Classification of herbs by FT-NIR Spectroscopy using the Thermo Scientific Antaris II Analyzer

Authors

Martin Hollein, Nicolet CZ s.r.o.
Prague, Czech Republic

Todd Strother
Thermo Fisher Scientific
Madison, WI, USA

Key Words

Antaris, TQ Analyst, closed sample cup, cosmetics, discriminant analysis, FT-NIR, herbs, pharmaceuticals, sample cup spinner

Introduction

Vibrational techniques like Fourier transform near-infrared (FT-NIR) spectroscopy are well-suited for raw material identification. This is because FT-NIR is sensitive to the characteristic molecular vibrations that occur for specific compounds. Because of this ability to detect molecular vibrations, NIR spectrometers are widely used in the pharmaceutical industry for classifying pure chemicals or their mixtures. Compounds usually contain highly absorbing functional groups (alkyls, phenyls, amines, thiols, hydroxyls, acids, esters, etc.) which allows for unambiguous identification within seconds. The Thermo Scientific™ Antaris™ line of FT-NIR spectrometers (Figure 1) is ideally suited for raw material identification of pure chemicals as well as complex biological samples used in pharmaceutical, nutraceutical and related industries.

Figure 1: The Antaris II FT-NIR analyzer showing different available detection options

While herbs and their extracts are often used in cosmetics and nutraceuticals, they additionally find use as raw materials in classic pharmaceutical industries. Often, these herbal active ingredients or their intermediates are isolated from raw materials using extraction or chromatographic techniques. Herbs used in pharmaceuticals must be treated like any another raw material in the manufacturing facility and must fulfill all of the regulatory demands as any other "chemical" raw material.

Researchers at IREL, spol. s.r.o. (Brno, Czech Republic) explored FT-NIR spectroscopy as a means to rapidly identify herbs used to produce extracts found in nutraceuticals, cosmetics and pharmaceuticals. Their biological origin means that many herbs contain similar components, like cellulose, proteins and sugars, which tend to result in similar spectra. The purpose of this application note is to demonstrate that complex herbal materials can be correctly classified using FT-NIR spectroscopy much more rapidly than conventional techniques.

Thermo Fisher
SCIENTIFIC

Experimental

The standard herbs used to develop the analysis method were supplied by IREL, spol s.r.o. and used directly to produce the fluid extracts – no additional grinding or sieving was performed. Particle size varied substantially among the samples. Leaves were typically thin and as large as 6 x 6 mm. Roots, galls and orange peels were crushed to maximum particle sizes of 10 x 10 x 5 mm. The list and description of the 14 herbs used in the data collection is summarized in Table 1.

Herbal Name	Scientific Names
Agrimony – top	Agrimonia eupatoria L.
Buckbean – leaf	Menyanthes trifoliata L.
Calamus – root	Acorus calamus L.
Camomile – flowers	Matricaria recutita
Gentian – root	Gentiana lutea L.
Hot pepper – powdered fruit	Capsicum sp.
Myrrh – powdered gum resin	Commiphora molmol E.
Oak Apple – crushed gall	Diplolepis quercus-folii
Orange Peel	Citrus Aurantium L.
Sage – top	Salvia officinalis L.
Valerian – root	Valeriana officinalis L.
Walnut – leaf	Juglans regia L.
Witch Hazel – leaf	Hamamelis virginiana L.
Wormwood – top	Artemisia absinthium L.

Table 1: List of herbal samples used in analysis

Three or four samples of each herb were scanned with an Antaris II FT-NIR system in the range between 10,000 and 4,000 cm^{-1}. The materials were placed in a closed rotating sample cup with a 30 mm window over the integrating sphere (Figure 2). Spectra and other calibrations data were collected and archived using the validated Thermo Scientific™ RESULT software. Each sample spectra was the result of 50 scans using 4 cm^{-1} resolution. These parameters allowed the material to be scanned through two full rotations of the sample cup in less than one minute.

