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

Technical Report:

Neural network and principal component regression in non-destructive soluble solids content assessment: a comparison*

Kim-seng CHIA†, Herlina ABDUL RAHIM†‡, Ruzairi ABDUL RAHIM

(*Department of Control and Instrumentation, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia*)

†E-mail: kschia2@live.utm.my; herlina@fke.utm.my

Received May 31, 2011; Revision accepted Aug. 29, 2011; Crosschecked Dec. 6, 2011

Abstract: Visible and near infrared spectroscopy is a non-destructive, green, and rapid technology that can be utilized to estimate the components of interest without conditioning it, as compared with classical analytical methods. The objective of this paper is to compare the performance of artificial neural network (ANN) (a nonlinear model) and principal component regression (PCR) (a linear model) based on visible and shortwave near infrared (VIS-SWNIR) (400–1000 nm) spectra in the non-destructive soluble solids content measurement of an apple. First, we used multiplicative scattering correction to pre-process the spectral data. Second, PCR was applied to estimate the optimal number of input variables. Third, the input variables with an optimal amount were used as the inputs of both multiple linear regression and ANN models. The initial weights and the number of hidden neurons were adjusted to optimize the performance of ANN. Findings suggest that the predictive performance of ANN with two hidden neurons outperforms that of PCR.

Key words: Artificial neural network (ANN), Principal component regression (PCR), Visible and shortwave near infrared (VIS-SWNIR), Spectroscopy, Apple, Soluble solids content (SSC)

doi:10.1631/jzus.B11c0150 **Document code:** A **CLC number:** TS207.3; S123

1 Introduction

Food safety and quality, which are directly related to the consumers' health and social progress, are important issues throughout the world (Cen and He, 2007). In particular, high quality fruits are crucial to human health because abundant nutritional components can be found in fruits (Lin and Ying, 2009). Therefore, it is necessary to sort out high quality fruits based on their internal quality for the fresh market.

The internal quality of fruit can be evaluated based on its acidity, pH, firmness, total solids content, and soluble solids content (SSC). Among these internal quality attributes, SSC is one of the main factors that influence consumer decisions regarding apple purchases (Zhu *et al.*, 2009). Conventionally, SSC can be measured by using a refractometer. However, a refractometer can only assess the liquid sample. In other words, the SSC measurement of fruit using a refractometer cannot be carried out without destroying the fruit. Therefore, the development of a non-invasive SSC assessment technology is necessary.

Visible and shortwave near infrared (VIS-SWNIR) spectroscopy is a non-destructive, green, and rapid technology for estimating the components of interest without conditioning it in the way classical analytical methods do. The wavelength ranges of

‡ Corresponding author

* Project (No. UTM.J.10.01/13.14/1/127/1 Jld 3(48)) supported by the Zamalah Scholarship from the Universiti Teknologi Malaysia

visible spectra and SWNIR spectra are from 380 to 780 nm and from 780 to 1 100 nm, respectively. VIS-SWNIR spectra consist of useful information for fruit internal quality assessment. Visible spectra, on the one hand, contain the information of plant pigment groups, e.g., chlorophyll, anthocyanins, and carotenoids (Herold *et al.*, 2008). On the other hand, SWNIR spectra or the Herschel region provides the information of overtones of molecular vibrations, e.g., N-H second overtone, O-H second overtone, and C-H third overtone (Osborne, 2006). Overtone is a series of absorptions that are multiples of the fundamental vibration frequency (Bosco, 2010). As a result, these hydrogen bonds between the molecules and functional groups in the near infrared region make VIS-SWNIR spectroscopy promising in food internal quality assessment (Sundaram *et al.*, 2009).

However, the information from SWNIR spectra is broad and overlapped with the absorption peaks of other absorbers (Herold *et al.*, 2008). Therefore, a calibration model is required to establish the relationship between the component of interest and the spectrum. In fact, a combination of near infrared spectroscopy and calibration model has been successfully implemented in non-destructive SSC assessment of various fruits such as apple (Ventura *et al.*, 1998; Peirs *et al.*, 2003; Liu and Ying, 2004; Alamar *et al.*, 2007; Peng and Lu, 2007; Zou *et al.*, 2007; Fan *et al.*, 2009; Zhu *et al.*, 2009), citrus (Lu *et al.*, 2006; Zude *et al.*, 2008), loquat (Fu *et al.*, 2009) mango (Yu and He, 2009), mangosteen (Teerachaichayut *et al.*, 2007), melon (Long, 2005; Tian *et al.*, 2007), and pineapple (Pathaveerat *et al.*, 2008).

