

Identification of pummelo cultivars by using Vis/NIR spectra and pattern recognition methods

Xun-lan Li^{1,2,3} · Shi-lai Yi³ · Shao-lan He³ · Qiang Lv³ ·
Rang-jin Xie³ · Yong-qiang Zheng³ · Lie Deng³

Published online: 18 December 2015
© Springer Science+Business Media New York 2015

Abstract Vis/NIR spectroscopy was used in combination with pattern recognition methods to identify cultivars of pummelo (*Citrus grandis* (L.) Osbeck). A total of 240 leaf samples, 60 for each of the four cultivars were analyzed by Vis/NIR spectroscopy. Soft independent modeling of class analogy (SIMCA), partial least square discriminant analysis (PLS-DA), back propagation neural network (BPNN) and least squares support vector machine (LS-SVM) were applied to the spectral data. The first 8 principal components extracted by principal component analysis were used as inputs in building the BPNN and the LS-SVM models. The results showed that a 97.92 % of discrimination accuracy was achieved for both the BPNN and the LS-SVM models when used to identify samples of the validation set, indicating that the performance of the two models was acceptable. Comparatively, the results of the PLS-DA and the SIMCA models were unacceptable because they had lower discrimination accuracy. The overall results demonstrated that use of Vis/NIR spectroscopy coupled with the use of BPNN and LS-SVM could achieve an accurate identification of pummelo cultivars.

Keywords Vis/NIR spectroscopy · Pummelo · SIMCA · PLS-DA · BPNN · LS-SVM

✉ Lie Deng
liedeng@163.com

Xun-lan Li
lixunlan2009@126.com

¹ Chongqing Academy of Agricultural Sciences, Jiulongpo, Chongqing 401329, China

² College of Horticulture and Landscape, Southwest University, Beibei, Chongqing 400715, China

³ Citrus Research Institute, Southwest University – Chinese Academy of Agricultural Sciences, Chongqing 400712, China

Introduction

Pummelo is one of the most important species in *Citrus*. Pummelo fruits are attractively large with flesh colors varying from white, yellow and pink to red and are high in vitamin C content. Pummelo peels are very thick and contain significant amount of antioxidant flavonoids, mostly naringine and neohesperidin, and dietary fiber pectin, and their leaves, flowers and fruits are rich in essential oils (Du and Chen 2009; Huang and Chen 1998). For these reasons, the pummelo industry is very important in China and in other pummelo growing countries.

There exists no simple and rapid method for the identification of pummelo cultivars. Many pummelo cultivars are derived from elite lines of seed progenies, so they are morphologically very similar to each other. Existing methods for the identification of pummelo cultivars are mainly based on morphology, palynology, biochemistry, cytology and molecular biology (Uzun et al. 2009), and are usually time-consuming and costly, and are therefore inconvenient to be used where rapid identification is needed.

Vis/NIR spectroscopy is a powerful technique widely used in quantitative and qualitative analysis of materials including plant materials with characteristic spectroscopic profiles. The technique is inexpensive and simple to use, requires very small amount of sample and can, at the same time, achieve a rapid and highly sensitive measurement. It has been used in identification and classification of various agricultural materials including strawberries (Sánchez et al. 2012), wolf berries (Du and Gong 2013), milk powders (Wu et al. 2008b) and natural textile raw materials (Zhou et al. 2008). A wide range of plant materials can be used in Vis/NIR analysis, and leaves should be much better in case of identifying a cultivar for the following reasons: first, leaves are available nearly all year around compared to other plant organs, especially on evergreen fruit trees; second, a leaf spectrum is easy to obtain.

Vis/NIR spectroscopy is usually used in combination with pattern recognition methods in identification and classification research including identification of pharmaceutical excipients (Candolfi et al. 1999), instant milk teas (Liu et al. 2009), apple cultivars (Luo et al. 2011) and moldy chestnuts (Zhou et al. 2011). Soft independent modeling of class analogy (SIMCA), partial least square discriminant analysis (PLS-DA), back propagation neural network (BPNN) and least squares-support vector machine (LS-SVM) are the most widely used pattern recognition methods in qualitative analysis.

