Journal of Zhejiang University SCIENCE B
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Application of principal component-radial basis function neural networks (PC-RBFNN) for the detection of water-adulterated bayberry juice by near-infrared spectroscopy*

Li-juan XIE, Xing-qian YE, Dong-hong LIU, Yi-bin YING†‡

(*College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China*)

†E-mail: ybying@zju.edu.cn

Received Feb. 24, 2008; revision accepted June 21, 2008; CrossCheck deposited Oct. 28, 2008

Abstract: Near-infrared (NIR) spectroscopy combined with chemometrics techniques was used to classify the pure bayberry juice and the one adulterated with 10% (w/w) and 20% (w/w) water. Principal component analysis (PCA) was applied to reduce the dimensions of spectral data, give information regarding a potential capability of separation of objects, and provide principal component (PC) scores for radial basis function neural networks (RBFNN). RBFNN was used to detect bayberry juice adulterant. Multiplicative scatter correction (MSC) and standard normal variate (SNV) transformation were used to preprocess spectra. The results demonstrate that PC-RBFNN with optimum parameters can separate pure bayberry juice samples from water-adulterated bayberry at a recognition rate of 97.62%, but cannot clearly detect water levels in the adulterated bayberry juice. We conclude that NIR technology can be successfully applied to detect water-adulterated bayberry juice.

Key words: Near-infrared (NIR) spectroscopy, Principal component-radial basis function neural networks (PC-RBFNN), Bayberry juice, Adulteration, Chemometrics technique

doi:10.1631/jzus.B0820057 **Document code:** A **CLC number:** O43

INTRODUCTION

The adulteration of food products is an issue, drawing an increasing attention of researchers and consumers. Economic adulteration involves the substitution of less costly or lower quality ingredients for those declared on the label. The most common adulteration methods for fruit juice include dilution with water, addition of sugars, addition of inferior products to concentrated pressed juice, or addition of the less expensive fruit juice (Nagy, 1997).

Bayberry (*Myrica rubra* Sieb. et Zucc.) is cultivated in southeastern China for more than 2000 years (Chen, 1996). The special sweet with moderate acid and exquisite flavor in its abundant juice makes bayberry taste delicious. In China, it has been traditionally used to deal with gastric intestinal problems, such as diarrhea and gastroenteritis (Fang *et al.*, 2007). Bayberry is also processed to produce sweets, jam, juice, syrup, and wine. Water and carbohydrates such as glucose, sucrose, and fructose with average concentrations of 12%~13% (w/w), are the main components of bayberry juice. Due to these chemical compositions, the main issues regarding authenticity of bayberry products are addition of water and sugar.

The main strategy for detecting adulteration in bayberry juice is to analyze and compare the chemical compositions. There are several methods reported in the literature for determining components in bayberry, including determination of phemolic compounds by high performance liquid chromatography (HPLC) coupled to photodiode array detection

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 60778024 and 30825027) and the National Basic Research Program (973) of China (No. 2006BAD11A12)

and electrospray ionization mass spectrometry (Fang *et al*., 2007). Although these chemical methods provide valuable information, they are tedious, complicated, time-consuming, and destructive, thus underlining a need for better techniques to quickly and reliably inspect out the adulterated bayberry juice.

Near-infrared (NIR) spectroscopy has become a rapid and well-established technique for the quantitative and qualitative analyses of agricultural products, and has been widely applied in the field of food chemistry. It can record the response of the molecular bonds of chemical constituents to the NIR spectrum (e.g., C-H, N-H and O-H bands which are the primary structural components of organic molecules) and thereby build a characteristic spectrum that behaves as a fingerprint of the sample (Cozzolino *et al*., 2004; Liu *et al*., 2006). This method can perform reliable quantization in small amounts of samples (Gautz *et al*., 2006). Compared with classical chemical and physical methods, the NIR spectroscopy is nondestructive, adaptable for almost all kinds of samples at low cost without any pretreatment, fast and simple to apply, requiring minimal testing samples, and easy to be automated (Chen *et al*., 2004; Gestal *et al*., 2004; Fu *et al*., 2008).