Recently, numerous studies have been published concerning the use of linear models [e.g., multiple linear regression (MLR) (Ventura *et al.*, 1998; Peng and Lu, 2007; Jha and Garg, 2010) and partial least square (PLS) (Peirs *et al.*, 2003; Xing *et al.*, 2005; Alamar *et al.*, 2007; Zou *et al.*, 2007; Fan *et al.*, 2009; Zhao *et al.*, 2009; Zhu *et al.*, 2009; Jha and Garg, 2010)] for the internal quality measurement of apples, using spectroscopic technology. However, the implementation of nonlinear models [e.g., least squares support vector machine (LS-SVM) (Shao and He, 2009; Shao *et al.*, 2009) and least squares support vector regression (LS-SVR) (Liu *et al.*, 2010)] has been reported to be better than that of linear models (e.g., PLS). Since the utilization of artificial neural network (ANN) (nonlinear model) has received comparatively little attention in the literature compared to linear models, this study investigated the ability of ANN in SSC assessment of apple based on VIS-SWNIR spectra compared to the principal component regression (PCR).

2 Materials and methods

2.1 Sample preparation

A total of 116 Fuji apples used in this experiment were bought directly from a local supermarket in Malaysia. First, all apples were washed using distilled water carefully, and labeled individually before the experiment. To avoid any significant temperature difference among the apples, all apples were stored in a laboratory (Control Laboratory 2, Faculty of Electrical Engineering, Universiti Teknologi Malaysia) at room temperature (~25 °C) for 24 h before spectrum acquisition. After that, all samples were randomly separated into calibration and prediction sets of 84 and 32 samples, respectively. The conventional approach of the SSC measurement via refractometer was carried out immediately once the spectrum acquisition of each apple had been completed.

2.2 Spectrum measurement

The reflectance spectrum of each apple was acquired using FieldSpec HandHeld Spectroscopy (Analytical Spectral Devices, Boulder, USA) with a 325–1075 nm wavelength range and 10° field-of-view. The VIS-SWNIR light source used in this work was a tungsten halogen bulb (ASD Pro Lamp). The optical standard reference (white reference) was a Teflon® disk.

The distance between spectroscopy and the surface of an apple was approximately 11 cm. The position of spectroscopy was fixed to approximately 90° at the top of the apple. The light source was placed in the position that the VIS-SWNIR light can reach to the surface of the apple without shadow on it. After that, the position of light source was fixed throughout the experiment. Lastly, to avoid a stray light effect from the circumstance, the system was shielded with a black cotton material.

The light source had been turned on for at least 30 min to warm up the halogen bulb before the

spectrum acquisition began. Only one reflectance spectrum from the middle position of each apple was acquired. The spectrum was the average of 300 successive scans with 6-nm intervals. All spectral data were stored in a computer and processed via the RS3 software (Analytical Spectral Devices, Boulder, USA).

2.3 Soluble solids content reference measurement

The SSC reference acquisition was started immediately using a digital hand-held "pocket" refractometer (ATAGO Co., Ltd., Japan) with automatic temperature compensation after spectrum acquisition of each apple was completed. The accuracy of its refractive index is ±0.2%. The measurement range and resolution of this instrument are 0–95% (°Brix) and 0.1% (°Brix), respectively.

A portion of flesh (depth of 1 cm and diameter of 2 cm) from the same position of spectral acquisition was cut and squeezed directly into a beaker. After that, a few drops of the juice from the beaker were used to assess SSC by using the refractometer. The standard reference used to calibrate the refractometer was reverse osmosis water.

