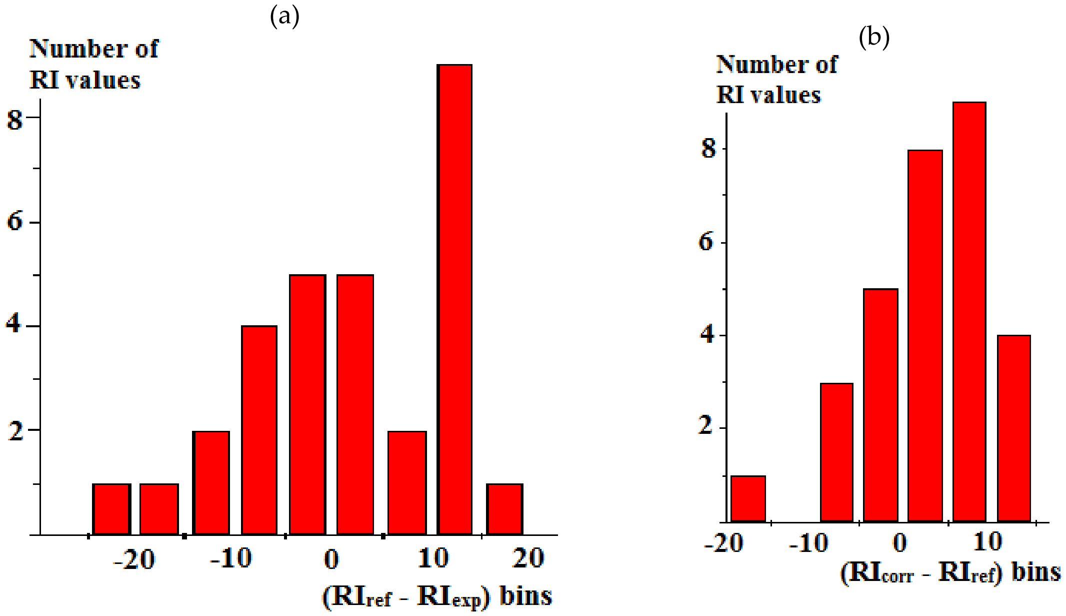

Figure 4. Histograms illustrating the results of comparing the RI data for essential oil components [24] with the reference data [21] for semi-standard non polar stationary phases (a) before and (b) after their correction. The bin sizes for both histograms are 5 i.u.

Hence, the algorithm proposed minimizes the shift of the differences of the initial experimental and reference RI values relative to zero and makes the distribution of these values narrower and more symmetric.

An important illustration of this algorithm is the situation when the reference RI values are determined under the conditions close to those of the experimental determination of RI values. It arises when the types of stationary phases are the same and the temperature conditions are close. In these cases, both coefficients "*a*" and "*b*" of equation (3) are obviously close to zero. To confirm this, let us consider the last Table 5, which illustrates the results of experimental testing of the algorithm using as an example a simple (only 10 constituents whose content exceeds 0.1% of the total peak area) sample of *Lavandula angustifolia* essential oil. Both experimental and reference data relate to the same semi-standard nonpolar stationary phases: HP-5 (experimental) and DB-5 (reference [21]). Therefore, the coefficients of the equation ($RI_{ref} – RI_{exp}$) = $aRI_{exp} + b$ (3) appear to be minimal among all examples considered above: $a = 0.008$, $b = –8.5$; the difference $\Delta_{corr\text{-}ref} = RI_{corr} – RI_{ref}$ is only 2 ±1, and the sum of residuals is $S_0 = 5$.

Table 4. Comparison of the retention indices of components of *Lavandula angustifolia* essential oil with reference data from [21] (all RI values were measured on semi-standard nonpolar stationary phases).

	RI_{exp}	RI_{ref} [21]	$\Delta RI_{ref-exp}$	RI_{corr}	ΔRI_{corr}
Camphene	949	946	-3	948	+2
Myrcene	993	988	-5	992	-1
(*E*)-Ocimene	1038	1044	+6	1037	-1
γ-Terpinene	1049	1054	+5	1048	-1
Linalool	1100	1095	-5	1098	+3
Alloocimene	1136	1128	-8	1134	-2
Linalyl acetate	1256	1254	-2	1253	-3
$C_{15}H_{24}$*	1365	-	-	1361	-
$C_{15}H_{24}$*	1384	-	-	1380	-
Aromadendrene	1443	1439	-4	1439	-4
Average difference $\Delta_{ref-exp}$:			4.8		
Average difference $\Delta_{corr-ref}$:					2 ± 1

Footnotes:*) For identification of the sesesquiterpenes, see discussion in the text. Parameters of the equation $\Delta RI_{ref-exp} = aRI_{ref} + b$: $a = -0.007 \pm 0.012$, $b = 6.0 \pm 13.9$, $R = -0.23$, and $S_0 = 5.2$.