As far as the authors are aware, there has been no report on the identification of pummelo cultivars using Vis/NIR spectroscopy techniques. In this study, the applicability of the technique in identifying pummelo cultivars was evaluated, and models were established using methods of SIMCA, PLS-DA, BPNN and LS-SVM. The objectives of this study were to develop a rapid and reliable method for the identification of pummelo cultivars using leaf spectra obtained by Vis/NIR spectroscopy.

Table 1 The plant materials used in the study

Cultivars	Code	Scientific name
Lingnanshatianyou	LN	<i>Citrus grandis</i> (L.) Osbeck cv. Lingnanshatianyou
Guanximiyou	GX	<i>Citrus grandis</i> (L.) Osbeck cv. Guanximiyou
Pingshanyou	PS	<i>Citrus grandis</i> (L.) Osbeck cv. Pingshanyou
Gulaoqianshatianyou	GL	<i>Citrus grandis</i> (L.) Osbeck cv. Gulaoqianshatianyou

Materials and methods

Sample preparation

The four pummelo cultivars used (Table 1) in this study were located in the National Citrus Germplasm Repository of Chongqing, China. Healthy mature leaves were randomly collected from spring shoots growing inside the tree canopy in mid-September of 2013. Sampled leaves were temporarily kept in sealed plastic bags that were immediately placed in a storage box chilled with iced-bags. A total of 240 leaf samples (60 leaves for each cultivar) were obtained. In order to develop calibration models, 192 samples (48 for each cultivar) were randomly selected as the sample set for calibration and the remaining 48 samples (12 for each cultivar) were used as the sample set for validation.

Equipment and spectra acquisition

The Vis/NIR spectra of pummelo leaves were collected using a FieldSpec Handheld spectrometer (Analytical Spectral Devices, Inc. Boulder, Colorado, USA) in a darkroom. The light source was provided by a halogen lamp (ASD Pro Lamp) at a 45° angle. The distance between light source and the surface of the leaf sample was 300 mm. The spectroradiometer was approximately 150 mm from the surface of the leaf being analyzed. The detectable wavelength range of the spectrometer was between 325 and 1075 nm, with a sampling interval of 1.6 nm and a spectral resolution of 1 nm. The field-of-view (FOV) of the spectrometer was 25°. The light source was turned on 30 min earlier to warm up the halogen bulb before analysis. Before spectra acquisition, the instrument was calibrated with a standard whiteboard. Leaf was placed on a 150 mm × 300 mm black fabric. Ten spectra were recorded and averaged for every leaf sample.

All spectra were read and exported from the instrument by the RS³ software (Analytical Spectral Devices, Inc. Boulder, Boulder, Colorado, USA). The software Unscrambler V9.7 (CAMO ASA, Oslo, Norway) was used for data preprocessing, and for the SIMCA and the PLS-DA models. The BPNN model was implemented in Matlab R2010a (The MathWorks, Inc. Massachusetts, USA) and free LS-SVM toolbox (LS-SVM v 1.7, Suykens, Leuven, Belgium) were used to build the LS-SVM model.

Pre-processing

To avoid low signal-to-noise ratio, the spectral range of 400–1000 nm was utilized in this study. The raw spectral data were pre-processed to reduce the noise by using moving average smoothing with nine segments (Gao et al. 2004). To remove baseline drifts and enhance spectral differences, the second order derivative was utilized by using the Savitzky–Golay convolution method with a second-order polynomial and a 5-point segment (Chu et al. 2004).

Principal component analysis

Principal component analysis (PCA) was used to simplify the complex spectral data (He et al. 2006). In the end, the first eight principal components (PCs) were extracted and used in relevant analysis.

