ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN – POLONIA
VOL. XXII, N 2, 10, 2009 SECTIO DDD DOI: 10.2478/v10080-008-0181-z

Chair and Department of Medicinal Chemistry, Medical University of Lublin

ALEKSANDRA BOJARCZUK, ROBERT SKIBIŃSKI, ŁUKASZ KOMSTA

Multivariate analysis of UV spectra of complex herbal mixtures and essential oils

Wielowymiarowa analiza widm UV złożonych wyciągów roślinnych i olejków eterycznych

The spectra of various chemical samples are often processed chemometrically as multivariate signal-like data. They contain complex information regarding the structure of sample ingredients [4, 7]. Multivariate treatment of the spectra can result in interesting conclusions about similarity and dissimilarity between samples and correlation of such dissimilarities with sample composition or properties.

The spectra in UV region are treated as highly inspecific and difficult to interpret. Prediction of compound structure from its UV spectra, or prediction of sample constituents only from complex UV spectrum is almost impossible [5]. Therefore, almost no papers deal with this problem. An exceptional study resulting in good conclusions was given by Luthria et al. [6]. They successfully used ANOVA-PCA and UV spectrometry to trace cultivars and growing conditions of broccoli (Brassica oleacea). This paper proves that in certain conditions, UV spectrum of complex mixture can contain some useful information. It inspired us to perform a multivariate analysis of UV spectra of complex mixtures from pharmaceutical field of interest – essential oils and herbal extracts.

MATERIAL AND METHODS

The essential oils and other plant fluid extracts (Table 1 gives the reference numbers) were of pharmaceutical purity from various manufacturers. They were bought in a local drugstore in Lublin, Poland. The essential oil samples were prepared by taking 10 µl amount of the oil (with a microsyringe) and dissolving in 4 ml of methanol (spectroscopic grade, POCH, Gliwice, Poland). The resulting solutions were diluted by the same solvent in a ratio 1:10 twice to obtain the final concentration 25 µl/L. The fluid extract samples were prepared in the same way, but in the final concentration 250 µl/L (one dilution).

The spectra were recorded in 1 cm quartz cells on Hitachi UV-2001 double-beam spectrophotometer in the range 200–400 nm, with 0.5 nm resolution (400 absorbance values) against pure methanol. The data were then transferred into R-project open source statistical environment as an absorbance matrix with 44 rows and 400 columns.

No.	Plant	Extract type	No.	Plant	Extract type
1	*Abies sibirica*	*Oleum*	23	*Gingko biloba*	*Tinctura*
2	*Adonis Vernalis*	*Tinctura*	24	*Hypericum perforatum*	*Intractum*
3	*Aesculus hippocastanum*	*Intractum*	25	*Hypericum perforatum*	*Succus*
4	*Aniba rosaeodora*	*Oleum*	26	*Lavenda officinalis*	*Oleum*
5	*Arctium lappa*	*Succus*	27	*Malaleuca alternifolia*	*Oleum*
6	*Arnica Montana*	*Tinctura*	28	*Melissa officinalis*	*Oleum*
7	*Betula pendula*	*Succus*	29	*Melissa officinalis*	*Intractum*
8	*Calendula officinalis*	*Tinctura*	30	*Mentha piperita*	*Tinctura*
9	*Capsicum annuum*	*Tinctura*	31	*Mentha piperita*	*Oleum*
10	*Cedrus libani*	*Oleum*	32	*Ocimum basilicum*	*Oleum*
11	*Cinnamomum verum*	*Oleum*	33	*Origanum majorana*	*Oleum*
12	*Citrus bergamia*	*Oleum*	34	*Picea abies*	*Oleum*
13	*Citrus limon*	*Oleum*	35	*Pimpinella anisum*	*Oleum*
14	*Citrus sinensis*	*Oleum*	36	*Pinus silvestris*	*Oleum*
15	*Citrus x paradise*	*Oleum*	37	*Salvia officinalis*	*Tinctura*
16	*Convallaria majalis*	*Tinctura*	38	*Santalum album*	*Oleum*
17	*Crataegus oxycantha*	*Intractum*	39	*Syzygium aromaticum*	*Oleum*
18	*Crataegus oxycantha*	*Tinctura*	40	*Taraxacum officinale*	*Succus*
19	*Cynara scolymus*	*Tinctura*	41	*Tussilago farfara*	*Succus*
20	*Echinacea purpurea*	*Tinctura*	42	*Urtica dioica*	*Succus*
21	*Eucalyptus globulus*	*Oleum*	43	*Valeriana officinalis*	*Tinctura*
22	*Geranium graveolens*	*Oleum*	44	*Viscum album*	*Intractum*

RESULTS AND DISCUSSION

Before further processing, the spectra were standardized by means of Standard Normal Variate transformation (SNV) [1]. This practice is common and removes from the matrix differences caused by sampling error or inadequate spectrophotometer zeroing. All spectra become then standardized, i.e. they have mean absorbance equal to zero and variance equal to one. Figure 1 presents the spectral matrix before (A) and after (B) transformation.

