Pattern Recognition Methods for Discrimination of Essential Oils (Rose Oils) by Their Gas Chromatograms

By N. Dimov and A. Tsoutsoulova, Chemical Pharmaceutical Institute, Sofia; and E. Stojanov, Chemical and Technological Institute, Bourgas, Bulgaria

It is possible to distinguish between essential oils of the same type and their mixtures by analysing their gas chromatograms using a pattern recognition method. An approach for the selection of peaks used in the calculation of the differentiation criterion is demonstrated. The accuracy and scope of the method are demonstrated using sixteen rose oil samples (genuine or blended) and five-fold repeat analyses of the same rose oil. The selectivity of this method permits an objective confirmation of the differences or similarities in the composition of the compared essential oils.

Essential oils are used mainly in the perfumery, cosmetic and food industries. Their market prices depend primarily on their quality as determined by organoleptic tests. This assessment however, because of its subjective nature, cannot be used with sufficient confidence in solving technological, commercial or other issues. Moreover, the physicochemical indices determined by conventional chemical methods, applied in detecting blending or adulteration, show too much variation to be reliable when small deviations are concerned.

Better results can be achieved when applying chromatographic and spectroscopic methods. Visual comparison of gas chromatograms of different types of essential oils normally is sufficient for distinguishing between them. However, such evaluations cannot be considered as reliable when a great number of samples of specific types of oils or oil mixture are involved.

0272-2666/87/0006-4501$04.00/00—

Table I. Peaks Used in the Mathematical Calculations of the Differentiation Criterion

Compounds	Rose Oil 1 Mean Content (%)	Five Injections Relative Error	Geranium Oil Content (%)
cis-rose oxide	0.22	4.6	0.64
trans-rose oxide	0.122	4.2	0.24
linalool	1.86	2.7	5.38
isoborneol	0.35	3.2	0.27
caryophyllene	0.32	1.6	0.69
citronellylacetate	0.56	4.9	0.88
neral	0.51	1.6	0.79
geranial neryl acetate	1.26	3.6	1.49
citronellol	22.50	2.4	26.01
geraniol	22.20	2.8	14.04
$n\text{-}C_{18}H_{38}$	0.22	4.0	0.10
phenylethylalcohol	0.90	8.8	1.88
$n\text{-}C_{19}H_{40}, n\text{-}C_{19}H_{38}$	15.74	7.0	0.10
$n\text{-}C_{20}H_{42}$	1.00	8.7	0.10
methyl eugenol	1.44	6.7	0.21
$n\text{-}C_{21}H_{44}$,eugenol	4.48	4.7	0.31
$n\text{-}C_{22}H_{46}$	0.40	5.8	0.10
farnesol	1.60	1.4	0.71
$n\text{-}C_{23}H_{48}$	0.90	4.3	0.52
unknown	0.20	7.8	0.10
$n\text{-}C_{24}H_{50}$	0.25	6.2	0.22
$n\text{-}C_{25}H_{52}$	0.45	4.8	0.10

Note: The identification of the peaks for this treatment is not necessary. The peaks with equal t'_R should be used.

Recently, in such cases the data have been treated by some of the methods for pattern recognition.[1-9] The aim of the present paper is to test whether it is possible to distinguish between oils of one type from different geographic regions and mixtures of them by applying a method of pattern recognition in analysing their gas chromatograms.

Experimental

The analyses were carried out on a model 2B Perkin Elmer gas chromatograph with flame ionization detector. The following conditions were used:

Column—WCOT, Carbowax 20M, d_f = 0.13 m, 0.3mm x 50m
Temperature—oven: 50°C to 195°C (15 min hold) at 2°C/min heating rate; detector: 260°C; injector: 240°C
Carrier Gas—argon: 0.74 ml/min or u = 17.45 cm/sec
g-pack-value[10,11]—1.212 (lower than value reported for non-capillary columns[11] but on the basis of results obtained totally satisfactory)
^{n}eff linalool[12]—25995
Sample—0.1 μl, split 1:92

The investigations were carried out with eight rose oil samples taken from different regions, three commercial lots, four samples obtained by mixing of four rose oils from different regions, one Turkish rose oil sample and also one geranium oil sample, the composition of which is similar to that of the rose oils. To determine the repeatability of the analysis, one oil was injected five times on different days. The coefficients of variation obtained for components in which content is above 0.1% were varied from 2 to 9.5%.