2.4 Data pre-processing

First, the reflectance spectra were converted into the most commonly used absorbance terms [$\log(1/R)$], and then transformed into ASCII format by using the ASD ViewSpecPro software V5.6.10 (Analytical Spectral Devices, Boulder, USA). Due to big scattering effects at the beginning and the end of the spectrum, this study utilized the spectrum from 400 to 1000 nm only. After that, all data (both spectra and their respective SSC reference values) were loaded as matrix format into Matlab workspace [MATLAB V7.4.0.287 (R2007a)]. Data were randomly separated into calibration and prediction sets in such a way that the range of SCC in calibration set was larger than or equal to the range of SCC in prediction set to avoid extrapolation problems.

Multiplicative scattering correction (MSC) is the most general method used for reducing the scatter effect, adjusting the baseline shifts and removing imperfections from the data matrix before model calibration begins (Rinnan *et al.*, 2009). The procedure of the implementation of MSC is summarized as follows. First, the mean of calibration set spectra was used as the reference value for MSC. Second, an ordinary least square method was applied to estimate the multiplicative and additive correction coefficients of each spectrum based on the reference value. Last, each individual original spectrum was corrected based on its calculated correction coefficients.

2.5 Linear model

PCR is a combination of principal component analysis (PCA) and MLR. The first step in PCR is to decompose a spectral data matrix using PCA. Generally, there are two types of decomposition techniques. The first technique is by computing eigenvectors and eigenvalues. The second technique is using singular value decomposition (SVD). In this work, we used SVD to decompose the spectral data matrix. This is because SVD is generally accepted as the most stable and numerically accurate technique (Gemperline, 2006).

SVD decomposes spectral data into column-mode eigenvectors, singular values, and row-mode eigenvectors. The products of column-mode eigenvectors and singular values, which are the so-called principle components (PCs), were used as the input variables of calibration models to avoid the redundancy and collinearity effects. Row-mode eigenvectors were used to transform the spectral data of the prediction set to PCs for prediction purpose. In order to determine the optimal number of PCs, a graph of root mean square error (RMSE) versus the number of PCs was plotted. After that, the optimal number of PCs was used as the input of MLR to perform PCR (de Maesschalck *et al.*, 1999).

2.6 Nonlinear model

After completing the calibration of the linear model, the same number of PCs was used as the input variables of ANN. The architecture of the ANN in this work is: nonlinear transfer function of log-sigmoid was used in the hidden layer so that the network would be capable to handle nonlinearity problem; linear transfer function of purelin was selected in the output layer; Levenberg-Marquardt algorithm was chosen as the back-propagation network training algorithm; gradient descent with momentum weight and bias learning function was applied as back-propagation weight/bias learning function; and, mean square error was implemented as performance function.

The training process of a feedforward back-propagation ANN consists of two steps, i.e., forward pass and error back-propagation (Despagne and Luc Massart, 1998). In the first step, the predicted outcome will be produced based on the architecture of ANN, the initial weights and input data. After that, the difference between predicted outcomes and desired outcomes will be calculated and utilized to adjust the weights of the ANN in such a way that errors of the next predicted outcome can be reduced. These two steps are repeated until convergence is reached.

Next, in order to optimize the performance of ANN, the number of hidden neurons was increased to raise the complexity of ANN model. In addition, the best initial weight was chosen from 100 different initial weights. This is because the performance of ANN is dependent heavily on the initialization of weights (Timotheou, 2009).

3 Results and discussion

3.1 Brix distribution

Table 1 summarizes the distribution of the SSC of the apples. The SSCs of 116 apples were fairly normally distributed around the mean value of 13.33 °Brix with a standard deviation (SD) of 0.88 °Brix. The range of prediction set (11.6–14.8 °Brix) was within the range of calibration set (11.2–15.0 °Brix). The purpose of this condition is to avoid extrapolation prediction.

3.2 Optimal number of principal components

In this work, the optimal number of PCs was estimated based on the performance of PCR model with minimum root mean square error of prediction (RMSEP) value and without over-fitting. This method is similar to the approach in previous studies for the determination of the optimal input number in PLS (Liu *et al.*, 2010) and PCR (Serneels and Verdonck, 2009). However, this method contrasts with the approach that based on the cumulative percentage of explained data variance (Mouazen *et al.*, 2010). Although the determination of the optimal PCs number still remains as an open question (Abdi and Williams, 2010), PCs with small variance (less than 0.1% variance) should not be ignored without proper investigation and evaluation. This is because some of these PCs may contain important and relevant information that can improve the robustness and predictive accuracy of a calibration model.