The list of compounds in Table 4 includes two minor sesquiterpenes $C_{15}H_{24}$ with the retention indices of 1365 and 1384, remaining unidentified. The application of the algorithm considered gives corrected values of their retention indices, namely, 1361 and 1380. Hence, in the reference set of data [21] we should find sesquiterpenes with RIs within the intervals 1356–1366 and 1375–1385. Screening in the first of them leads to the rather "exotic" silfiperfol-4,7(14)-diene with RI 1358. However, an analog of this compound previously was identified in some species of *Lavandula* genus [26], and RI of silfiperfol-4,7(14)-diene determined in [27] (1367) is also close to RI_{corr} in Table 5. However, the identification of this compound should be considered as tentative.

For the next sesquiterpene with RI within the interval 1375–1385, according to databases [2] and [21], there are several possible candidates:

RI	Compound	Number of RI values for standard/semi-standard phases in [2]
1376 [2]	α-Copaene	377/698
1377 [21]	Silfiperfol-6-ene	None
1379 [21]	β-Patchoulene	10/23
1380 [21]	Daucene	16/14
1381 [21]	β-Panansinene	1/5

Let us additionally take into account the number of averaged RI values available for every compound in database [2] for standard/semi-standard nonpolar phases. It allows using such auxiliary (probability) criterion as the number of previous mentions of a particular compound [28]. Four compounds with the suitable RI from [21] appear to be little mentioned compared to α-copaene with RI 1376 [2] (377/698). However, this identification should also be considered as tentative.

3. Materials and Methods

The principal feature of this work is the possibility of using practically unlimited number of examples taken from the literature. However, for special consideration we selected the RI values of alkylarenes on Apiezon L phase [23] (can be classified as semi-standard),data for essential oil constituents on standard DB-1 nonpolar phase [25], and experimental data for essential oil of *Lavandula angustifolia* L. on semi-standard HP-5phase. The reference RI data for the standard and semi-standard nonpolar phases were taken from the NIST17 database [2].

The sample of Lavandula angustifolia essential oil (Technical Specification 20.53.10-006-74840603-2018, Mirrolla Lab., Leningrad district, n_D^{20} 1.4568) was purchased in a regular pharmacy. Isobutyl alcohol was of chemically pure grade (GOST (State Standard) 6016-77, Angarsk Chemical Reagent Plant, Russia); the internal references (*n*-tridecane and *n*-tetradecane) and other reference *n*-alkanes C_7–C_{18} were of chemically pure grade for chromatography (Reakhim, Moscow, Russia).

Gas chromatographic analysis of *Lavandula angustifolia* essential oil was carried out using its 10% solution in isobutyl alcohol with a Khromatek-Kristall 5000.2 gas chromatograph (Yoshkar-Ola, Russia) equipped with a flame ionization detector and a WCOT capillary column (length 10 m, internal diameter 0.53 mm, semi-standard HP-5 stationary phase, film thickness 2.65 µm). The analysis was performed with programmed heating from 90 to 240°C, ramp 6 deg min^{-1}. The carrier gas was nitrogen, flow rate 3.8 mL min^{-1}, linear velocity 34 cm s^{-1}, split ratio 1 : 3. The injector and detector temperature was 200°C. A 10 µL microsyringe volume was used for injecting 1.0 µL samples. The solution of *n*-alkanes C_7–C_{18} for determining the retention indices was injected separately.

Before ***GC–MS analysis***, the previous sample of *Lavandula angustifolia* essential oil was diluted 100-fold with the same solvent. The analysis was performed with a Shimadzu QO 2010 SE gas chromatograph–mass spectrometer equipped with an Optima 5 MS GC column, length 30 m, internal diameter 0.32 mm, and film thickness 0.25 µm, with programmed heating in the range 90–270°C, ramp 6 deg min^{-1}. The carrier gas was helium, flow rate 1.82 mL min^{-1}, linear velocity 53.3 cm s^{-1}, split ratio 1 : 10. The injector temperature was 200°C, and the interface and ion source temperature was 250°C. The ionization energy was 70 eV, the mass range was 40–500 Da,and the chromatogram recording start delay time was 2.0 min.

Processing and presentation of the results. Excel (Microsoft Office 2010) and Origin (versions 4.1 and 8.1) software was used for the statistical data processing and construction of the histograms. The QBasic program was used for calculating the linear-logarithmic RIs [1].