The absorption peaks of NIR spectra are broad and overlapping, making visual examination of the NIR spectra miss discrimination between authentic and adulterated products (Liu *et al*., 2006). Useful multivariate data analysis tools such as principal component analysis (PCA), linear discriminant analysis (LDA), soft independent modeling of class analogy (SIMCA), artificial neural networks (ANN), partial least-squares discriminant analysis (PLSDA), and so on, open the possibility to unravel and interpret the optical properties of samples in question and allow a classification without the need to input chemical information beforehand (Liu *et al*., 2006). Several studies have reported the use of NIR spectroscopy to detect adulteration in products, such as juice, honey, oil, meat, coffee, hamburger, seafood, and so on (Gayo *et al*., 2006).

The main objective of this study was to investigate the usefulness of NIR spectroscopy to differentiate the authenticity of pure bayberry juice from the one adulterated with distilled water [10% (w/w) and 20% (w/w), respectively]. The capabilities of different chemometrics techniques were also studied.

MATERIALS AND METHODS

Samples

A total of 43 bayberry samples at different ripe stages from 14 varieties and 9 different areas in Zhejiang Province, China, were used in this research. They were cultivated in June to July 2006 from the following varieties: 'Biqi' (Cixi region, 4 maturities; Xianju region, 4 maturities; Wenzhou region, 1 maturity), 'Dongkui' (Xianju region, 4 maturities; Xiangshan region, 4 maturities), 'Wandao' (Ninghai region, 1 maturity), 'Lizhizhong' (Hangzhou region, 1 maturity), 'Tanmei' (Hangzhou and Ninghai regions, each with 1 maturity), 'Wuzi' (Xiangshan region, 4 maturities), 'Dingqiao' (Wenzhou region, 1 maturity), 'Zaoda' (Linhai region, 1 maturity), 'Chida' (Cixi region, 1 maturity), 'Fenhong' (Shangyu region, 1 maturity), 'Shuijin' (Shangyu region, 1 maturity), 'Baiyangmei' (Ninghai region, 1 maturity), 'Zaose' (Xiaoshan region, 1 maturity) and 'Chise' (Xiaoshan region, 1 maturity). Some varieties, such as Biqi, had three sizes.

The bayberry samples were pitted, and the juice was extracted from the fruits using a centrifugal juicer without filtering and was stored frozen (−60 °C) in screw-capped glass vials to ensure that the bayberry juice quality was maintained until February, 2007. The samples were kept overnight in a cold dark room for defrosting and then in the lab at room temperature (20~25 °C) to equilibrate before spectra collection. The juice samples were filtered in order to separate the dispersed solid particles and were placed in the same temperature-controlled room where the spectrometer was located before performing the analysis. The samples added with 10% (w/w) and 20% (w/w) distilled water were prepared at the time of spectral collection. A total of 129 pure and adulterated juices were thus prepared for analysis.

Spectral measurement

Bayberry juices were measured in transmission mode using a Nexus FT-NIR spectrometer (Thermo Nicolet Corporation, Madison, WI, USA) equipped with an interferometer, a cooled InGaAs detector, and a wide-band quartz halogen light source (Quartz Tungsten Halogen, 50 W). Samples were scanned in an optical path-length rectangular quartz cuvette with air as reference at room temperature. The mirror

velocity was 0.9494 cm/s, and the resolution was 1 cm^{-1}. For each sample, the recorded spectra consisted of the average of 64 scans. The total number of data points was 18046 for each spectrum. Each sample was scanned twice. The replicates per sample were averaged to obtain a new dataset. The background spectrum was acquired using the same instrumental conditions as those employed for samples. Because of the high background stability, it was considered unnecessary to scan a new background spectrum after the measurement of each sample. The rectangular quartz cuvette was cleaned after each sample was scanned to minimize cross-contamination. With the increase of pathlength, the sensitivity for ingredient determination increases, but the wavelength range at which absorbance measurements can be done reduces (Rambla *et al*., 1997). Therefore, a pathlength of 1 mm was chosen through this study.