Figure 2: Antaris II analyzer with the closed sample cup and spinner

The resulting spectra were evaluated with Thermo Scientific™ TQ Analyst software using a Discriminant Analysis algorithm in order to classify the samples. Multiplicative Signal Correction (MSC) pathlength type was chosen because it is designed for analysis of samples, such as solids, where it is difficult to obtain an independent pathlength measurement. Because of light scattering, pathlength is then treated as a multiplicative contribution to the spectral signal. The software applies the same function to the standards and the unknown samples. The raw spectra in the range between 9,900 and 4,100 cm^{-1} were used for the calibration algorithm. A linear removed baseline correction was used but no other mathematical smoothing or derivative functions were applied.
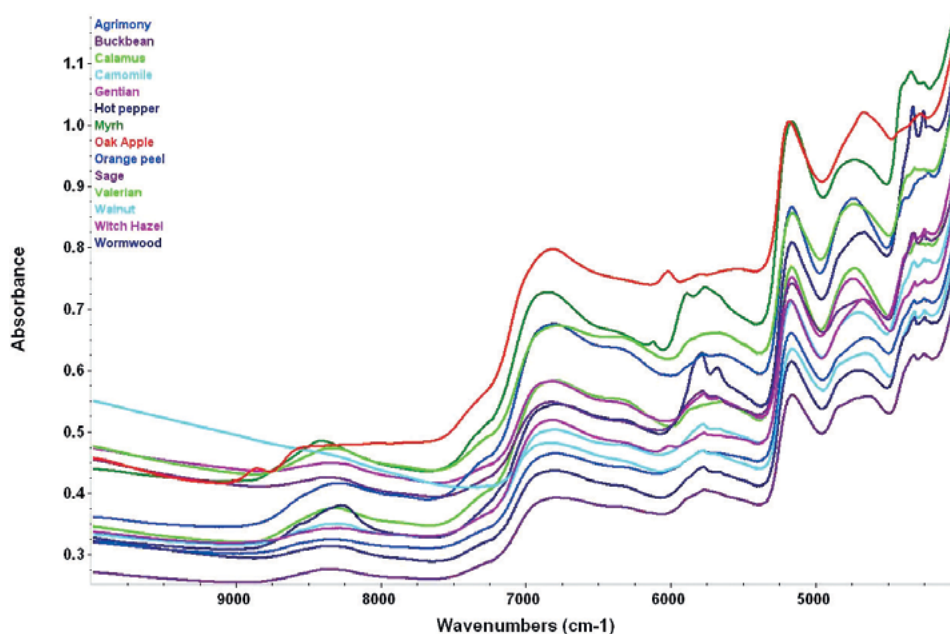

Figure 3: Representative spectra of the 14 samples used in the analysis.

Results and discussion

Figure 3 is the plot of average diffuse reflection spectra of the 14 herbs. The samples show substantial spectral variation that can be exploited for building the classification model. Using Discriminant Analysis as the chemometric method 99.5% of the spectral variability was described with five principal components. Principal components are orthogonal vectors that describe spectral variation in a set of standards. The first principal component will account for most of the spectral variation and subsequent components account for the remaining variation. Generally, the fewer components required to describe the total variation results in a more robust method.

This method correctly classified all of the standards as well as randomly chosen validation spectra from each class. One means of determining the quality of a chemometric method is through Mahalanobis distance analysis. The Mahalanobis distance can be described as the spectral distance a particular sample is from the "center of mass" of the group. For samples that are spectrally similar to a particular class or group, the Mahalanobis distances will be

relatively small. Samples that are spectrally very different than the class will have high Mahalanobis distances. For the samples under consideration, the Mahalanobis distances of the nearest incorrect class were at least twice the distance to the correct class; indicating good spectral separation and accurate classification. The distances of the nearest and next nearest classes are shown in Table 2.