In addition, the determination of the input dimension is crucial to avoid both under-fitting and over-fitting problems. An under-fitting model is not capable to achieve its full predictive potential. In contrast, an over-fitting model is always less robust and not reliable, i.e., this kind of model always provides too optimistic results from the calibration set, but poor results from new data set. Therefore, a compromise between under-fitting and over-fitting is essential as pointed out by Næs and Mevik (2001) and Næs *et al.* (2002).

Fig. 1 illustrates the root mean square error of calibration (RMSEC) and RMSEP of a PCR model against the number of PCs as its input variables. By inspection, the root mean square errors (RMSEs) of PCR models in both calibration and prediction sets

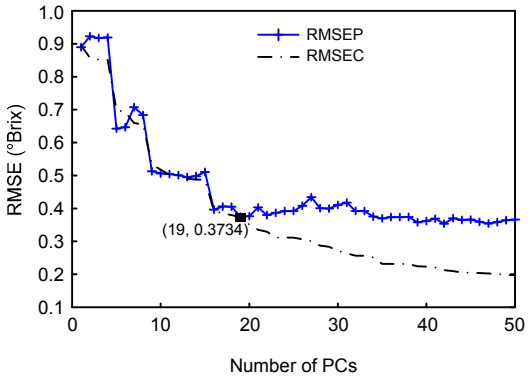

Fig. 1 Root mean square errors (RMSEs) of calibration and prediction sets with different numbers of principal components (PCs)

Table 1 Soluble solids content (SSC) distribution of apples

Sample set	Number of sample	SSC (°Brix)				Coefficient of variation (%)
		Min.	Max.	Mean	SD	
Calibration set	84	11.2	15.0	13.26	0.88	6.64
Prediction set	32	11.6	14.8	13.51	0.86	6.37
Total	116	11.2	15.0	13.33	0.88	6.60

are decreased dramatically as the number of inputs (i.e., PCs) is increased. Although the RMSEC decreases when the number of inputs increases, it is found that the predictive performance (RMSEP) does not improve or improve insignificantly when the input number is more than 19. Therefore, this observation indicates that the optimal number of PCs is 19.

3.3 Artificial neural network with principal components (PCs-ANN)

The performance of PCs-ANN with one hidden neuron indicated that the predictive ability was not improved significantly compared to that of PCR, i.e., PCs-ANN with one hidden neuron was 11.4% better than PCR in terms of RMSEP. It was not an unexpected result since one hidden neuron is normally not capable enough to solve complicated problems. However, it is worth noting that the implementation of a nonlinear function in ANN makes the predictive model superior to that of the linear predictive model. This is in agreement with Cen *et al.* (2007) showing that ANN has better predictive performance than linear models (e.g., PLS) in orange variety classification. In fact, the Beer-Lambert law is only valid for a dilute and non-saturated system (Despagne and Luc Massart, 1998). Therefore, some useful information may have nonlinear characteristics.

Recently, Mouazen *et al.* (2010) advocated that ANN with latent variables (LVs) was the best when compared with PCs-ANN, PCR, and PLS models in soil property assessments. However, this is worth to highlight that latent variables are generally not recommended to be the inputs of ANN. This is because the information of latent variables is linearly correlated with the dependent variable (Despagne and Luc Massart, 1998).

Next, the performance of PCs-ANN with two hidden neurons was significantly better than that of PCR, i.e., PCs-ANN with two hidden neurons was 29.3% better than PCR in terms of RMSEP. Fig. 2 shows the predicted SSC versus the reference SSC of both PCs-ANN (with two hidden neurons) and PCR models.