NIR spectra were collected using OMNIC software (Thermo Nicolet Corporation, Madison, WI, USA). Sample spectra were scanned in 800~2400 nm range. Both sample and background spectra were collected in absorbance. After NIR spectra were collected, spectra were exported from OMNIC software and imported directly into TQ Analyst software (Thermo Nicolet Corporation, Madison, WI, USA) for spectral processing and chemometrics analysis. The water content of biological samples poses a limitation in the use of NIR spectroscopy because water absorbs strongly and contributes to a significant amount of light scattering (Bechmann and Jorgensen, 1998). In addition, water absorption bands can interfere with the spectral features of the chemical parameter (Gayo and Hale, 2007). Because bayberry juices are aqueous solution, NIR spectral region in which water has intense absorbance bands cannot be used for analytical purpose (León *et al*., 2005). Therefore, in this research, the available spectral-analytical windows were from 800 to 1850 nm and from 2050 to 2350 nm. The spectral regions from 1850 to 2050 nm and from 2350 to 2400 nm were not available due to the high absorbance of water in these regions.

HPLC analysis

Prior to the NIR spectral analysis, all samples were identified by HPLC to determine glucose, sucrose, fructose, malic acid and citric acid contents. Sugar analysis was done as follows. Bayberry juice was diluted, centrifuged at 10000 r/min for 20 min to remove solids, and passed through a 0.45 μm cellulose acetate filter. The injected sample volume was 10 μl. The HPLC settings were: Hypersil NH_2 column (150 mm×4.6 mm i.d.); mobile phase 0.75% (w/v) acetonitrile in water; flow rate 1.0 ml/min; run time 10 min; column temperature 35 °C; and refractive index detection (model 2414). To check the reproducibility of the HPLC measurement, each sample was measured twice. Quantification was performed by integrating the peak areas of the HPLC results using a computer-assisted optimization program, DryLab2000 (LC Resources).

HPLC analysis of organic acids was done on an Alliance 2695 separations module (Waters, Milford, MA, USA) equipped with a Photodiode Array Detector 2996 (Waters, Milford, MA, USA). The diluted bayberry juice was filtered through a Millipore membrane (0.45 μm) and then 10 μl was injected into a Diamonsil C18 column (250 mm×4.6 mm i.d.). The mobile phase was 50 mmol/L $(NH_4)_2HPO_4$ at a flow rate of 1.0 ml/min, and its pH value was adjusted to pH 2.7. The eluent was monitored at 210 nm for the two organic acids.

Data analysis

Chemometrics analysis, which can highlight the chemical differences between samples and reduce variation due to physical effects, was performed using TQ Analyst software. In this study, the radial basis function neural network (RBFNN), a kind of ANN, was used to distinguish authentic and adulterated juices. The ratio of the number of samples to the number of adjustable parameters in the ANN should be kept as large as possible (Akhlaghi and Kompany-Zareh, 2005). One way to realize it is to compress the input data, especially, when they consist of absorbance recorded at several hundred wavelengths (Akhlaghi and Kompany-Zareh, 2005). The most popular method for data reduction in chemometrics is PCA.

PCA was used for a preliminary examination of the spectral dataset, and was applied to extract the main information on the NIR spectra recorded on juice samples, reduce the number of variables and express the total variation in the dataset in only a few principal components (PCs). Each spectrum will have its own unique set of scores; therefore, a spectrum can

be represented by its PCA scores in the factor space instead of intensities in the wavelength space (Park *et al*., 2003). By plotting the PCs, one can view inter-relationships between different variables, detect and interpret sample patterns, groupings, similarities or differences (Mouazen *et al*., 2006).

Recently, there is a growing interest in the use of RBFNN for its short training time and being guaranteed to reach the global minimum of error surface during training (Liu *et al*., 2004). RBFNN is used to solve modeling and classification problems (Pulido *et al*., 1999). It has superiorities in function approximation and learning speed (Qu *et al*., 2007). The RBFNN can be considered as a three-layer feed-forward neural network with a simple architecture. In RBFNN, the first layer does not process the information and only input a distributor. In our case, the inputs were scores from the PCA of different preprocessed spectra data. Each neuron of the hidden layer represents a radial function with equal dimensions to the input data and the number of radial functions depends on the problem to be solved (Fidêncio *et al*., 2001). The RBFNN discussed in this study has two or three neurons in the output layer, for two or three classes should be determined. Fig.1 shows a detailed structure of RBFNN. As shown, the training parameters are very small.