Methodology

The Anders approach forms the basis for the calculations in our study.[13] The vector elements in this paper are the percentage contents (from the peak areas or their heights) of the various constituents. All of the vector elements are divided each by the other A_i [J] / A_i [L]. A triangular *matrix is developed from these ratios. Its solution* gives the differentiation criterion, Z. The Z-value by the Anders method depends on another criterion, M, whose value is chosen arbitrarily.

The influence of its values on the differentiating power of the Z criterion was critically evaluated by Dimov et al.[14] It was shown that, for a correct estimation, time-consuming calculations are necessary.

Stojanov et al. proposed a new criterion of similarity (TK):

$$TK = \frac{2 \sum_{i=2}^{n-1} \sum_{j=1}^{i-1} \left(1 - \frac{x_{(i,j)}}{a_{(i,j)}} \right)^2}{n\ (n-1)}$$

where $x_{(i,j)}$ and $a_{(i,j)}$ are the correlation matrix of the two samples (x_i and a_i—percentage content of i-th peak in the sample x and a) and n is the number of the peaks in the samples. When $x_{(i,j)}/a_{(i,j)} > 1$, the reciprocal value is used. The lowest value (TK_0) is connected with the error of the applied analytical method. Its value depends on the mean error of the content (in %) of the selected for the distinguishing peaks:

$$TK_0 = \left[n/\ (n-1) \right] \bullet \left[2.\ \alpha^2/\ (1+\alpha)^2 \right]$$

where n is the number of the selected peaks and α is calculated from the experimental data of 3 to 5 repeatable injections confidence interval of their content. Therefore knowing the repeatability of the analysis, one can find the TK_0 value. The $TK \leq TK_0$, the samples are undistinguishable.

Result and Discussion

Applying a pattern recognition method to the analysis of the various rose oil samples first required a criterion for determining which gas chromatographic peaks should be used. The following approach was applied:

- —Only the data for peaks with the best reproducibility with regard to retention characteristics and quantification were used
- —Peaks which emerge close to each other and which can overlap with only slight changes in the test conditions were eliminated.

The adjusted retention times (t'_R), relative retention ($r_{1,2}$), heights (h) and percentage composition (calculated by normalisation) for each peak of the selected 16 rose oil samples were experimentally determined. The $r_{1,2}$ values are almost constant while t'_R values vary within a larger range. For this reason t'_R values were used in our further study thus testing the recognition method under more stringent conditions.

On this basis the peaks of α-terpenyl acetate and borneol with a difference in t'_R of 10 sec., α-terpineol and n-hexadecane—28 sec.; geranyl acetate and n-heptadecane—18 sec.; phenylethyl acetate and nerol—25 sec. were eliminated from the initial 36 chosen peaks. Thus, 26 peaks remained for further treatment.

A subsequent reduction was made by discarding those peaks in the beginning of the chromatogram originating from components with high volatility which, for that reason, will show lack of repeatability in quantification. Thus, 22 peaks (Table I) remained for mathematical treatment.

On the basis of the experimentally obtained data, we assume 10% as the maximum permissible interval of variation. This is a margin, big enough to accommodate results from different laboratories and make them representative. The TK_0 value, obtained under such conditions, is 0.08.

