You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice

**Title:** The HPLC/DAD Fingerprints and Chemometric Analysis of Flavonoid Extracts from the Selected Sage (Salvia) Species

**Author:** Mieczysław Sajewicz, Dorota Staszek, Michał S. Wróbel, Monika Waksmundzka-Hajnos, Teresa Kowalska

UNIWERSYTET ŚLĄSKI W KATOWICACH          Biblioteka Uniwersytetu Śląskiego          Ministerstwo Nauki i Szkolnictwa Wyższego

*Research Article*

# The HPLC/DAD Fingerprints and Chemometric Analysis of Flavonoid Extracts from the Selected Sage (*Salvia*) Species

**Mieczysław Sajewicz,[1] Dorota Staszek,[1] Michał S. Wróbel,[1] Monika Waksmundzka-Hajnos,[2] and Teresa Kowalska[1]**

[1] *Institute of Chemistry, University of Silesia, 9 Szkolna Street, 40-006 Katowice, Poland*
[2] *Department of Inorganic Chemistry, Faculty of Pharmacy, Collegium Pharmaceuticum, Medical University of Lublin, 4a Chodźki Street, 20-093 Lublin, Poland*

Correspondence should be addressed to Teresa Kowalska, teresa.kowalska@us.edu.pl

The results of spectrophotometric and HPLC/DAD analysis are discussed, and a comparison is made of selectively extracted flavonoid fractions derived from twenty six sage species belonging to the *Salvia* genus. The sage samples were harvested in the vegetation seasons 2007, 2008, and 2009. It was a goal of this study to find out which species contain the highest yields of flavonoids (recognized for their free-radical-scavenging activity), as those with the highest yields could be applied in official medicine. It was spectrophotometrically established that the four sage species can be recognized for their highest flavonoid levels, while the HPLC/DAD analysis pointed out to the four other species. The source of the discrepancy between the two evaluation approaches was discussed. Moreover, the HPLC/DAD fingerprints of the flavonoid fraction underwent a chemometric pre-treatment, and then the purified fingerprints were analyzed by means of Principal Component Analysis (PCA) for the differences in the harvesting period. A difference was revealed between the herbs harvested in the 2007 season, and those harvested in 2008 and 2009. The main source of this difference could be the seasonal weather variation and the relatively longest storage period with the plants harvested in 2007.

## 1. Introduction

In spite of a vast number of the different species and a wide popularity of the plants belonging to the *Salvia* genus, relatively little attention has been paid to phytochemical analysis of these plants, prior to our own systematic research (e.g., [1–4]). In the course of the centuries, different sage species have gained the repute for their outstanding therapeutic, culinary, and decorative valor. However, the official European medicine recognizes just one sage species for its curative properties, which is *Salvia officinalis*. It is a well-known fact that the curative properties of many plants are due to the high contents of phenolics, which act as the free-radical scavengers. Thus, it became an objective of this study to analyze a selection of the different sage species popular in Central and South Europe (where they grow both in a natural habitat and as cultivars) and to fingerprint the flavonoid fraction present therein by means of HPLC/DAD. It was our intent to ultimately point out to these *Salvia* species, which might compete with *S. officinalis* in terms of the high levels of flavonoids.

To this effect, we spectrophotometrically determined an overall content of flavonoids in the sage extracts which were selectively obtained following the pharmacopeial procedure [5]. This preliminary assessment allowed selecting four out of twenty six sage species with the highest overall contents of flavonoids (which were *S. glutinosa*[a], *S. pratensis ssp. Haematodes*, *S. staminea*, and *S. triloba*). For these four sage species and additionally for *S. officinalis*, a comparison was carried out of their respective HPLC/DAD fingerprints. In our earlier study [6], an analogous comparison was performed of the total sage extracts derived with the use of methanol and, hence, containing a wider spectrum of the polar components. It was expected that the fingerprints

registered for the selectively derived flavonoid fraction could better characterize the sage species than the nonselectively derived methanol extracts.

However, the flavonoid contents in the sage species, as monitored through the prism of the sums of all the separated chromatographic peak areas had to be arranged in a different order and, in this case, the four out of twenty-six sage species with the highest overall contents of flavonoids were *S. nemorosa*, *S. forskahlei*, *S. azurea*, and *S. amplexicaulis*.

The source of the discrepancy between the two assessment approaches was discussed and, additionally, a chemometric comparison was performed by means of principal component analysis (PCA) of all the sage species considered in this study, which based on the HPLC/DAD chromatograms of the flavonoid fractions. Upon the obtained results, certain conclusions were drawn regarding seasonal differences in flavonoid composition among the individual plant species.

## 2. Experimental

*2.1. Herbal Material and Reagents.* Samples of the twenty-six different sage species (which are listed in Table 1 as species 1–5, 7–26, and 28) investigated in this study were collected in the Pharmacognosy Garden of the Medical University, Lublin, Poland, in three harvesting seasons (2007, 2008, and 2009). Botany specialists identified each investigated species, and the voucher specimens were deposited in the herbarium of the Department of Pharmacognosy, Medical University, Lublin, Poland. This plant material was dried for 40 h in an oven with a forced air flow at 35 to 40°C. The obtained dry material was stored in a refrigerator until commencement of the analysis. Species 6 was *S. glutinosa*[a], which originated from the natural habitat in the Ostrowsko region of south Poland and species 27 was *S. officinalis*, which originated from the natural habitat in the Zlatibor region of central Serbia. Species 6 was harvested in all three vegetation seasons (2007, 2008, and 2009), and species 27 was harvested in the summer of 2008 only and purchased dried. Summarized information about all the investigated herbal material is given in Table 1.