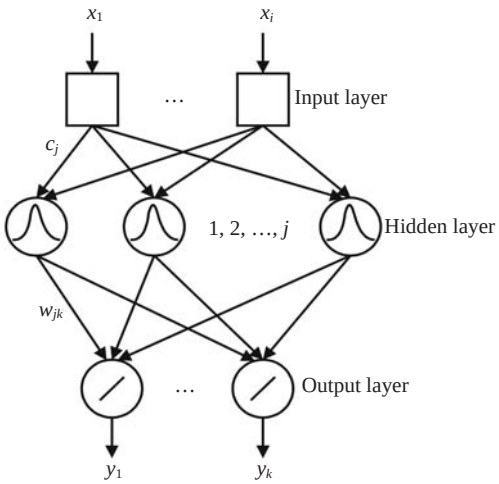

Fig.1 The structure of RBFNN

The type of activation function of RBFNN in this report is Gaussian function (Derks *et al*., 1995; Walczak and Massart, 1996). The successful implementation of these networks is to find suitable centers for such a function, which is characterized by two parameters, i.e., center (c_j) and peak width (σ).

$$K = \exp\left(\frac{-\| \boldsymbol{x} - c_j \|}{2\sigma^2}\right), \quad j=1, 2,\ldots, m, \qquad (1)$$

where K represents the radial basis function, and $\|\boldsymbol{x}-c_j\|$ is the Euclidean distance between $\boldsymbol{x}$ input vector and c_j. The outputs from the radial functions are fully connected to the neurons of the output layer by the strength of weight coefficients w_{jk}. The relation between the output value and the input variable can be represented by:

$$y_k = \sum_{j=1}^{m} w_{jk} K + b_j, \quad k=1, 2, \ldots, p, \qquad (2)$$

where m is the total number of hidden layer neurons, j represents the jth node in the hidden layer, b_j is the bias. w_{jk} connects the output layer, and the hidden layer is adjusted to minimize the mean square error of the net output. When the error of network output reaches the pro-set error goal value in RBFNN, the procedure of adding hidden neurons will stop (Qu *et al*., 2007).

Spectral data pretreatment plays very important roles in improving model performance. Common approaches to preprocessing NIR spectral data are standard normal variate (SNV) transformation and multiplicative scatter correction (MSC), which were implied in this study.

RESULTS

Spectral analysis

The average raw absorbance spectra of authentic and adulterated (10% (w/w) and 20% (w/w) water) juice samples were shown in Fig.2. The spectra revealed peaks at 955~973, 1147~1192, 1434~1456, and 1777~1800 nm. Valleys at 1650~1720, 1812~1830, and 2187~2244 nm were also observed. These regions show slight differences by magnitude. Apart from these regions, no significant spectral difference can be observed between the three spectra and the spectra are highly overlapped. Fig.2 also illustrates that the lowest molecular absorptivities were in the

short wavelength region (833~1408 nm) with higher values in the first overtone region (1490~1850 nm) and still higher absorbance levels in the combination region (2083~2350 nm). The average spectrum of the sample blended with 20% (w/w) water has a higher absorbance than other spectra because of the absorbance of a larger amount of water. As the level of water decreases, the absorbance decreases. The absorptions at 970 and 1450 nm are dominated by water absorption bands, 970 nm corresponding to O-H band stretching and water second overtone and 1450 nm related to O-H bond stretching and water first overtone. The region from 800 to 1350 nm is related to the second and third overtones and is characterized by low intensity and a low signal-to-noise ratio (Gierlinger *et al*., 2004). C-H first overtone stretch vibration modes in CH_3 and CH_2 groups occur around 1660~1760 nm region, whereas the bands located between 2200 and 2300 nm are attributed to the first set of C-H combination bands which are characteristic sugar bands (León *et al*., 2005; Iñón *et al*., 2005). The wavelengths at 2320 and 2350 nm correspond to the lipid C-H combination and the second overtone (Laporte and Paguin, 1999). The absorption band at 2270 nm is assigned to the C-H stretch and deformation combination. Fig.2 shows that the spectra were rather rough because of high resolution. However, noise is an undesirable feature in the spectra (Pedro and Ferreira, 2005). To solve this problem, the original spectra can be preprocessed by mean smoothing, which will not eliminate important features of the spectra. All the relevant chemical information can be retained for modeling. The number of variables will also be reduced, which helps to improve modeling speed. The smooth average (data not shown) spectra had the same trend with the raw spectra while there was less noise in the smooth spectra and absorbance peaks were more evident.