Soft independent modeling of class analogy

Soft independent modeling of class analogy (SIMCA) is a pattern recognition method that is based on PCA. PCA was applied in the calibration set to develop PCA class models. The PCs to describe most of the variation within each class was determined by cross-validation (Wold 1976). Unknown samples were then identified by the distances to different PCA models. The model distance limit S_{\max} is calculated for the class model m as follows in Eq. 1:

$$S_{\max(m)} = S_o(m)\sqrt{F_C} \quad (1)$$

where S_o is the average distance within the model, F_c (Fisher criterion) is the critical value. The performance of SIMCA is evaluated by recognition rate and rejection rate (Ning et al. 2008).

Partial least squares discriminant analysis

Partial least squares discriminant analysis (PLS-DA) is a supervised pattern recognition method that is based on partial least squares regression. In PLS-DA, each sample in the calibration set needs to be assigned a dummy variable as a categorical variable Y for each cultivar (0001 for LN, 0010 for GX, 0100 for PS and 1000 for GL). A regression analysis between categorical variable Y and spectral data of calibration set X is then conducted to establish the PLS regression model. The general model of PLS-DA can be written as Eqs. 2 and 3:

$$X = TP^T + E \quad (2)$$

$$Y = UQ^T + F \quad (3)$$

where X represents the matrix of spectral data, T is a factor score matrix, P is the X loadings and E is the residual or a noise term. Y is a matrix of the categorical variable, U is the scores for Y , Q is the Y loadings, and F is the residuals. Full cross-validation is used to evaluate the quality and prevent over-fitting of the calibration model. Finally, predicted categorical variable values (Yp) are determined by the PLS-DA model. If the values between a sample of the validation set and a categorical variable meet the following conditions: $|Yp - Y| > 0.5$ and deviation value < 0.5 , the sample belongs to this class (Yang et al. 2008).

Back propagation neural network

As one of the most popular neural network topologies, the BPNN is commonly used in the present Vis/NIR spectroscopy analysis. The BPNN is consisting of an input layer, one or more hidden layers and an output layer. Training procedures of BPNN are comprised of a forward propagation of input signal and a backward propagation of error. The signal is successively transmitted to input layer, hidden layer and output layer, and the status of every layer can only be affected by next layer. If anticipant outcome cannot be gotten in output layer, then change to back-propagation. According to the error signal of back propagation, the back-propagation algorithm network adjusts weights in each successive layer to reduce the errors at each level until all the errors are within the required tolerance

(Yang et al. 2011). A schematic diagram of multilayer neural network architecture is shown in Fig. 1.

The transfer function of net-layer takes the form:

$$f(x) = \frac{1}{1 + e^{-xQ}} \quad (4)$$

where, Q is the parameter of function Sigmoid.

When constructing a BPNN model, parameters such as the structure, the initial connection weights and thresholds, the learning rate and expected error need to be considered. In theory, a BPNN with three layers can solve arbitrary classification. In this study, a three layers BPNN was constructed to identify different varieties of pummelo. PCA was used to simplify the complex spectral data (He et al. 2006). PCs extracted by PCA were used as inputs. Dummy variable values (0001 for LN, 0010 for GX, 0100 for PS and 1000 for GL) were used as outputs.

Least square-support vector machine

Least square-support vector machine (LS-SVM) is a statistical learning theory that adopts a least squares linear system as a loss function. LS-SVM has the capability for linear and non-linear multivariate calibration and solving the multivariate calibration problems in a relatively faster way, so that LS-SVM has been commonly applied in the present Vis/NIR spectroscopy analysis. The details of the theoretical background to the LS-SVM algorithm are given in Wu et al. (2008a) and Suykens and Vandewalle (1999). The LS-SVM regression model can be expressed as:

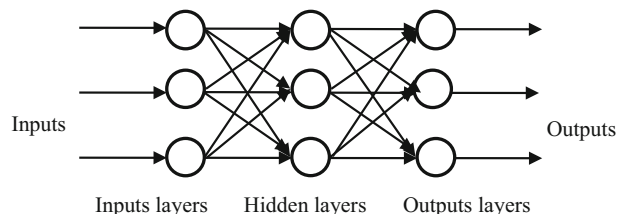
$$Y(X) = \sum_{i=1}^n \alpha_i K(X, X_i) + b \quad (5)$$

where α_i is the Lagrange multiplier called support value, X_i is the input vector, $K(X, X_i)$ is the kernel function mapping input into high dimensional space and b is the bias term, respectively. Determination of optimal input feature subset, proper kernel function and optimum kernel parameters are three crucial problems that need to be solved in LS-SVM. Because radial basis function (RBF) kernel can handle the non-linear relationships, reduce the computational complexity of the training procedure and give a good performance under general smoothness assumptions (Borin et al. 2006), RBF kernel was used as the kernel function of LS-SVM in this study. RBF kernel can be expressed as:

$$\text{Radial basis function : } K(X, X_i) = \exp\left(-\frac{X - X_i^2}{\sigma^2}\right) \quad (6)$$

where σ is the kernel parameter.

Fig. 1 Schematic diagram of the topological structure of BPNN



Grid-search technique and cross-validation were applied to find out the optimal parameter values. To simplify the complex spectral data, PCA was applied in the spectral analysis. PCs extracted by PCA were used as inputs to LS-SVM.

Results and discussion

Principal component analysis

The 3D scatter plots of the four pummelo cultivars were generated following a PCA of their respective leaf spectra. As shown in Fig. 2, leaf samples from the four cultivars were clustered into four groups in 3D space. However, there was a moderate overlap between samples of GX and LN and a small overlap between samples of PS and GL. This indicated that the leaves of GX and LN were more similar to each other in leaf spectra or only three PCs with a total accumulated reliability of 64.4 % were not representative enough to express their differences. Table 2 shows the accumulated reliability: 81.7 % of it was represented by the first 6 PCs much larger than 77.1 % represented by the first 5 PCs; the accumulated reliability had little change represented by the first 6 PCs to the first 10 PCs. So the BPNN and the LS-SVM models were established based on different numbers of PC from 6 to 10. The optimal number of PC was determined as 8 after comparing the final prediction results of BPNN models and LS-SVM models respectively. Thus, the first 8 PCs were used as the input variables in building the next BPNN model and LS-SVM model.

Results using SIMCA

The discrimination results based on SIMCA are shown in Table 3. When the F_c was set to 25 %, the appropriate number of PCs determined by cross-validation was 6, 4, 5 and 3 for LN, GX, PS and GL respectively. In the sample set for calibration, a recognition rate of 100 %, with the smallest percentages of rejection, could be reached for PS. The average rates of recognition and rejection for the calibration sample set were 95.83 and 83.33 %

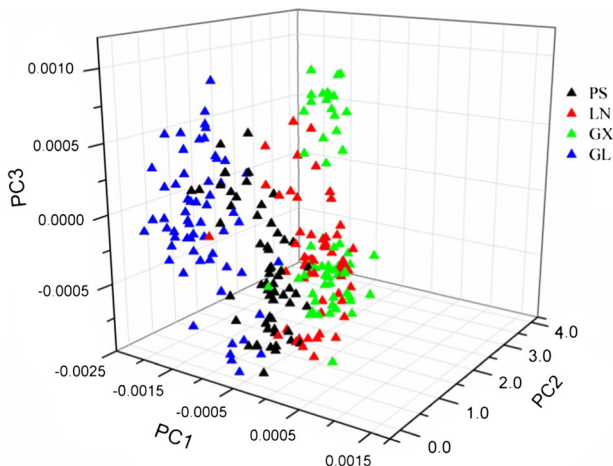


Fig. 2 3D-scatter plots of PC1 vs PC2 vs PC3 extracted by PCA of four cultivars of pummelo