4. Conclusion

The suggested algorithm for comparing the experimental and reference GC retention indices as an important element of chromatographic and GC–MS identification of organic compounds is aimed at eliminating a significant element of uncertainty inherent in many contemporary recommendations: the use of fixed limiting differences between the experimental and reference GC retention indices, $\Delta RI = (RI_{ref} - RI_{exp}) \leq RI_{lim}$.The new algorithm is the most effective for complex multicomponent mixtures in which reliably identified components with known reference RI values can be revealed. It includes calculating the coefficients of regression equations $\Delta RI = (RI_{ref} - RI_{exp}) = aRI_{exp} + b$, followed by using these relations for recalculating of all other RI values into corrected data $RI_{corr} = RI_{exp} + \Delta RI$. This algorithm allows interpreting retention indices measured on standard nonpolar polydimethylsiloxane stationary phases using reference data for semi-standard nonpolar phases like polydimethylsiloxanes containing 5% phenyl groups and *vice versa*. It is applicable both to statistical processing of reference data (presented in the format "average arithmetic value ± permissible deviation") and to the results of single determinations. Furthermore, it allows using different RI databases and, if necessary, comparing different databases with each other and even recalculating the reference data from one database (e.g., semi-standard) into another one (standard).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: This experimental part of this work was carried out at the Center for Chemical and Materials Research of the Research Park of St. Petersburg State University. The authors are grateful to the staff of this Center for assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kovats' Retention Index System. In *Encyclopedia of Chromatography*. Ed. J. Cazes. 3rd Edn. Taylor & Francis. New York. **2010**, *2*, 1304-1310.
2. The NIST Mass Spectral Library (NIST/EPA/NIH EI MS Library, 2017 Release). Software/Data Version; NIST Standard Reference Database, Number 69, August 2017. National Institute of Standards and Technology, Gaithersburg, MD 20899: http://webbook.nist.gov (Accessed: September 2025).
3. Isidorov, V.A.; Zenkevich, I.G. Gas chromatographic – mass spectrometric determination of impurities of organic compounds in atmosphere. Khimia Publ. Leningrad, **1982**. (In Russian).
4. Sokolov, S.; Karnofsky, J.; Gustafson, P. The Finnigan Library Search Program. *Finn. Appl. Rep*., **1978**, № 2.
5. Zenkevich, I.G.; Ukolova, E.S. Dependence of chromatographic retention indices on a ratio of amounts of target and reference compounds // *J. Chromatogr. A.* **2012**, *1265*, 133-143. doi: 10.1016/j. chroma.2012.09.076.
6. Rospondek, M.J.; Narynowski, L.; Chachaj, A.; Gora, M. Novel aryl polycyclic aromatic hydrocarbons: phenylphenanthrene and phenylanthracene identification, occurrence and distribution in sedimentary rocks // *Org. Geochem*. **2009**, *40*, 986-1004. doi: 10.1016/j.orggeochem.2009.06.001.
7. Costa, R.; Pizzimenti, F.; Marotta, F.; Dugo, P.; Santi, L.; Mondelo, L. Volatiles from steam-distilled leaves of some plant species from Madagascar and New Zealand and evaluation of their biological activity // *Nat. Prod. Commun*. **2010**, *5*(11), 1803-1808.
8. Wang, X.-L.; Yu, W.-J.; Zhou, Q.; Han, R.-F.; Huang, D.F; Metabolic response of pakchoi leaves to amino acid nitrogen // *J. Integrative Agricult*. **2014**, *13*(4), 778-788.
9. Bodner, M.; Vagalinski, B.; Makarov, S.E.; Antic, D.Z.; Vujisic, L.V.; Leis, H.-J.; Raspotnig, G. "Quinone millipedes" reconsidered: evidence for a mosaic-like taxonomic distribution of phenol-based secretion across the julidae // *J. Chem. Ecol.* **2016**, *42*, 249-258. doi: 10.1007/s10886-016-0680-4.
10. Jung, E.P.; Alves, R.C.; Rocha, W.F. de C.; Monteiro, S. da S.; Ribeiro, L. de O.; Moreira, R.F.A. Chemical profile of the volatile fraction of Bauhinia forficata leaves: an evaluation of commercial and in natura samples // *Food Sci. Technol.* **2022**, *42*, № e34122.
11. Trovato, E.; Micalizzi, G.; Dugo, P.; Utczas, M.; Mondello, L. Use of linear retention indices in GC-MS. In: Baser K.H.C., Buchbauer G. *Handbook of essential oils. Science, technology, and applications*. 2015. Boca Raton (CA.). CRC Press, **2015**.
12. Weissburg, J.R.; Johanningsmeier, S.D.; Dean, L.L. Volatile compounds profiles of raw and roasted peanut seeds of the runner and Virginia market-types // *J. Food Res.* **2023**, *12*(3), 47-68. doi: 10.5539/jfr.v12n3p47.
13. Wesolowska, A; Jadczak, D.; Zyburtowich, K. Influence of distillation time and distillation apparatus on the chemical composition and quality of Lavandula angustifolia Mill, essential oil // *Polish J. Chem. Technol*. **2023**, *25*(4), 36-43.
14. Albarico, G.; Unbanova, K.; Houdkova, M.; Bande, M.; Tulin, E.; Kokoskova, T.; Kokoska, L. Evaluation of chemical composition and anti-staphylococcal activity of essential oils from leaves of two indigenous plant species, Litsea leytensis and Piper philippinum // *Plants*. **2024**, *13*, № 3555. doi: 10.3390/plants13243555.
15. Castillo, L.N.; Calva, J.; Ramirez, J.; Vidari, G.; Arnijos, C. Chemical analysis of the essential oils from three populations of Lippia dulcis Trevir. grown at different locations in southern Ecuador // *Plants*. **2024**, *13*, № 253. doi: 10.3390/plants13020253.
16. Smith, D.H.; Achenbach, M.; Yeager, W.I.; Andersson, P.J.; Fitch, W.H.; Rindfleisch, T.C. // Quantitative comparison of combined gas chromatographic/mass spectrometric profiles of complex mixtures // *Anal. Chem*. **1977**, *49*(11), 1623-1632. doi: 10.1021/ac50019a041.
17. Zenkevich, I.G.; Babushok, V.I.; Linstrom, P.J.; White, V E.; Stein, S.E. Application of histograms in evaluation of large collections of gas chromatographic indices // *J. Chromatogr. A.* **2009**, *1216*, 6651-6661. doi: 10.1016/j.chroma.2009.07.065.
18. Bizzo, H.R.; Brilhante, N.S.; Nolvachai, Y.; Marriott, P.J. Use and abuse of retention indices in gas chromatography // *J. Chromatogr. A.* **2023**, *1708*, № 464376. doi: 10.1016/j.chroma.2023.464376.
19. Matheson, A. Identifying and rectifying the misuse of retention indices in GC // *LC-GC International*. **2024**, *20*(12), 2-8.
20. Zenkevich, I.G. Information maintenance of gas chromatographic identification of organic compounds in ecoanalytical investigations // *J. Anal. Chem*. **1996**, *51*(11), 1140-1148. (In Russian).