Because of the similarity of spectra, chemometrics techniques were required to distinguish between spectra of authentic and adulterated juice samples.

PCA

Following PCA of the entire raw spectral dataset, a score plot of the first two principal components is shown for a preliminary examination of the spectral data (Fig.3), which reveals the feasibility of discrimination between the samples. The plot of components 1 and 2, which respectively account for 88.291% and 7.828% of the variation in the spectra, showed that although no special grouping can be observed for pure and adulterated bayberry juices, groups can also be observed. From Fig.3, we can find that there is no clear boundary in samples and that many points overlap each other. Pure bayberry juice samples are located on the top left of the plot, while the water-adulterated samples mainly located on the right half and the bottom left. It can also be seen that adulterated samples cannot be divided and some overlaps occur between these two classes. We could see two apparent clusterings of the adulterated samples that were not related to the extent of adulteration, and this may be attributed to the different characteristics of samples, differences in geographical origin, effects of harvest time, and so forth. This clustering phenomenon can also be found in authentic samples, although the boundary is not so clear. Table 1 shows the descriptive statistics for the glucose, sucrose, fructose, malic acid and citric acid concentrations of these two groups adulterated with water. The

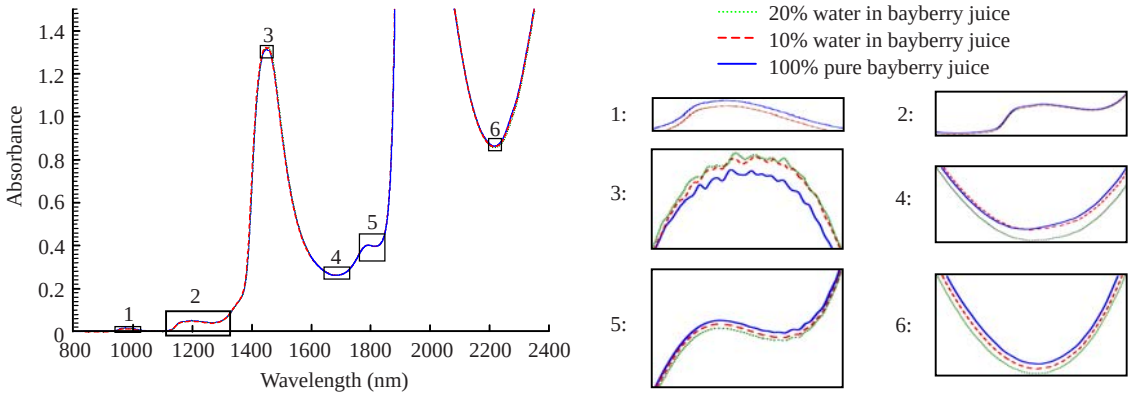

Fig.2 Average spectra of authentic and adulterated [10% (w/w) and 20% (w/w)] bayberry juice

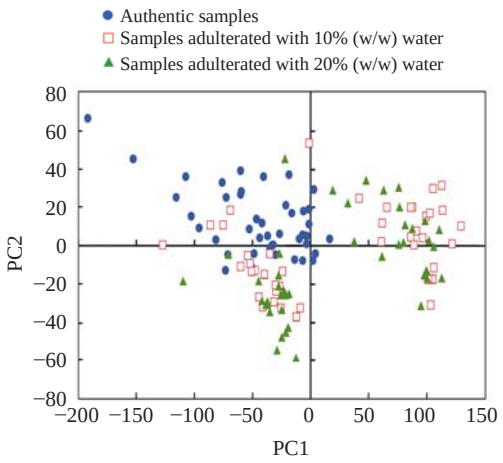

Fig.3 The first two PC scores' plot of bayberry juices using raw NIR spectra

Table 1 Statistics (range, mean, standard deviation (*SD*)), and significance of differences for the glucose, sucrose, fructose, malic acid and citric acid concentrations of the two groups