Although the findings indicated that the predictive accuracy of PCs-ANN with two hidden neurons was superior to PCR model in terms of RMSEP, surprisingly, some results from PCs-ANN were worse than those from PCR model in the sample with SSC values of 12.5, 13.3, 13.6, and 14.6 °Brix. This observation suggests that better outcomes could be achieved by combining both nonlinear and linear models based on ensemble idea. This is because by combining the strengths of both nonlinear and linear models, an ensemble predictive model should have higher accuracy and robustness than a single model (Hastie *et al.*, 2009). In addition, the weakness of both models may be complemented (Beyer *et al.*, 2009) by each other to increase its predictive performance.

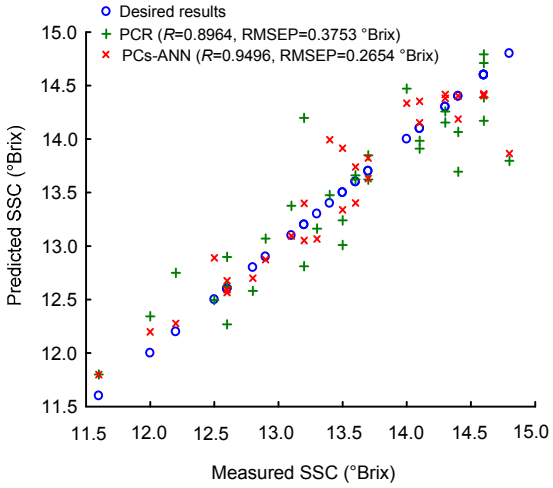

Fig. 2 Predicted SSC against measured SSC based on PCs-ANN model with two neurons in the hidden layer and PCR model

All performance indicators in this paper are stated in terms of RMSEC, RMSEP, and coefficient of correlation (*R*-value) as tabulated in Table 2. A marginally difference between RMSEC and RMSEP indicates that the robustness of these models is satisfactory, with no over-fitting problems. The relationship between SSC and VIS-SWNIR spectrum is strong, based on the high coefficient of correlation.

4 Conclusions

The performances of PCs-ANN and PCR were compared in the SSC assessment of 'Fuji' apples based on VIS-SWNIR data. The results indicate that the nonlinear predictive model (PCs-ANN) is superior to the linear predictive model (PCR) in the SSC prediction of apples. The accuracies of PCs-ANN with two hidden neurons and PCR in terms of RMSEP were 0.2654 and 0.3734 °Brix, respectively.

Table 2 Summary of predictive models and their performances

Model	Number of PCs	Number of neurons in hidden layer	RMSEC (°Brix)	RMSEP (°Brix)	*R*-value	
					Calibration set	Prediction set
PCR	19		0.37	0.38	0.90	0.90
PCs-ANN	19	1	0.35	0.33	0.92	0.93
PCs-ANN	19	2	0.27	0.27	0.95	0.95

In addition, the PCs-ANN with two hidden neurons achieved better predictive performance compared to the PCs-ANN with one hidden neuron. Therefore, the findings suggest that PCs-ANN with appropriate parameters, which are the initial weights and the number of hidden neurons, is capable to achieve a better accuracy in SSC measurement based on VIS-SWNIR data compared to PCR.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.*, **2**(4):433-459. [doi:10.1002/wics.101]

Alamar, M.C., Bobelyn, E., Lammertyn, J., Nicolaï, B.M., Moltó, E., 2007. Calibration transfer between NIR diode array and FT-NIR spectrophotometers for measuring the soluble solids contents of apple. *Postharv. Biol. Technol.*, **45**(1):38-45. [doi:10.1016/j.postharvbio.2007.01.008]

Beyer, J., Heesche, K., Hauptmann, W., Otte, C., Kruse, R., 2009. Ensemble Learning for Multi-source Information Fusion. Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Springer, Berlin Heidelberg, p.748-756.

Bosco, G.L., 2010. James L. Waters Symposium 2009 on near-infrared spectroscopy. *TrAC Trends Anal. Chem.*, **29**(3):197-208. [doi:10.1016/j.trac.2010.01.003]

Cen, H., He, Y., 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends Food Sci. Technol.*, **18**(2):72-83. [doi:10.1016/j.tifs.2006.09.003]

Cen, H., He, Y., Huang, M., 2007. Combination and comparison of multivariate analysis for the identification of orange varieties using visible and near infrared reflectance spectroscopy. *Eur. Food Res. Technol.*, **225**(5-6):699-705. [doi:10.1007/s00217-006-0470-2]

de Maesschalck, R., Estienne, F., Verdú-Andrés, J., Candolfi, A., Centner, V., Despagne, F., Jouan-Rimbaud, D., Walczak, B., Massart, D.L., Jong, S.D., *et al*., 1999. The Development of Calibration Models for Spectroscopic Data Using Principal Component Regression. Available from http://www.vub.ac.be/fabi/tutorial/PCR.pdf [Accessed on Dec. 9, 2011].