Unscaled Principal Component Analysis, as a probably most often used tool for multivariate data projection [4], was used first. This computational technique decorrelates the variables (absorbances) and converts them into linear combinations, called principal components (PCs). The first two principal components explain the maximum available overall variance, so their plot can be used for a comparison of the compounds or methods. In our case, the first PC explained 66% of total variance, both of them explain 80% of variance. Therefore, the compression of dataset was very good and enough to perform the projection onto the plane, where the distances between samples are the most efficient approximation of the multivarate Euclidean distance between them.

This preliminary analysis showed that two spectra (11 and 35) are large outliers, due to intensive absorption bands around 260 and 290 nm, absent in all other samples. Therefore, we decided to perform the final analysis by robust PCA [4], according to the method given by Croux and Ruiz-

Gazen [2]. This prevented the distortion of data projection caused by outlying objects and gave the possibility to estimate the similarity in a more efficient way. The software implemented by by Croux et al. [3] and availalble as the R package "pcaPP" was used without any modifications.

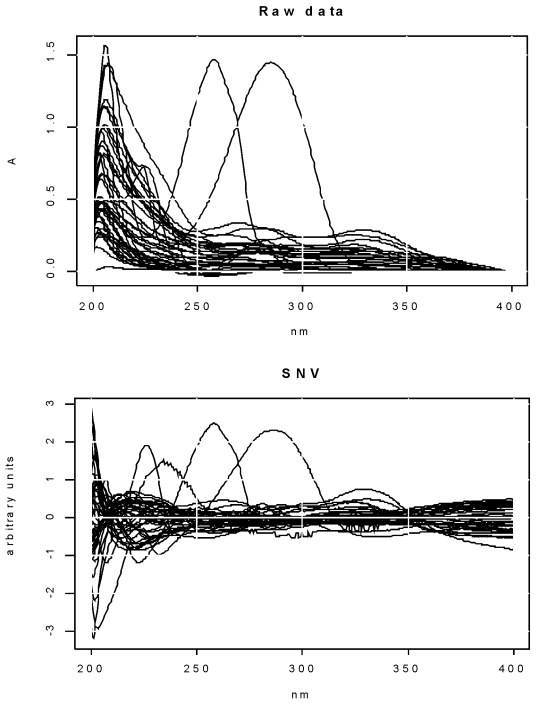

Fig. 1. The analyzed spectra without any pretreatment (A) and after Standard Normal Variate and centering (B)

The comparison made by robust PCA is presented in Figure 2 as scores (A) and first two loading vectors (B). The following observations can be drawn from the analysis:

1. Most of the essential oils form a cluster with a large value of PC1. The exceptions are: 10, 11, 31, 32, 35, 39. Therefore, from the loading vectors, it can be concluded that most of the oils have a strong band just above 200 nm, weak absorbance around 225 nm and strong absorbance above 300 nm. The sign of PC1 (mostly positive for oils, negative for other extracts) can be treated as the simplest discriminant for classifying samples. Only oils with strong aromatic constituents fail with such discrimination.

2. The second PC explains intensity of absorption around 250 nm, which is correlated with weaker absorption above 300 nm.

3. There is no strong correspondence between similarity and chemical composition. For example, oils 10 and 39 are very different in composition, but very similar in spectral properties.

4. There is a significant difference between spectrum of the oil and tinctura from the peppermint (pair 30–31). It is caused by the chemical difference between oil and other extracts. Quite a large difference also exists between intractum and tinctura from Crataegus spp. (17–18). On the contrary, intractum and succus from St. John's worth (24–25) are very similar.