Table II. Value of the Differentiation Criterion (TK) for 8 Rose Oils, Turkish Rose Oil and Geranium Oil ($TK_0 = 0.08$)

	1		2	3	4	5	6	7	8		
Sample	2nd Inj.	Mean Content								Turkish Rose Oil	Geranium Oil
1. 1st injection	0.078	0.044									
1. 2nd injection		0.053									
1. Mean content			0.11	0.10	0.10	0.11	0.17	0.18	0.21	0.19	0.49
2.				0.12	0.15	0.17	0.17	0.22	0.20	0.26	0.50
3.					0.13	0.14	0.16	0.19	0.20		0.48
4.						0.18	0.22	0.21	0.25		0.48
5.							0.17	0.18	0.23	0.29	0.51
6.								0.22	0.19		0.51
7.									0.26		0.55
8.											0.46

Note: The underlined values refer to undistinguishable oils.

The value of the differentiation criterion TK from two injections with the greatest differences in percentage of one and the same sample of rose oil and the mean content (see Table I) are shown in Table II. Data for seven other rose oils from different regions are included as well as for one Turkish rose oil and one geranium oil.

Obviously, the selected injections of the same sample with the greatest percentage difference are classified as undistinguishable by TK_0. They are also undistinguishable when comparing them with their mean content.

All other rose oils can be differentiated because they have a TK value higher than TK_0. Besides, the value of TK is different for every oil. The Turkish rose oil was compared with the Bulgarian rose oils 2 and 5 because they have similar TK values with respect to rose oil 1. The obtained TK values confirm the similarity between oils 2 and 5 and at the same time better obviate the existing slight differences between them. Thus comparison of oils with one or several other oils facilitates their differentiation and reveals discrete differences in their composition due to varying ecological conditions.[16-18]

The results from Table II show that regardless of the differences between the rose oils, all have TK values in the range of 0.10 to 0.29. The TK value of the Turkish oil also lies in this range. Geranium oil, used at the beginning of the twentieth century for rose oil adulteration because of its similar composition,[19] has a TK value considerably out of the range of rose oil values.

This good differentiating power of TK and the possibility it gave for clustering encouraged us in using it for distinguishing three commercial lots and four samples obtained by blending rose oils from different regions. By the method proposed, as shown in Table III, it is possible to distinguish between both the commercial lots and the various mixtures.

The results from Table III confirm the narrow range of TK values in which the rose oils fall. There are no TK values above 0.29, and the value obtained when comparing commercial lot B with the mixture II shows that their compositions are very similar, i.e., almost undistinguishable. Repeat organoleptic tests (by the 10 point scale) of the same two samples (every two hours, during 24 hours interval) showed a difference of 0.3 points, i.e., according to the organoleptic tests they also have a closely similar odour.

The above results show that the proposed method allows the differentiation of oils belonging to one type. The obtained values of the differentiation criterion TK (Table II) give a quantitative estimate for the similarity in the composition (organoleptic quality) and potential interchangeability.

The results in Table III confirm that combining rose oils, from different rose valley microregions and harvests, yields blends of Bulgarian rose oil with different compositions and different odour, shades and notes respectively. The proposed method gives greater veritability and reliability when solving technological, criminal and

Table III. Values of the Differentiation Criterion (TK) for 4 Mixtures and 3 Commercial Lots of Rose Oils (TK_0 = 0.08)

	Mixtures			Commercial Lots		
Sample	II	III	IV	A	B	C
I	0.12	0.12	0.22	0.20	0.18	0.18
II		0.26	0.18	0.27	0.09	0.15
III			0.28	0.15	0.25	0.12
IV				0.28	0.12	0.26
A					0.28	0.18
B						0.15

other problems and offers an objective confirmation of the organoleptic test, which although subjective, is still of decisive importance in the perfumery and cosmetic industry.

Conclusion

A mathematical method is proposed for differentiating genuine or blended essential oils (rose oils) from different regions, using their gas chromatograms. A new differentiation criterion, TK, is applied. Its value depends on the error of the analytical method used. All samples with TK lower than a critical Tk_0 value can be considered as undistinguishable.

The method was used for differentiating sixteen rose oils and its scope and accuracy were tested in five-fold repeat analyses of the same rose oil.