Methenamine was purchased from Pharma Cosmetics (Cracow, Poland), and methanol, ethyl acetate, acetone, glacial acetic acid, aluminium chloride, and hydrochloric acid used for the experiments were of analytical purity grade and purchased from POCh (Gliwice, Poland). Water was double distilled and deionized in the laboratory conditions by means of Elix Advantage model Millipore system (Molsheim, France).

*2.2. Selective Extraction of Flavonoids from Herbal Material.* Stock extract solution of each investigated sage species was prepared from 1 g medium powdered crude plant material. To this plant material, 20 mL acetone, 2 mL HCL (281 g L$^{-1}$), and 1 mL methenamine (5 g L$^{-1}$) were added. The entity was kept boiling on the water bath under the reflux for 30 min. Hydrolysate was filtered to the volumetric flask (100 mL). The separated plant material was extracted for the second and the third time with the 20 mL portions of

TABLE 1: Basic characteristics of the investigated plant material.

| Sample no. | Sage species | Harvesting year |
| --- | --- | --- |
| 1 | *S. amplexicaulis* | 2007, 2008, 2009 |
| 2 | *S. azurea* | 2007, 2008, 2009 |
| 3 | *S. cadmica* | 2007, 2008, 2009 |
| 4 | *S. deserta* | 2007, 2008, 2009 |
| 5 | *S. forskahlei* | 2007, 2008, 2009 |
| 6 | *S. glutinosa*[a] | 2007, 2008, 2009 |
| 7 | *S. hians* | 2007, 2008, 2009 |
| 8 | *S. jurisicii* | 2007, 2008, 2009 |
| 9 | *S. nemorosa* | 2007, 2008, 2009 |
| 10 | *S. pratensis ssp. Haematodes* | 2007, 2008, 2009 |
| 11 | *S. sclarea* | 2007, 2008, 2009 |
| 12 | *S. staminea* | 2007, 2008, 2009 |
| 13 | *S. stepposa* | 2007, 2008, 2009 |
| 14 | *S. tesquicola* | 2007, 2008, 2009 |
| 15 | *S. triloba* | 2007, 2008, 2009 |
| 16 | *S. verticillata* | 2007, 2008, 2009 |
| 17 | *S. officinalis* | 2007, 2008, 2009 |
| 18 | *S. lavandulifolia* | 2007, 2009 |
| 19 | *S. atropatana* | 2007 |
| 20 | *S. canariensis* | 2007 |
| 21 | *S. argentea* | 2008, 2009 |
| 22 | *S. austriaca* | 2008, 2009 |
| 23 | *S. nutans* | 2008, 2009 |
| 24 | *S. regeliana* | 2008, 2009 |
| 25 | *S. superba* | 2008, 2009 |
| 26 | *S. glutinosa* | 2008 |
| 27 | *S. officinalis*[b] | 2008 |
| 28 | *S. pratensis* | 2009 |

[a]Sample originating from the natural habitat in the Ostrowsko region of south Poland.
[b]Sample originating from the natural habitat in the Zlatibor region of central Serbia.

acetone kept boiling for 10 min. All extracts were filtered to the same volumetric flask, and acetone was added to make up to 100 mL. Then, 20 mL of the obtained solution was transferred to the separation funnel, 20 mL water was added, and the entity was extracted with ethyl acetate (firstly with the 15 mL, and then three times with the 10 mL portions of ethyl acetate). The separated organic layers were collected jointly and twice washed with the 40 mL portions of water. The organic layer was filtered to the 50 mL volumetric flask and filled up to the volume with ethyl acetate. Selective extraction of flavonoids from the investigated plant material was carried out in triplicate from the three different plant samples belonging to each individual batch of the sage species. The procedure of selective extraction is described in [5].

*2.3. Spectrophotometric Determination of the Overall Content of Flavonoids.* Flavonoids were determined spectrophotochemically according to the procedure described in [5],

using the selectively obtained flavonoid fraction extracts. Each result presented in this study is a mean value from the three independent spectrophotometric measurements obtained for each individual extract. An overall (%) content of flavonoids was recalculated for hyperoside, using the recalculation factor ($k$).

For each spectrophotometric measurement, two solutions were prepared. Solution 1 was prepared in the following way: to the 10 mL stock extract solution, 2 mL aluminium chloride ($20 \, g \, L^{-1}$) solution was added and filled up to 25 mL with the 1 : 19 mixture of acetic acid and methanol. Solution 2 was the reference sample, and it was prepared as follows: 10 mL stock solution was filled up to 25 mL with the 1 : 19 mixture of acetic acid and methanol. After 45 min from the preparation of these two solutions, absorbance of sample 1 was measured at the wavelength $\lambda = 425$ nm, using sample 2 as a blank. The percentage (%) content of total flavonoids ($X$) was calculated, using the following formula [5]:

$$X = \frac{A \cdot k}{m}, \tag{1}$$

where $A$ is the absorbance of the examined solution, $k$ is the recalculation factor for hyperoside ($k = 1.25$), and $m$ is the weight of the crude plant material (in grams).