5. There is no clustering against the type (form) of the extract (all extracts are mixed).

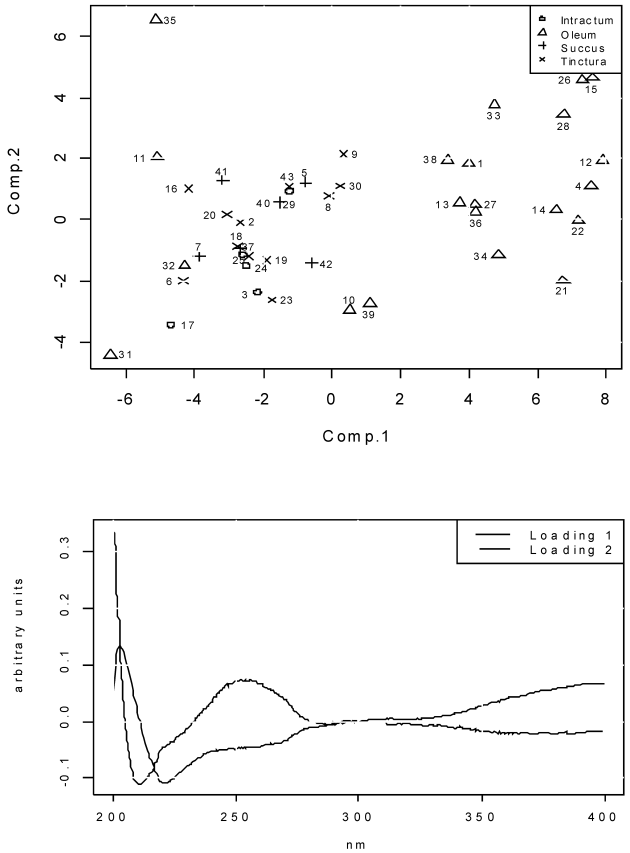

Fig. 2. A multivariate comparison of the spectra by robust PCA (A) and two first loading vectors over the wavelength

CONCLUSIONS

Although there is no way to distinguish subsequent extracts or oils, the difference between oils and extracts forms a general trend, which can be used for discriminant analysis of such complex samples.

REFERENCES

1. Barnes R. J., Dhanoa M. S., Lister S. J.: Standard Normal Variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectr., 43, 772, 1989.
2. Croux C., Ruiz-Gazen A.: High Breakdown Estimators for Principal Components: the Projection-Pursuit Approach Revisited', J. Multivariate Anal., 95, 206, 2005.
3. Croux C., Filzmoser P., Oliveira M.: Algorithms for Projection-Pursuit Robust Principal Component Analysis, Chemom. Intell. Lab. Syst., 87, 218, 2007.
4. Daszykowski M., Walczak B., Massart D. L.: Projection methods in chemistry, Chemom. and Intell. Lab. Syst., 65, 97, 2003.

5. Komsta Ł., Skibiński R., Gałaszkiewicz A., Grech-Baran M.: Mapping of the UV spectra of drugs by principal and independent component analysis (PCA and ICA), Annales UMCS. sect. DDD, 20, 115, 2007.
6. Luthria D. L., Mukhopadhyay S., Robbins R. J. et al.: UV Spectral Fingerprinting and Analysis of Variance – Principal Component Analysis: a Useful Tool for Characterizing Sources of Variance in Plant Materials. J. Agricult. Food Chem., 56, 5457, 2008.
7. Mark H., Workman J.: Chemometrics in Spectroscopy, Academic Press, 2007.

SUMMARY

44 UV spectra of essential oils and plant extracts were analyzed in multivariate way by robust principal component analysis. Although the samples were not clustered against their chemical composition, an interesting dependence was found. Most of the oils had the first PC1 positive, other samples – negative, which was connected with an intensive band just above 200 nm and above 300 nm. These results suggest the ability to distinguish between oils and other extracts by discriminant analysis models, and first PC1 acts as a good discriminant.

STRESZCZENIE

44 widma UV olejków eterycznych i wyciągów roślinnych poddano wielowymiarowej analizie porównawczej techniką stabilnej analizy składowych głównych. Mimo iż wiele różnych próbek wykazywało podobne do siebie widma, zaobserwowano wyraźną zależność pomiędzy olejkami a pozostałymi wyciągami, związaną z pasmem absorpcji nieco powyżej 200 nm i powyżej 300 nm. Zdecydowana większość olejków posiadała pierwszą składową główną dodatnią, zaś wyciągów – ujemną. Wyniki te świadczą o tym, iż istnieje możliwość tworzenia modeli dyskryminacyjnych odróżniających olejki od innych wyciągów, a pierwsza składowa główna jest dobrym przykładem takiej dyskryminanty.