Table 2 The first 10 principal components of the accumulated reliability

PCs	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Reliability (%)	38.89	13.71	11.81	6.72	6.02	4.61	2.88	1.86	1.17	0.77
Accumulative reliability (%)	38.89	26.30	64.40	71.12	77.13	81.74	84.62	86.49	87.65	88.4

Table 3 Results of calibration set and validation set using SIMCA

Cultivars	No. of PCs	Calibration set		Validation set	
		Recognition (%)	Rejection (%)	Recognition (%)	Rejection (%)
LN	6	97.92	97.92	83.33	100
GX	4	85.42	100	50	83.33
PS	5	100	79.62	91.67	83.33
GL	3	85.42	100	91.67	66.67
Total		95.83	83.33	79.17	83.33

respectively. In the validation sample set, the highest rate of recognition was 91.67 %, and the average rates of recognition and rejection for the sample set were 79.17 and 83.33 % respectively which were not acceptable for discriminating the four pummelo cultivars.

Results using PLS-DA

The discrimination results of PLS-DA model are shown in Table 4. Using the model, 98.44 % of the samples in the calibration sample set were correctly identified. For the samples in the validation set, those from GX and GL were correctly discriminated, but two and one of the samples were incorrectly identified in PS and LN respectively. In total, 93.75 % of the samples in validation set were correctly discriminated. Compared with the SIMCA model, the performance of the PLS-DA model was better, though a discrimination rate of 83.33 % for PS was still insufficient.

Results using BPNN

In building the BPNN model, the momentum and the least learning rate were set at 0.6 and 0.01, while the threshold residual error and the training time were set at 0.001 and 1500 s, respectively. Hyperbolic tangent sigmoid and logarithmic sigmoid transfer functions were

Table 4 Results of calibration set and validation set using PLS-DA

Cultivars	Calibration set				Validation set			
	LN	GX	PS	GL	LN	GX	PS	GL
LN	46	0	1	1	11	1	0	0
GX	0	48	0	0	0	12	0	0
PS	0	0	48	0	1	0	10	1
GL	1	0	0	47	0	0	0	12
Overall accuracy (%)	98.44				93.75			

Table 5 Results of calibration set and validation set using BPNN

Cultivars	Calibration set				Validation set			
	LN	GX	PS	GL	LN	GX	PS	GL
LN	46	0	1	1	12	0	0	0
GX	0	48	0	0	0	12	0	0
PS	0	0	48	0	1	0	11	0
GL	1	0	0	47	0	0	0	12
Overall accuracy (%)	98.44				97.92			

applied in the hidden and the output layers respectively. After adjustment of the parameters, a three-layer BPNN model was built, in which the number of neurons was 8, 9, 4 for the input layer, the hidden layer and the output layer respectively.

As shown in Table 5, the BPNN model could identify 98.44 % of the calibration samples correctly, which is comparable to the result of the PLS-DA model. The overall accuracy of the BPNN model in identifying the samples of the validation set was 97.92 %. In comparison, the BPNN results were the best among the three models. Particularly, the recognition rate of 97.92 % for the validation samples was much higher than that obtained by either the SIMCA model or the PLS-DA model. This indicated that Vis/NIR spectroscopy combined with BPNN could be utilized to identify pummelo cultivars.

Results using LS-SVM

LS-SVM is usually applied in two-class classification. In order to use it in multi-class classification, the minimum output coding (MOC) scheme was introduced into LS-SVM in this study (Van Gestel et al. 2002). The first 8 PCs obtained from PCA were used as inputs in LS-SVM modeling. The adjustment parameter γ and the RBF kernel parameter σ^2 were 43.3768 and 0.017499 respectively, both of which were optimized by a grid-search technique and cross-validation. As shown in Table 6, for all samples in the calibration set, the overall accuracy was 98.44 %, which was the same as those obtained by the PLS-DA and the BPNN models. For all samples in the validation set, a prediction accuracy of 97.92 % was obtained, which was the same as that obtained by BPNN.