Comp.	Group	Range (g/100 g)	Mean (g/100 g)	*SD* (g/100 g)	*P*
Glucose	1	0.46~1.04	0.68	0.1854	<0.05
	2	0.31~1.16	0.65	0.1884	
Sucrose	1	1.28~3.39	2.39	0.6037	
	2	1.60~3.57	2.79	0.5750	
Fructose	1	2.23~6.60	4.62	0.9982	<0.05
	2	2.55~6.70	4.88	1.1481	
Malic acid	1	0.36~2.07	0.85	0.3991	<0.05
	2	0.42~2.56	1.11	0.5601	
Citric acid	1	4.45~20.38	9.00	4.6348	<0.05
	2	4.76~13.10	6.83	1.9361	

Group 1: samples located on the right half of the plot shown in Fig.2; Group 2: samples on the bottom left of Fig.2

significant differences in the values of all items, except for sucrose content, were determined by using a Student's *F* test ($P<0.05$).

PCA gives very important information about the basic data structure regarding a potential capability of separation of objects (Andre, 2003). The result suggests that the discrimination between authentic and adulterated samples, at least some of them, may be feasible, and that different spectral attributes of samples are associated with characteristics of the samples.

PC-RBFNN classification

In RBFNN, the spread and the number of the radial basis function are the two important parameters affecting the performance of RBFNN (Liu *et al.*, 2004). In order to find the optimum values of these two parameters and prohibit the overfitting of the model, leave-one-out cross-validation of the whole training set was performed. Different numbers of PCs for the input layer constructed the RBFNN architectures. The optimal RBFNN models were listed in Table 2, where the samples were divided into two classes, authentic and adulterated samples. Parameters of optimum RBFNN, the number of neuron in the hidden layer and the value of peak width (σ) were determined corresponding to different preprocess methods. Table 3 shows the classification results in the validation set (14×3). The use of SNV of NIR spectral data with the number of neuron=17 and peak width (σ)=1.8 produced a very high level of classification rate. Overall a classification accuracy of 97.62% was reached, which demonstrated a good discriminatory power to differentiate authentic and adulterated samples. For authentic juice samples, only one sample was incorrectly assigned. For adulterated ones, the correct classification rate was 100%. The worst model that produced the lowest correct classification rate on validation sample sets involved 22 neurons and σ of 0.8 was obtained using raw spectral data. Five authentic samples were incorrectly classified (64.29% accurate classification) and eight adulterated samples were misclassified (71.43% accurate classification). Overall identification rate was 69.05%, indicating that RBFNN method with proper spectral preprocessing produced a good classification for authentic and adulterated juice samples.

Table 2 Parameters of optimum RBFNN using raw, SNV, MSC spectral data

	σ	Number of neuron
Raw spectra	0.8	22
SNV spectra	1.8	17
MSC spectra	0.2	20

Table 3 Overall prediction results for the validation set objects

Pretreatment	Real class	Predicted class		Identification rate (%)
		AUS	ADS	
Raw spectra	AUS	9	5	69.05
	ADS	8	20	
SNV spectra	AUS	13	1	97.62
	ADS	0	28	
MSC spectra	AUS	13	1	95.24
	ADS	1	27	

AUS: authentic sample; ADS: adulterated sample

In the present study, we also found that 13 of authentic juice samples were misclassified into the adulterated samples [10% (w/w) or 20% (w/w) water] according to different spectral data (not shown), and about half of the 10% (w/w) water-blended samples were incorrectly assigned to the 20% (w/w) water-blended samples. We were not able to distinguish those two water-adulterated classes from each other, no matter what pretreatment methods were used, indicating that RBFNN method could not quantitatively detect the level of water blended in bayberry juice.