The obtained quantitative results (i.e., the overall contents of flavonoids in the percentage scale) for all the investigated sage species are given in form of the bar diagram in Figure 1.

### 2.4. The HPLC/DAD Fingerprinting.

The HPLC analysis was carried out with a Varian 920-LC model liquid chromatograph (Harbor City, Calif, USA) equipped with a 900-LC model autosampler, gradient pump, 330 model DAD, and the Galaxie software for data acquisition and processing. The analyses were carried out in the gradient mode using a Pursuit C18 (5 $\mu$m particle size) column (250 × 4.6 mm id; Varian; Cat. no. 1215-9307). Methanol (A) and water (with a 1% ($v/v$) amount of glacial acetic acid; (B)) were used in the following A + B gradient program ($v/v$): from 0 to 17 min, 50 + 50; from 18 to 26 min, 80 + 20; from 27 to 46 min, 90 + 10; from 47 to 58 min, 100 + 0; from 59 to 70 min, 50 + 50. The flow rate was 0.6 mL min$^{-1}$. The chromatograms were registered at the wavelength $\lambda = 254$ nm.

### 2.5. Chemometric Baseline and Noise Correction of Chromatograms.

Dealing with the chromatograms of natural samples is not an easy task, although the fingerprinting approach has long been used for rapid screening of complex analytical signals. It has also been used in our earlier phytochemical study of the sage species [7]. In this section, a short description of the applied chemometric techniques is given, aiming to remove the background and noise from the HPLC/DAD chromatograms, and in that way to prepare the input data for the further exploration and visualization thereof with the use of the principal component analysis (PCA).

The first step was elimination of the background component from the chromatograms. One of numerous baseline

elimination techniques is the penalized asymmetric least squares approach (PALS) [8]. This method applies the least squares approach to fit a baseline to the signal. Each point of a signal gets a different weight which locates it above or below the original signal. The weights are modified according to an iterative procedure such that the points above the original signal have very small weights and the points below this signal have the weights close to 1. There are two adjustable parameters, that is, the order of the differences and the penalty parameter, which are to be optimized. Usually, the order of the differences is set as equal to 2. The larger the penalty parameter is, the smoother baseline is obtained.

Chromatographic signals were smoothed to suppress the white (Gaussian) noise. In this study, the Savitzky-Golay differentiation filter [9] was used. This filter helps to reduce the peak overlapping and the linear baseline drift by constructing the first and the second derivative spectra. The Savitzky-Golay filter technique resembles the local polynomial regression with a window of at least $f + 1$ points, where $f$ is the polynomial degree.

Also, the other undesired effects that could be present in the raw data were eliminated using the standard normal variate transformation (SNV) [10]. The pretreated data were used as the input data for principal component analysis (PCA) and discriminant partial least squares (DPLS).

### 2.6. Principal Component Analysis (PCA).

Principal component analysis (PCA) is the data exploration and visualization technique [11]. It allows to construct a set of new variables called principal components (PCs). The principal components are the orthogonal vectors that are linear combinations of the original variables and represent the data structure by maximizing the description of data variance. The PCA model consists of $k$ principal components, where $k$ is selected by the user. The original data matrix $\mathbf{X}(m \times n)$ is decomposed according to the following formula:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E}, \tag{2}$$

where $\mathbf{T}(m \times k)$ is the matrix of scores, $\mathbf{P}(n \times k)$ is the matrix of loadings, $\mathbf{E}(m \times n)$ is the residual matrix, and the superscript T denotes transposition of the matrix.

As a projection method, PCA enables projection of the objects or variables on the planes which they define [12]. Projection of the samples on the plane defined by the selected pairs of PCs allows studying the similarities among the samples (in form of the score plots). The loading plots are the projections of the variables on the planes of the selected principal components and allow tracing correlations among the data variables.