The results showed that both the BPNN model and the LS-SVM model were superior to the PLS-DA model and the SIMCA model for the classification of pummelo cultivars. However, there were cases where LN was miss-identified as GL or vice versa when the calibration samples were analyzed by PLS-DA, BPNN and LS-SVM. This is perhaps related to the fact that the GL and LN are genetically closer to each other.

Table 6 Results of calibration set and validation set using LS-SVM

Cultivars	Calibration set				Validation set			
	LN	GX	PS	GL	LN	GX	PS	GL
LN	46	0	1	1	11	1	0	0
GX	0	48	0	0	0	12	0	0
PS	0	0	48	0	0	0	12	0
GL	1	0	0	47	0	0	0	12
Overall accuracy (%)	98.44				97.92			

Conclusions

In this study, leaf Vis/NIR spectra of four pummelo cultivars were obtained and analyzed, and four pattern recognition methods, SIMCA, PLS-DA, BPNN and LS-SVM were compared. PCA was used to extract PCs that were used as inputs for LS-SVM and BPNN models. It was found that the BPNN and LS-SVM models performed better than the SIMCA and the PLS-DA models as shown by their respective discrimination accuracy. Satisfactory discrimination accuracy was achieved, which was 98.44 % for the BPNN model and the LS-SVM model, when samples of the calibration set were analyzed. Discrimination accuracy of 97.92 % was achieved for the BPNN model and the LS-SVM model, when samples of the validation set were analyzed. The PLS-DA model data also showed that the PLS-DA model could achieve the same identification rate which was 98.44 % as the BPNN and LS-SVM models achieved in samples of the calibration set, yet its discrimination accuracy was only 83.33 % in validation set of PS. Thus, the PLS-DA model was insufficient for discriminating pummelo cultivars. In contrast, the BPNN model and the LS-SVM model, as non-linear methods, performed better in discriminating pummelo cultivars.

This research demonstrated that Vis/NIR spectroscopy combination with BPNN and LS-SVM has the potential ability to identify the different cultivars of pummelo and offered a new approach to the fast discriminating. This fast discriminating system will reduce considerable pressure on labor for identifying cultivars of pummelo by the crop expert, and makes the rapid and on-the-go identification of pummelo cultivars in the field possible, especially in identifying the cultivars of nursery stock of *Citrus* before field planting.

Moreover, as our future work, further studies are still needed to cover more cultivars of pummelo and a more robust discrimination model for application of this fast discriminating system in the fields is desired.

Acknowledgments This research was supported by Program of China (863 Program) (Project No. 2012AA101904), International Science and Technology Cooperation program of China (Project No. 2013DFA11470), International Science & Technology Cooperation Program of Chongqing (Project No. CSTC2011ghz80001), National Key Technology R&D Program (Project No. 2012BAD35B08-3), The national sparking plan project (Project No. 2012GA8110017).

References

- Borin, A., Ferrão, M. F., Mello, C., Maretto, D. A., & Poppi, R. J. (2006). Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. *Analytica Chimica Acta*, 579(1), 25–32.
- Candolfi, A., De Maesschalck, R., Massart, D. L., Hailey, P. A., & Harrington, A. C. E. (1999). Identification of pharmaceutical excipients using NIR spectroscopy and SIMCA. *Journal of Pharmaceutical and Biomedical Analysis*, 19(6), 923–935.
- Chu, X. L., Yuan, H. F., & Lu, W. Z. (2004). Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique. *Progress in Chemistry*, 16(4), 528–542.
- Du, X. H., & Chen, Y. (2009). QSRR study on volatile components of essential oil from pomelo peel. *Food Science*, 30(19), 61–64.
- Du, M., & Gong, Y. (2013). Rapid identification of wolfberry fruit of different geographic regions with sample surface near infrared spectra combined with multi-class SVM. *Spectroscopy and Spectral Analysis*, 33(5), 1211–1214.
- Gao, R. Q., Fan, S. F., Yan, Y. L., & Zhao, L. L. (2004). Preprocessing of near infrared spectroscopic data. *Spectroscopy and Spectral Analysis*, 24(12), 1563–1565.