CONCLUSION

In this study, PCA was used to reduce the dimensionality of NIR spectra data and evaluate possible classes among samples. RBFNN was trained and used effectively to process the scores of PCs to discriminate authentic and adulterated samples, and good performance was achieved, with a total recognition rate of 97.62% using SNV spectral data. However, adulterated samples cannot be distinguished by RBFNN models according to the level of water added. It could be concluded that PC-RBFNN was an available alternative for juice adulteration detection based on NIR spectroscopy. Further studies are needed to further develop this combined approach for the identification and quantification of water that is blended in bayberry juices.

References

Akhlaghi, Y., Kompany-Zareh, M., 2005. Comparing radial basis function and feed-forward neural networks assisted by linear discriminant or principal component analysis for simultaneous spectrophotometric quantification of mercury and copper. *Analytica Chimica Acta*, **537**(1-2): 331-338. [doi:10.1016/j.aca.2004.12.079]

Andre, M., 2003. Multivariate analysis and classification of the chemical quality of 7-aminocephalsporanic acid using near-infrared reflectance spectroscopy. *Analytical Chemistry*, **75**(14):3460-3467. [doi:10.1021/ac026393x]

Bechmann, I.E., Jorgensen, B.M., 1998. Rapid assessment of quality parameters for frozen cod using near infrared spectroscopy. *Lebensmittel-Wissenschaft und-Technologie*, **31**(7-8):648-652. [doi:10.1006/fstl.1998.0418]

Chen, J., Arnold, M.A., Small, G.W., 2004. Comparison of combination and first overtone spectral regions of near-infrared calibration models for glucose and other biomolecules in aqueous solutions. *Analytical Chemistry*, **76**(18):5405-5413. [doi:10.1021/ac0498056]

Chen, Z.L., 1996. The history of bayberries. *Journal of Fruit Science*, **13**(1):59-61 (in Chinese).

Cozzolino, D., Kwiatkowski, M.J., Parker, M., Cynkar, W.U., Dambergs, R.G., Gishen, M., Herderich, M.J., 2004. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta*, **513**(1):73-80. [doi:10.1016/j.aca.2003.08.066]

Derks, E.P.P.A., Sánchez, M.S., Buydens, L.M.C., 1995. Robustness analysis of radial base function and multi-layered feed-forward neural network models. *Chemometrics and Intelligent Laboratory Systems*, **28**(1): 49-60. [doi:10.1016/0169-7439(94)00082-T]

Fang, Z., Zhang, M., Wang, L., 2007. HPLC-DAD-ESIMS analysis of phenolic compounds in bayberries (*Myrica rubra* Sieb. et Zucc.). *Food Chemistry*, **100**(2):845-852. [doi:10.1016/j.foodchem.2005.09.024]

Fidêncio, P.H., Ruisánchez, I., Poppi, R.J., 2001. Application of artificial neural networks to the classification of soils from São Paulo state using near-infrared spectroscopy. *Analyst*, **126**(12):2194-2200. [doi:10.1039/b107533k]

Fu, X.P., Ying, Y.B., Zhou, Y., Xie, L.J., Xu, H.R., 2008. Application of NIR spectroscopy for firmness evaluation of peaches. *Journal of Zhejiang University SCIENCE B*, **9**(7):552-557. [doi:10.1631/jzus.B0720018]

Gautz, L.D., Kaufusi, P., Jackson, M.C., Bittenbender, H.C., Tong, C., 2006. Determination of kavalactones in dried kava (*Piper methysticum*) power using near-infrared reflectance spectroscopy and partial least-squares regression. *Journal of Agricultural and Food Chemistry*, **54**(17): 6147-6152. [doi:10.1021/jf060964v]

Gayo, J., Hale, S.A., 2007. Detection and quantification of species authenticity and adulteration in crabmeat using visible and near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, **55**(3):585-592. [doi:10.1021/jf061801+]

Gayo, J., Hale, S.A., Blanchard, S.M., 2006. Quantitative analysis and detection of adulteration in crab meat using visible and near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, **54**(4):1130-1136. [doi:10.1021/jf051636i]

Gestal, M., Gómez-Carracedo, M.P., Andrade, J.M., Dorado, J., Fernández, E., Prada, D., Pazos, A., 2004. Classification of apple beverages using artificial neural networks with previous variable selection. *Analytica Chimica Acta*, **524**(1-2):225-234. [doi:10.1016/j.aca.2004.02.030]

Gierlinger, N., Schwanninger, M., Wimmer, R., 2004. Characteristics and classification of Fourier-transform near infrared spectra of the heartwood of different larch species (*Larix* sp.). *Journal of Near Infrared Spectroscopy*, **12**(22):113-119.