Principal Component Scores plots show spectra variation for samples in two dimensions. Figure 4 is a Principal Component Scores plot that shows clusters of different classes described by the first two principal components. Most of the classes are tightly clustered with standards that are near each other. Extremely homogeneous samples were most tightly clustered (i.e., powdered hot pepper). Those that have dispersed clusters are classes with irregular shaped particles or heterogeneous samples (i.e., walnut leaf, oak apple-crushed gall). The Principal Component Scores plot shows no overlap between the different classes when viewed in multidimensional space indicating that the samples can be successfully and easily classified.

Class ID	Distance	Next Distance	Next Class
Agrimony	0.5	4.0	Camomile
Agrimony	0.7	3.6	Camomile
Agrimony	1.0	3.5	Buckbean
Buckbean	0.2	3.3	Camomile
Buckbean	0.2	3.3	Camomile
Buckbean	0.4	3.1	Camomile
Calamus	0.7	2.8	Orange Peel
Calamus	0.7	3.7	Orange Peel
Calamus	0.4	3.0	Orange Peel
Camomile	1.0	2.1	Wormwood
Camomile	0.9	2.1	Wormwood
Camomile	0.6	2.2	Wormwood
Gentian	0.2	5.0	Calamust
Gentian	0.4	5.0	Calamust
Gentian	0.4	4.8	Calamust
Hot Pepper	0.1	13.2	Wormwood
Hot Pepper	0.1	13.3	Wormwood
Hot Pepper	0.1	13.2	Wormwood
Myrrh	0.3	7.4	Sage
Myrrh	0.4	7.3	Sage
Myrrh	0.5	7.3	Wormwood
Oak Apple	0.5	5.5	Witch Hazel
Oak Apple	0.9	5.0	Witch Hazel

Class ID	Distance	Next Distance	Next Class
Oak Apple	0.8	6.2	Witch Hazel
Oak Apple	0.9	6.4	Witch Hazel
Orange Peel	1.3	3.1	Calamus
Orange Peel	0.8	2.8	Calamus
Orange Peel	0.7	3.6	Calamus
Orange Peel	1.2	3.6	Calamus
Sage	1.0	4.6	Wormwood
Sage	1.0	6.2	Wormwood
Sage	0.8	5.0	Wormwood
Valerian	1.1	7.0	Gentian
Valerian	0.9	6.6	Gentian
Valerian	0.7	6.6	Gentian
Walnut	1.6	10.6	Agrimony
Walnut	2.0	13.0	Agrimony
Walnut	1.1	10.2	Agrimony
Walnut	1.7	11.3	Agrimony
Witch Hazel	0.8	5.9	Agrimony
Witch Hazel	0.9	5.7	Oak-Apple
Witch Hazel	0.6	5.5	Oak-Apple
Wormwood	0.6	2.2	Camomile
Wormwood	0.7	1.8	Camomile
Wormwood	0.2	2.1	Camomile