- He, Y., Feng, S., Deng, X., & Li, X. (2006). Study on lossless discrimination of varieties of yogurt using the Visible/NIR-spectroscopy. *Food Research International*, 39(6), 645–650.
- Huang, Y. Z., & Chen, Q. Y. (1998). The chemical components of essential oils from the leaves of 110 species and cultivars of Citrus plants. *Acta Botanica Sinica*, 40(9), 846–852.
- Liu, F., Ye, X., He, Y., & Wang, L. (2009). Application of visible/near infrared spectroscopy and chemometric calibrations for variety discrimination of instant milk teas. *Journal of Food Engineering*, 93(2), 127–133.
- Luo, W., Huan, S., Fu, H., Wen, G., Cheng, H., Zhou, J., et al. (2011). Preliminary study on the application of near infrared spectroscopy and pattern recognition methods to classify different types of apple samples. *Food Chemistry*, 128(2), 555–561.
- Ning, Z., Dequan, Z., Shurong, L., & Qingpeng, L. (2008). Preliminary study on origin traceability of mutton by near infrared reflectance spectroscopy coupled with SIMCA method. *Transactions of the Chinese Society of Agricultural Engineering*, 24(12), 309–312.
- Sánchez, M. T., De la Haba, M. J., Benítez-López, M., Fernández-Navales, J., Garrido-Varo, A., & Pérez-Marín, D. (2012). Non-destructive characterization and quality control of intact strawberries based on NIR spectral data. *Journal of Food Engineering*, 110(1), 102–108.
- Suykens, J. A. K., & Vandewalle, K. J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9, 293–300.
- Uzun, A., Yesiloglu, T., Tuzcu, O., & Gulsen, O. (2009). Genetic diversity and relationships within Citrus and related genera based on sequence related amplified polymorphism markers (SRAPs). *Scientia Horticulturae*, 121(3), 306–312.
- Van Gestel, T., Suykens, J. A., Lanckriet, G., Lambrechts, A., De Moor, B., & Vandewalle, J. (2002). Multiclass LS-SVMs: Moderated outputs and coding-decoding schemes. *Neural Processing Letters*, 15(1), 45–58.
- Wold, S. (1976). Pattern recognition by means of disjoint principal components models. *Pattern Recognition*, 8(3), 127–139.
- Wu, D., He, Y., Feng, S. J., & Bao, Y. D. (2008a). Application of infrared spectra technique based on LS-support vector machines to the non-destructive measurement of fat content in milk powder. *Journal of Infrared and Millimeter Waves*, 27(3), 180–184.
- Wu, D., Yang, H., Chen, X., He, Y., & Li, X. (2008b). Application of image texture for the sorting of tea categories using multi-spectral imaging technique and support vector machine. *Journal of Food Engineering*, 88(4), 474–483.
- Yang, Z., Ren, H. Q., & Jiang, Z. H. (2008). Discrimination of wood biological decay by NIR and partial least squares discriminant analysis. *Spectroscopy and Spectral Analysis*, 28(4), 793–796.
- Yang, Y., Zhu, J., Zhao, C., Liu, S., & Tong, X. (2011). The spatial continuity study of NDVI based on Kriging and BPNN algorithm. *Mathematical and Computer Modelling*, 54(3), 1138–1144.
- Zhou, Z., Li, X., Li, P., Gao, Y., Zhan, H., & Liu, J. (2011). Near-infrared spectral detection of moldy chestnut based on GA-LSSVM and FFT. *Transactions of the Chinese Society of Agricultural Engineering*, 27(3), 331–335.
- Zhou, Y., Xu, H. R., & Ying, Y. B. (2008). NIR analysis of textile natural raw material. *Spectroscopy and Spectral Analysis*, 28(12), 2804–2807.