Iñón, F.A., Llario, R., Garrigues, S., Guardia, M., 2005. Development of a PLS based method for determination of the quality of beers by use of NIR: spectral ranges and sample-introduction considerations. *Analytical and Bioanalytical Chemistry*, **382**(7):1549-1561. [doi:10.1007/

s00216-005-3343-9]

Laporte, M.F., Paguin, P., 1999. Near-infrared analysis of fat, protein, and casein in cow's milk. *Journal of Agricultural and Food Chemistry*, **47**(7):2600-2605. [doi:10.1021/jf980929r]

León, L., Kelly, J.D., Downey, G., 2005. Detection of apple juice adulteration using near-infrared transflectance spectroscopy. *Applied Spectroscopy*, **59**(5):593-599. [doi:10.1366/0003702053945921]

Liu, H.X., Zhang, R.S., Yao, X.J., Liu, M.C., Hu, Z.D., Fan, B.T., 2004. Prediction of electrophoretic mobility of substituted aromatic acids in different aqueous-alcoholic solvents by capillary zone electrophoresis based on support vector machine. *Analytica Chimica Acta*, **525**(1):31-41. [doi:10.1016/j.aca.2004.07.033]

Liu, L., Cozzolino, D., Cynkar, W.U., Gishen, M., Colby, C.B., 2006. Geographic classification of Spanish and Australian Tempranillo red wines by visible and near-infrared spectroscopy combined with multivariate analysis. *Journal of Agricultural and Food Chemistry*, **54**(18):6754-6759. [doi:10.1021/jf061528b]

Mouazen, A.M., Karoui, R., de Baerdemaeker, J., Ramon, H., 2006. Classification of Soils into Different Moisture Content Levels Based on VIS-NIR Spectra. The 2006 ASABE Annual International Meeting. Oregon Convention Center, Portland, Oregon, 9-12 July, Paper No. 061067.

Nagy, S., 1997. Economic adulteration of fruit beverages. *Fruit Process*, **4**:125-131.

Park, B., Abbott, J.A., Lee, K.J., Choi, C.H., Choi, K.H., 2003. Near-infrared diffuse reflectance for quantitative and qualitative measurement of soluble solids and firmness of delicious and gala apples. *Transactions of the ASAE*, **46**(6):1721-1731.

Pedro, A.M.K., Ferreira, M.M.C., 2005. Nondestructive determination of solids and carotenoids in tomato products by near-infrared spectroscopy and multivariate calibration. *Analytical Chemistry*, **77**(8):2505-2511. [doi:10.1021/ac048651r]

Pulido, A., Ruisánchez, I., Rius, F.X., 1999. Radial basis functions applied to the classification of UV-visible spectra. *Analytica Chimica Acta*, **388**(3):273-281. [doi:10.1016/S0003-2670(99)00082-3]

Qu, N., Li, X., Dou, Y., Mi, H., Guo, Y., Ren, Y., 2007. Nondestructive quantitative analysis of erythromycin ethylsuccinate powder drug via short-wave near-infrared spectroscopy combined with radial basis function neural networks. *European Journal of Pharmaceutical Sciences*, **31**(3-4):156-164. [doi:10.1016/j.ejps.2007.03.006]

Rambla, F.J., Garrigues, S., Guardia, M., 1997. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. *Analytica Chimica Acta*, **344**(1-2):41-53. [doi:10.1016/S0003-2670(97)00032-9]

Walczak, B., Massart, D.L., 1996. The radial basis functions-partial least squares approach as a flexible non-linear regression technique. *Analytica Chimica Acta*, **331**(3):177-185. [doi:10.1016/0003-2670(96)00202-4]