Table 2: Chart showing the Mahalonobis distance of each sample to its identified class. The distance to the next nearest class is included as a quantitative indication of how close the samples are to an incorrect class. Lower numbers indicate spectral similarity to a class, higher numbers indicate spectral dissimilarity. All of the samples were correctly identified.
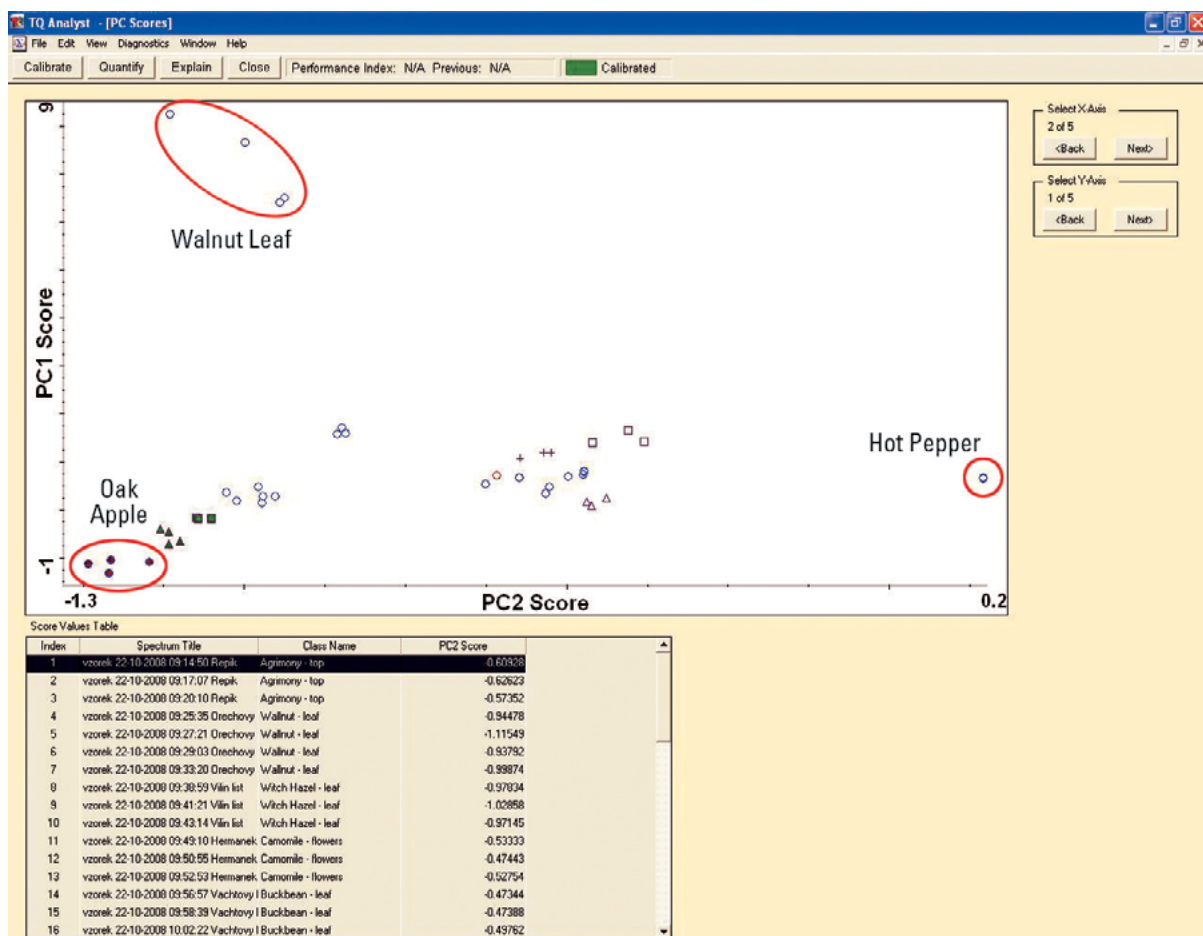
Figure 4: Principal components scores plots showing the class clustering as a function of the first and second principal component. Heterogeneous samples tended to have widely scattered clusters (Walnut Leaf, Oak Apple), while homogeneous samples were tightly clustered (Hot Pepper). There were well defined and no overlapping clusters for the classes when viewed in 5 dimensional principal component space.

Conclusion

The Antaris II FT-NIR analyzer offers an excellent alternative to traditional identification and classification methods (i.e., descriptive morphology, TLC and HPLC chromatography) of herbal extracts used in nutraceuticals, cosmetics and pharmaceuticals. There is no need for additional grinding, cutting, sieving or other preparation of the samples. This results in a great reduction in time and labor required to properly classify incoming raw materials used in these industries. Additionally, the sample cup spinner accessory with the closed sample cups allows for the reproducible collection of spectra from heterogeneous samples. The Discriminant Analysis method developed here for analysis of the 14 herbs is shown to be robust and reliable while providing answers within seconds.

Find out more at **www.thermofisher.com**

Thermo Fisher
SCIENTIFIC