### 2.7. Discriminant Partial Least Squares (DPLS).

The discriminant partial least squares approach (DPLS) [13] is widely applied in chemistry, because of the multivariate character of the data studied and the high correlation usually observed among the explanatory variables. With the DPLS model, a linear relationship between a property of interest and a set of the explanatory variables is described. The property of interest is usually a binary or a bipolar coded vector. The
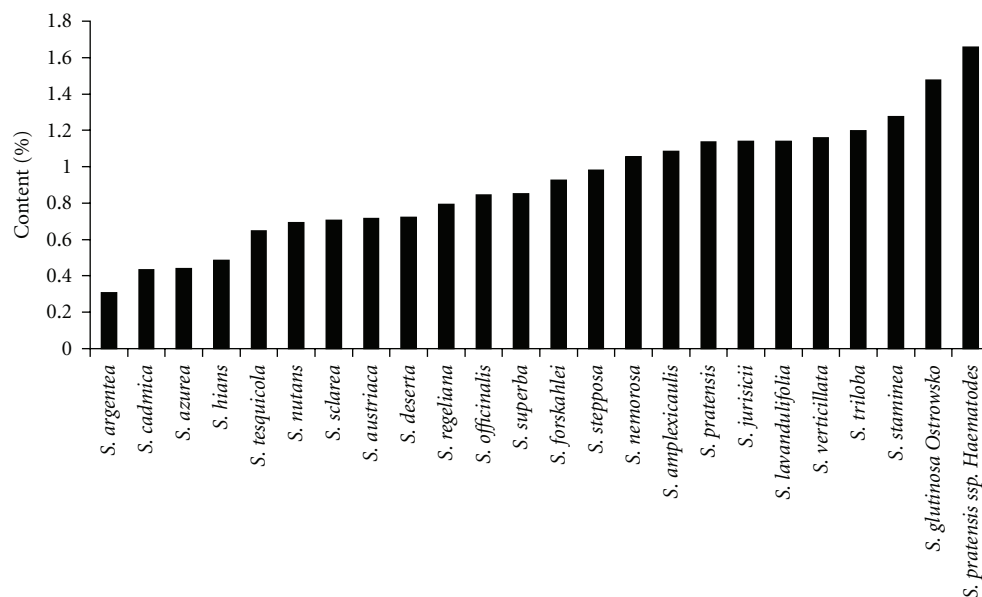
FIGURE 1: A bar diagram comparison of the overall contents of flavonoids in all the sage species harvested in 2009.

explanatory variables can be the sets of instrumental signals, for example, the chromatograms. Via the DPLS model, a set of a few orthogonal factors is constructed, aiming to maximize the covariance of the explanatory variables with the property of interest. When constructing the model, a number of orthogonal factors need to be estimated, which is usually done through the cross-validation mode. The final model is delivered in the form of the regression coefficients vector.

## 3. Results and Discussion

*3.1. The HPLC/DAD Fingerprinting and Spectrophotometric Results.* From a comparison of the HPLC/DAD fingerprints obtained for twenty-four selective extracts of flavonoids derived from the different sage species harvested in 2009, it was observed that the chromatogram of *S. nemorosa* shows both the highest number of the fourteen separated peaks and the highest sum of the separated peak areas (1244 mAV × min). The same species showed one of the highest numerical values of the sum of the separated peak heights (1893.5 mAV).

To the contrary, the chromatogram of the flavonoid extract from *S. cadmica* is characterized with the lowest number of the four separated peaks. Accordingly, the sums of the separated peak heights and the sums of the separated peak areas for this particular sage species were obtained among the lowest numerical levels (630.4 mAV × min and 1507.7 mAV, resp.).

On the basis of the chromatographic results, the following sage species are those with the highest sums of the separated peak areas and/or the highest sums of the separated peak heights: *S. amplexicaulis*, *S. azurea*, *S. forskahlei*, *S. hians,* and *S. nemorosa*. Sage species showing the lowest sums of the separated peak areas and/or the lowest sums of the separated peak areas are the following ones: *S. cadmica*,

*S. nutans*, *S. officinalis*, *S. regeliana,* and *S. triloba*. Numbers of the separated chromatographic peaks, sums of the separated peak areas, and sums of the separated peak heights for all the investigated sage species are given in Table 2. For the sake of graphical illustration, in Figure 2(a), we presented selected chromatographic fingerprints of the four sage species with the highest sums of the separated peak areas and also the fingerprint of *S. officinalis*. *Salvia officinalis* was compared with these four species, due to its unique position in the traditional European medicine, in spite of the lowest overall percentage content of flavonoids among the five compared species. The chromatogram of the *S. officinalis* extract fully confirmed the spectrophotometric findings. Its fingerprint is characterized with a relatively low number of the six separated peaks, and with the relatively low sums of the separated peak areas and the peak heights (676.1 mAV × min and 1377.4 mAV, resp.).

According to the spectrophotometric results (Figure 1), the following sage species: *S. glutinosa*[a], *S. pratensis ssp. Haematodes*, *S. staminea*, and *S. triloba* were characterized with the highest overall percentage contents of flavonoids, as recalculated to hyperoside.

*Salvia triloba* is one of the four sage species showing the higher overall percentage contents of flavonoids, as established spectrophotometrically. However, on the chromatogram of this species, we found seven separated peaks only, and the sums of their areas and heights are not very impressive either (although higher than with *S. officinalis*).

The second out of the four is *S. staminea*, which according to the spectrophotometric result is richer in flavonoids than *S. triloba* and *S. officinalis*. The chromatographic results confirmed the spectrophotometric ones. On the chromatogram of the flavonoid fraction derived from *S. staminea*, we found nine separated peaks, and the sums of their areas and heights are the highest ones among the selected four.