21. Adams, R. *Identification of essential oils components by GC/MS.* Version 4. Carol Stream, IL: Allured Publ. Corp., 2007. 804 p.
22. Babushok, V.I.; Linstrom, P.J.; Zenkevich, I.G. Retention Indices for Frequently Reported Compounds of Plant Essential Oils // *J. Phys. Chem. Ref. Data.* **2011**, *40*(4), № 043101. doi: 10.1063/1.3653552.
23. Brown, I. The composition of a Lurgi brown coal tar. III. The neutral oil fraction boiling from 30 to 130°C and 130–172°C // *Austr. J. Appl. Chem.* **1960**, *11*(4), 403-433.
24. Linnik, Yu.V. *Least squares method and fundamentals of processing the observations.* Moscow, Physical-Mathematical Literature Press, 1958. 334 p. (In Russian).
25. Engewald, W.; Knobloch, T.; Haufe, G.; Muller, M.; Pohris, V. A novel method for terpene pattern determination of essential oils by selectivity tuning in GC // *Fresenius' J. Anal. Chem*. **1991**, *341*(10), 641-643. doi: 10.1007/BF00322279.
26. Benadi, T.; Lemhadri, A.; Harboul, K.; Chtibi, H.; Khabbach, A.; Jadouali, S.M.; Quesada-Romero, L.; Louahliq, S.; Hammani, K.; Ghaleb, A.; Lee, L.-H.; Bouyahya, A.; Rusu, M.E.; Akhazzane, M. Chemical profiling and biological properties of essential oils of Lavandula stolchas L., collected from three Moroccan sites: in vitro and in silico investigations // *Plants*. **2023**, *12*(6), № 1413. doi: 10.3390/plants120614-13.
27. Obistioiu, D.; Cristina, R.T.; Schmerold, I.; Chizzola, R.; Stolze, K.; Nichita, I.; Chiurciu, V. Chemical characterization by GC-MS and in vitro activity against Candida albicans of volatile fractions prepared from Artemisia dracunculus, Artemisia abrotanum, Artemisia absinthium and Artemisia vulgaris // *Chem. Cent. J.*, **2014**, *8*(6), ID 24475951. doi: 10.1186/1752-153X-8-6.
28. Zenkevich, I.G. Non-traditional criteria for the identification of organic compounds by chromatography and chromatography – mass spectrometry // *J. Anal. Chem*. **1998**, *53*(8), 725-731. (In Russian).