Despagne, F., Luc Massart, D., 1998. Neural networks in multivariate calibration. *Analyst*, **123**(11):157R-178R. [doi:10.1039/A805562I]

Fan, G., Zha, J., Du, R., Gao, L., 2009. Determination of soluble solids and firmness of apples by Vis/NIR transmittance. *J. Food Eng.*, **93**(4):416-420. [doi:10.1016/j.jfoodeng.2009.02.006]

Fu, X.P., Li, J.P., Zhou, Y., Ying, Y.B., Xie, L.J., Niu, X.Y., Yan, Z.K., Yu, H.Y., 2009. Determination of soluble solid content and acidity of loquats based on FT-NIR spectroscopy. *J. Zhejiang Univ.-Sci. B*, **10**(2):120-125. [doi:10.1631/jzus.B0820097]

Gemperline, P.J., 2006. Principal Component Analysis. Practical Guide to Chemometrics, 2nd Ed. CRC Press, p.69-104. [doi:10.1201/9781420018301.ch4]

Hastie, T., Tibshirani, R., Friedman, J., 2009. Ensemble Learning. The Elements of Statistical Learning. Springer, New York, p.1-20.

Herold, B., Kawano, S., Sumpf, B., Tillmann, P., Walsh, K., 2008. VIS/NIR Spectroscopy. *In*: Zude, M. (Ed.), Optical Monitoring of Fresh and Processed Agricultural Crops: Contemporary Food Engineering. CRC Press, p.143-248.

Jha, S., Garg, R., 2010. Non-destructive prediction of quality of intact apple using near infrared spectroscopy. *J. Food Sci. Technol.*, **47**(2):207-213. [doi:10.1007/s13197-010-0033-1]

Lin, H., Ying, Y., 2009. Theory and application of near infrared spectroscopy in assessment of fruit quality: a review. *Sens. Instrum. Food Qual. Saf.*, **3**(2):130-141. [doi:10.1007/s11694-009-9079-z]

Liu, Y., Gao, R., Hao, Y., Sun, X., Ouyang, A., 2010. Improvement of near-infrared spectral calibration models for brix prediction in 'Gannan' navel oranges by a portable near-infrared device. *Food Bioprocess Technol.*, in press. [doi:10.1007/s11947-010-0449-7]

Liu, Y.D., Ying, Y.B., 2004. Measurement of sugar content in Fuji apples by FT-NIR spectroscopy. *J. Zhejiang Univ.-Sci.*, **5**(6):651-655. [doi:10.1631/jzus.2004.0651]

Long, R.L., 2005. Improving Fruit Soluble Solids Content in Melon (*Cucumis Melo* L.) (Reticulatus Group) in the Australian Production System. PhD Thesis, Central Queensland University, Rockhampton, Australia.

Lu, H.S., Xu, H.R., Ying, Y.B., Fu, X.P., Yu, H.Y., Tian, H.Q., 2006. Application Fourier transform near infrared spectrometer in rapid estimation of soluble solids content of intact citrus fruits. *J. Zhejiang Univ.-Sci. B*, **7**(10):794-799. [doi:10.1631/jzus.2006.B0794]

Mouazen, A.M., Kuang, B., de Baerdemaeker, J., Ramon, H., 2010. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma*, **158**(1-2):23-31. [doi:10.1016/j.geoderma.2010.03.001]

Næs, T., Mevik, B.H., 2001. Understanding the collinearity problem in regression and discriminant analysis. *J. Chemom.*, **15**(4):413-426. [doi:10.1002/cem.676]

Næs, T., Isaksson, T., Fearn, T., Davies, T., 2002. A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications, p.24-25.