TABLE 2: A comparison of the numbers of the separated chromatographic peaks, and of the sums of the separated peak heights and peak areas with twenty-four different sage species harvested in 2009. Numbering of the sage samples is in conformity with Table 1.

| Sample no. | Sage species | No. of separated peaks | Sum of separated peak heights (mAV) | Sum of separated peak areas (mAV × min) |
|---|---|---|---|---|
| 1 | S. amplexicaulis | 7 | 1906.3 | 1038.2 |
| 2 | S. azurea | 6 | 1856.3 | 1168.3 |
| 3 | S. cadmica | 4 | 1507.7 | 630.4 |
| 4 | S. deserta | 6 | 1666.3 | 740.8 |
| 5 | S. forskahlei | 8 | 2041.6 | 1238.3 |
| 6 | S. glutinosa[a] | 8 | 1630.3 | 895.8 |
| 7 | S. hians | 11 | 1951.7 | 1009.0 |
| 8 | S. jurisicii | 13 | 1643.7 | 847.0 |
| 9 | S. nemorosa | 14 | 1893.5 | 1244.0 |
| 10 | S. pratensis ssp. Haematodes | 12 | 1635.5 | 903.4 |
| 11 | S. sclarea | 5 | 1610.7 | 852.3 |
| 12 | S. staminea | 9 | 1980.8 | 948.0 |
| 13 | S. stepposa | 9 | 1646.1 | 632.3 |
| 14 | S. tesquicola | 9 | 1665.8 | 667.4 |
| 15 | S. triloba | 7 | 1586.4 | 669.4 |
| 16 | S. verticillata | 7 | 1711.7 | 878.8 |
| 17 | S. officinalis | 6 | 1377.4 | 676.1 |
| 18 | S. lavandulifolia | 9 | 1614.3 | 837.4 |
| 21 | S. argentea | 9 | 1634.1 | 981.5 |
| 22 | S. austriaca | 8 | 1561.0 | 930.0 |
| 23 | S. nutans | 7 | 1464.7 | 609.4 |
| 24 | S. regeliana | 7 | 1487.9 | 806.6 |
| 25 | S. superba | 6 | 1993.5 | 886.4 |
| 28 | S. pratensis | 9 | 1600.7 | 927.6 |

The third species (*S. glutinosa*[a]) showed the highest overall content of flavonoids among the four species. On its fingerprint chromatogram, eight separated peaks can be found, and the sums of their areas and heights are 895.8 mAV × min and 1630.3 mAV, respectively. These values are lower than those chromatographically obtained for *S. staminea*, yet higher than with *S. officinalis* and *S. triloba*.

The fourth spectrophotometrically selected species is *S. pratensis ssp. Haematodes* as that with the relatively highest percentage content of flavonoids. On its chromatogram, the highest number of the twelve separated peaks was observed among the four compared species (and *S. officinalis*). Sums of their areas and heights were, respectively, 903.4 mAV × min and 1635.5 mAV. For the sake of graphical illustration, in Figure 2(b), we presented selected chromatographic fingerprints of the four sage species with the highest overall percentage contents of flavonoids, as spectrophotometrically assessed and recalculated to hyperoside, and also the fingerprint of *S. officinalis*.

The perceptible discrepancy between the spectrophotometric and the chromatographic results (which can anyway be considered as semiquantitative only) is due to the different principles and also different sensitivities of the two analytical approaches. Spectrophotometric analysis assumes a very simplifying recalculation of the overall flavonoid contents to hyperoside. On the other hand, the chromatographic fingerprinting is certainly more sensitive, although in spite of the selective and flavonoids-oriented extraction, one cannot exclude the presence of the compounds other than flavonoids in the chromatographed extracts, which might result in a different source of the estimation error. This is the reason why these two approaches have been presented and compared in this study.

3.2. Chemometric Evaluation of Chromatographic Fingerprints. In this study, a set of herbal fingerprints obtained from HPLC/DAD for the *Salvia* species was analyzed with the use of the chemometric techniques. Firstly, we enhanced the signal-to-noise ratio. To this effect, the background removal was carried out by application of the PALS method. Also, the noise influence was reduced with the use of the Savitzky-Golay smoothing filter, which delivered smoothed signals. For all the assessed fingerprints, the penalty parameter used in the PALS method was set to $10^7$ and the Savitzky-Golay standard Matlab command of 51 frame size was applied. In Figure 3, we showed the baselines of the chromatograms and
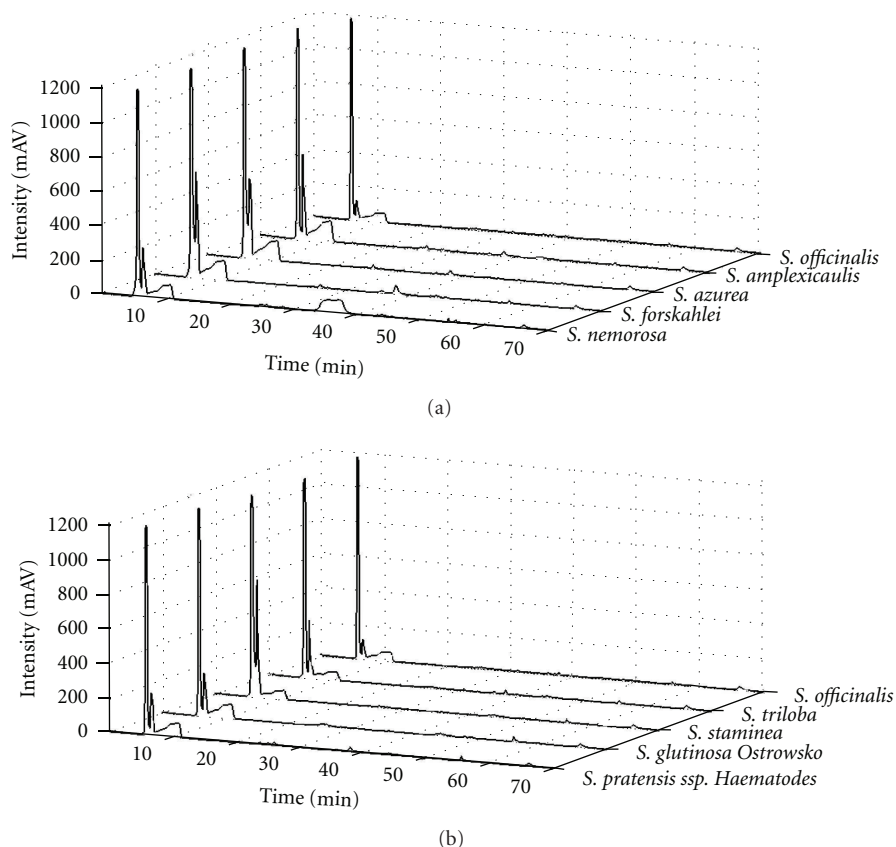
(a)