The value of the differentiation criterion TK enables rose oils to be classified irrespective of the year, the region and even the country of production. The value of the differentiation criterion TK and the quantitative estimation of similarity permits the clustering of the oils and can be used for assessing their interchangeability.

Acknowledgements

The authors are grateful to Dr. G. Stoev for preparing the highly efficient glass capillary column, on which the study has been carried out and to D.Ph. N. Nicolov for the organoleptic testing of the samples.

References

Address correspondence to Assoc. Prof. Neno Dimov, Chemical Pharmaceutical Institute, 1156 Sofia, Bulgaria.

1. P. C. Jurs, T. L. Isenhour, Chemical applications of pattern recognition, Wiley-Interscience, New York (1975)
2. G. E. Veress, Pattern recognition in analytical chemistry, Trends in Analytical Chemistry, **1**, (16), 374 (1982)
3. M. Forina and E. Tiscornia, Pattern recognition methods in the prediction of Italian olive oil origin by their fatty acid content, Annali di Chimica, **72**, 143, (1982)
4. P. Tétényi, J. Fischer, A. Csiszar, E. Héthelyi and P. Kaposi, Approach to infraspecific chemical taxonomy of Tanacetum vulgare L. by mathematical means, VII International Congress of Essential Oils, Kyoto, 198 (1977)
5. R. Cantagrel, Application de l'analyse statistique multidimensionnelle à la différenciation des huiles essentielles de lavandes et de lavandins, Parfums, Cosmetiques, Aromes, (61), 73 (1985)
6. J-C. Dore, J-N. Jaubert, Exemple d'étude systémique de chromatogrammes d'essences naturelles par analyse factorielle des correspondances, Parfumes, Cosmetiques, Aromes, (61), 79 (1985)
7. V. E. Stepanenko, Group analysis and pattern recognition as a basis of chromatographic identification, Zn Anal Khim., **40** (5), 881 (1985)
8. M. Chien, Analysis of complex mixtures by gas chromatography/mass spectrometry using a pattern recognition method, Anal. Chem., **57** (1), 348 (1985)
9. R. E. Lea, C. R. Bramston, and J. Tschida, Pattern recognition for identification and quantification of complex mixtures in chromatography, Anal. Chem., **55**(4), 626 (1983)
10. H. van den Dool, Standardisation of GC Analysis of Essential Oils, Proefschrift, Rijksuniversiteitte Groningen, Rotterdam (1974)
11. Analytical Methods Committee, Application of Gas-Liquid Chromatography to the Analysis of Essential Oils, Part VII. Fingerprinting of Essential Oils by Temperature-Programmed Gas-Liquid Chromatography Using a Carbowax 20M Stationary Phase, Analyst, **105**, 262 (1980)
12. ISO 7609: Essential oils—Analysis by gas chromatography on capillary columns—General method.
13. O. U. Anders, Ratio matching a statistical aid for discovering generic relationships among samples, Analytical Chemistry, **44** 1930 (1972)
14. N. Dimov, E. Stoyanov, G. Velikov, Differentiation of commodity production by one of the methods of image identification. Chimia i Industria, **5**, 210 (1984)
15. E. Stoyanov and N. Dimov in press.
16. N. Nicolov and I. Ognianov, Etudes et conditions actuelles de la Rose oléagineuse de Kazanlik et de L'essence de Rose Bulgare, III Congrès International des Huiles Essentielles, Plovdiv, 27 (1964)
17. N. Nicolov, A. Tsoutsoulova and N. Nenov, Essence de Roses et autres huiles essentielles Bulgare, Riv. Ital. EPPAOS, Luglio (1976)
18. N. Nicolov, P. Dragostinov, A. Tsoutsoulova, F. Portarska, B. Apostolova, and N. Nenov, Bulgarian Rose oil and its quality, VII International Congress of Essential Oils, Kyoto, 159 (1977)
19. V. Topalov and I. Irintchev, The Rose production in Bulgaria, Plovdiv (1967)