(b)

FIGURE 2: (a) A comparison of the HPLC/DAD fingerprints for the five different sage species (*S. nemorosa*, *S. forskahlei*, *S. azurea*, *S. amplexicaulis,* and *S. officinalis*). The chromatograms were registered at the wavelength $\lambda$ = 254 nm. All fingerprints except for *S. officinalis* are those with the highest sums of the separated chromatographic peak areas. (b) A comparison of the HPLC/DAD fingerprints for the five different sage species (*S. glutinosa*. *S. pratensis ssp. Haematodes*. *S. staminea*. *S. triloba*., and *S. officinalis*). The chromatograms were registered at the wavelength $\lambda$ = 254 nm. All fingerprints except for *S. officinalis* are those with the highest overall sums of flavonoids, as spectrophotometrically established.

also the signals after performance of the preprocessing step. Finally, the SNV transformation was applied to the signals which were analyzed further in that form.

On the score plots shown in Figure 4, three groups of the *Salvia* samples can be distinguished. One can easily notice that the sage samples collected in the 2007 vegetation season markedly differ from the remaining ones in the space of PC1, which describes nearly 70% of data variance (Figure 5). Real cause of this difference remains unknown, yet it can be due to the local weather changes and/or due to the relatively longest storage period with the plant samples collected in 2007.

We tried to distinguish samples originating from the 2007 vegetation season from the remaining ones in another way also, which was achieved by application of the discriminant partial least squares model (DPLS) [13]. All samples were split into the two sets using the Kennard and Stone approach [14], namely into the training and the test set. With this algorithm, all kinds of samples are included in the training set, what provides the representativeness of the training set. The training set consisted of 17 samples from each of the two classes (class 1 was valid for the 2007 vegetation season samples and class 2 for the 2008 and 2009 samples), and the test set contained all the remaining samples. Due to a rather limited number of the available samples, the cross-validation leave-3-out method was used to estimate the complexity of the DPLS model and the eight latent factors were chosen (Figure 6). The correct classification rate (CCR) was used as the model characterization parameter, and the sensitivity and specificity parameters were calculated (Table 3). The CCR for the model set characterizes the fitting of the model to the data, and for the test set it describes predictive power of the model. Finally, we obtained the 84.38% correctly classified samples from the independent test set, which was a reasonably satisfying result (additionally confirming correctness of distinguishing the 2007 sage samples from the remaining specimens studied). All calculations and chemometric treatment were applied via the R2010a Matlab by the MathWorks and its toolboxes.

## 4. Conclusions

Upon the spectrophotometric results, we compared twenty-four different sage species harvested in 2009 in terms of the overall percentage contents of flavonoids (recalculated
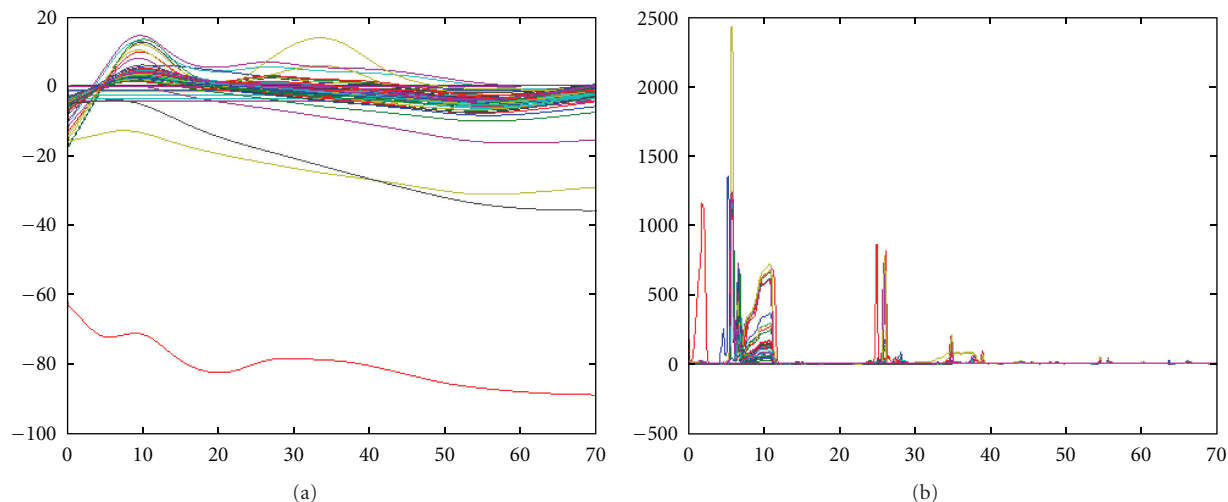
FIGURE 3: The HPLC/DAD profiles of the different *Salvia* species extracts: (a) with the baseline and (b) after the baseline and noise removal.
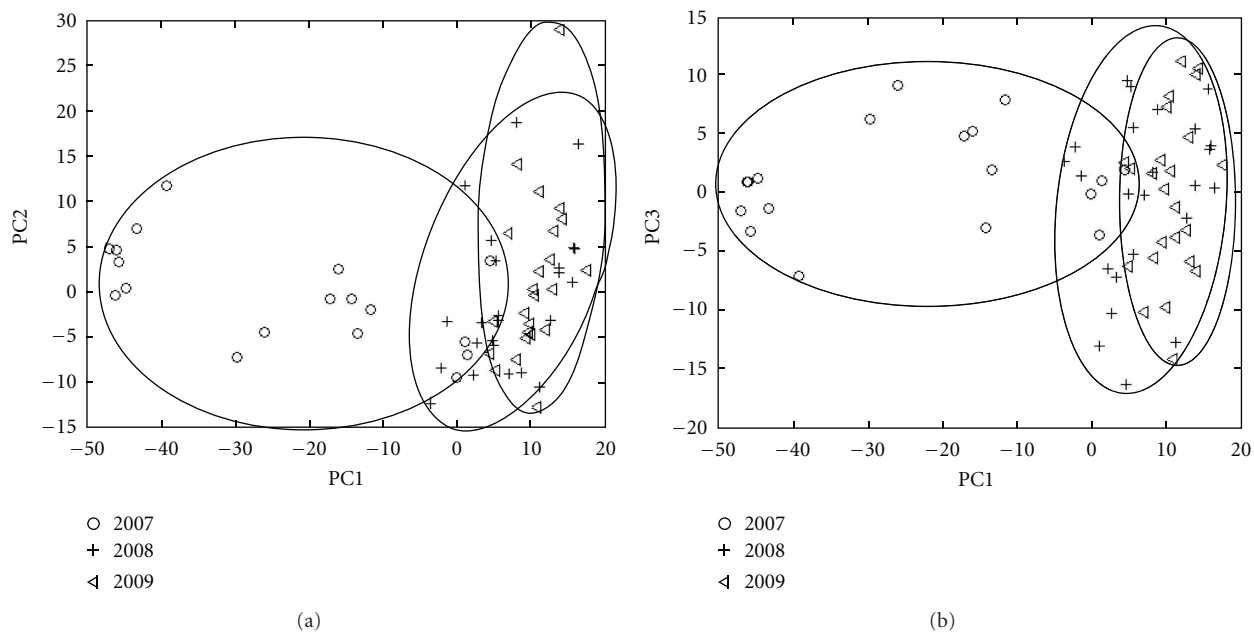


FIGURE 4: Plots of the *Salvia* samples on the plane determined by (a) the first and the second principal component and (b) the first and the third principal component.

to the contents of hyperoside) and, on this basis, we selected those showing the highest overall percentage contents (i.e., *S. glutinosa*[a], *S. pratensis ssp. Haematodes*, *S. staminea*, and *S. triloba*).

From the HPLC/DAD comparison of the fingerprints valid for the same twenty-four different sage species and from the comparison of the chromatograms with the highest sums of the chromatographic peak areas, the following sage species: *S. nemorosa*, *S. forskahlei*, *S. azurea*, and *S. amplexicaulis* could be selected as those with the highest overall contents of flavonoids.

A comparison of the spectrophotometric data with the results of the chromatographic fingerprinting allowed for

TABLE 3: Model parameters. where SE is sensitivity, SP is specificity, and CCR is the correct classification rate for the training and the test set, respectively.

|  | Training set | Test set |
| --- | --- | --- |
| SE | 94.18 | 88.00 |
| SP | 100 | 71.43 |
| CCR | 97.06 | 84.38 |

a conclusion as to the perceptible discrepancy between the results of these two approaches. This discrepancy apparently is due to the different principles and also different
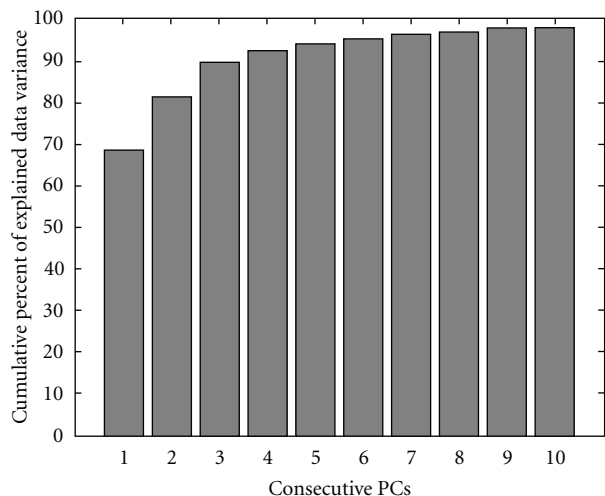
FIGURE 5: The cumulative percent of the explained data variance by the consecutive principal components.
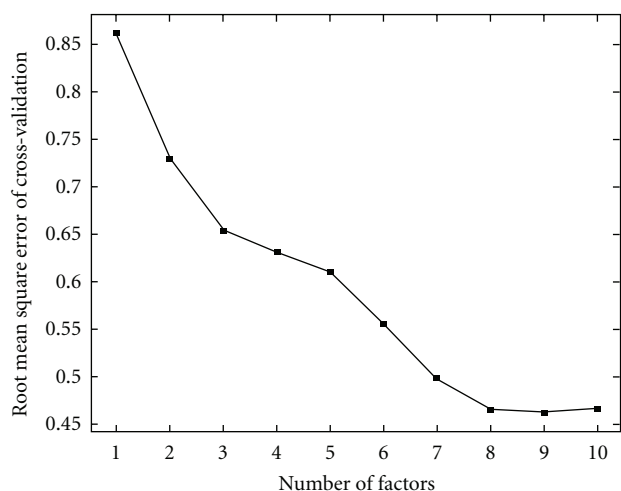


FIGURE 6: The cross-validation error for estimation of the model complexity.

sensitivities of each analytical technique applied. Due to completely different error sources in each approach and also to an unknown chemical composition of the fingerprinted extracts, for the time being, it seems noteworthy to pay roughly equal attention to the two series of the obtained results.

In spite of the differences in the spectrophotometric and chromatographic results, it can be concluded that, in terms of the flavonoid fraction contents, many individual sage species outperform *S. officinalis*.

The chromatographic fingerprints of the selectively derived flavonoid extracts proved useful for the construction of the chemometric models. In this study, they proved helpful in differentiating among the harvesting years with the investigated *Sage* species.

## References

[1] J. Rzepa, Ł. Wojtal, D. Staszek et al., "Fingerprint of selected *Salvia* species by HS-GC-MS analysis of their volatile fraction," *Journal of Chromatographic Science*, vol. 47, no. 7, pp. 575–580, 2009.

[2] Ł. Ciesla, M. Hajnos, D. Staszek, Ł. Wojtal, T. Kowalska, and M. Waksmundzka-Hajnos, "Validated binary high-performance thin-layer chromatographic fingerprints of polyphenolics for distinguishing different *Salvia* species," *Journal of Chromatographic Science*, vol. 48, no. 6, pp. 421–427, 2010.

[3] Ł. Cieśla, D. Staszek, M. Hajnos, T. Kowalska, and M. Waksmundzka-Hajnos, "Development of chromatographic and free radical scavenging activity fingerprints by thin-layer chromatography for selected *Salvia* species," *Phytochemical Analysis*, vol. 22, no. 1, pp. 59–65, 2011.

[4] M. Sajewicz, D. Staszek, M. Natić, Ł. Wojtal, M. Waksmundzka-Hajnos, and T. Kowalska, "TLC-MS versus TLC-LC-MS fingerprints of herbal extracts. Part II. Phenolic acids and flavonoids," *Journal of Liquid Chromatography and Related Technologies*, vol. 34, no. 10-11, pp. 864–887, 2011.

[5] Polish Pharmaceutical Society, *Polish Pharmacopoeia VI*, Polish Pharmaceutical Society, Warsaw, Poland, 2002.

[6] M. Sajewicz, D. Staszek, Ł. Wojtal, T. Kowalska, M. Hajnos, and M. Waksmundzka-Hajnos, "Binary HPLC-Diode Array Detector and HPLC-Evaporative Light Scattering Detector fingerprints of methanol extracts from the selected sage (*Salvia*) species," *Journal of AOAC International*, vol. 94, pp. 71–76, 2011.

[7] M. Daszykowski, M. Sajewicz, J. Rzepa et al., "Comparative analysis of the chromatographic fingerprints of twenty different sage (*Salvia* L.) species," *Acta Chromatographica*, vol. 21, no. 4, pp. 513–530, 2009.

[8] P. H. C. Eilers, "A perfect smoother," *Analytical Chemistry*, vol. 75, no. 14, pp. 3631–3636, 2003.

[9] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

[10] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Applied Spectroscopy*, vol. 43, no. 5, pp. 772–777, 1989.

[11] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, 1987.

[12] M. Daszykowski, B. Walczak, and D. L. Massart, "Projection methods in chemistry," *Chemometrics and Intelligent Laboratory Systems*, vol. 65, no. 1, pp. 97–112, 2003.

[13] D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. de Jong, P. J. Lewi, and J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, The Netherlands, 1997.

[14] R. W. Kennard and L. A. Stone, "Computer aided design of experiments," *Technometrics*, vol. 11, pp. 137–148, 1969.