Osborne, B.G., 2006. Near-Infrared Spectroscopy in Food Analysis. *In*: Meyers, R.A. (Ed.), Encyclopedia of Analytical Chemistry. John Wiley & Sons, Ltd., Chichester, p.1-14. [doi:10.1002/9780470027318.a1018]

Pathaveerat, S., Terdwongworakul, A., Phaungsombut, A., 2008. Multivariate data analysis for classification of pineapple maturity. *J. Food Eng.*, **89**(2):112-118. [doi:10.1016/j.jfoodeng.2008.04.012]

Peirs, A., Scheerlinck, N., Nicolaï, B.M., 2003. Temperature compensation for near infrared reflectance measurement of apple fruit soluble solids contents. *Postharv. Biol. Technol.*, **30**(3):233-248. [doi:10.1016/s0925-5214(03)00118-2]

Peng, Y., Lu, R., 2007. Prediction of apple fruit firmness and soluble solids content using characteristics of multispectral scattering images. *J. Food Eng.*, **82**(2):142-152. [doi:10.1016/j.jfoodeng.2006.12.027]

Rinnan, A., van der Berg, F., Engelsen, S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.*, **28**(10):1201-1222. [doi:10.1016/j.trac.2009.07.007]

Serneels, S., Verdonck, T., 2009. Principal component regression for data containing outliers and missing elements. *Comput. Stat. Data Anal.*, **53**(11):3855-3863. [doi:10.1016/j.csda.2009.04.008]

Shao, Y., He, Y., 2009. Measurement of soluble solids content and pH of yogurt using visible/near infrared spectroscopy and chemometrics. *Food Bioprocess Technol.*, **2**(2):229-233. [doi:10.1007/s11947-008-0180-9]

Shao, Y., Bao, Y., He, Y., 2009. Visible/near-infrared spectra for linear and nonlinear calibrations: a case to predict soluble solids contents and pH value in peach. *Food Bioprocess Technol.*, **4**(8):1376-1383. [doi:10.1007/s11947-009-0227-6]

Sundaram, J., Kandala, C., Butts, C., 2009. Application of near infrared spectroscopy to peanut grading and quality analysis: overview. *Sens. Instrum. Food Qual. Saf.*, **3**(3):156-164. [doi:10.1007/s11694-009-9081-5]

Teerachaichayut, S., Kil, K.Y., Terdwongworakul, A., Thanapase, W., Nakanishi, Y., 2007. Non-destructive prediction of translucent flesh disorder in intact mangosteen by short wavelength near infrared spectroscopy. *Postharv. Biol. Technol.*, **43**(2):202-206. [doi:10.1016/j.postharvbio.2006.09.007]

Tian, H.Q., Ying, Y.B., Lu, H.S., Fu, X.P., Yu, H.Y., 2007. Measurement of soluble solids content in watermelon by Vis/NIR diffuse transmittance technique. *J. Zhejiang Univ -Sci. B*, **8**(2):105-110. [doi:10.1631/jzus.2007.B0105]

Timotheou, S., 2009. A novel weight initialization method for the random neural network. *Neurocomputing*, **73**(1-3):160-168. [doi:10.1016/j.neucom.2009.02.023]

Ventura, M., de Jager, A., de Putter, H., Roelofs, F.P.M.M., 1998. Non-destructive determination of soluble solids in apple fruit by near infrared spectroscopy (NIRS). *Postharv. Biol. Technol.*, **14**(1):21-27. [doi:10.1016/s0925-5214(98)00030-1]

Xing, J., van Linden, V., Vanzeebroeck, M., de Baerdemaeker, J., 2005. Bruise detection on Jonagold apples by visible and near-infrared spectroscopy. *Food Control*, **16**(4):357-361. [doi:10.1016/j.foodcont.2004.03.016]

Yu, J., He, Y., 2009. Fast Measurement of Soluble Solid Content in Mango Based on Visible and Infrared Spectroscopy Technique. *In*: Li, D., Zhao, C. (Eds.), Computer and Computing Technologies in Agriculture II, Volume 1. IFIP Advances in Information and Communication Technology, Springer, Boston, Volume 293, p.89-95. [doi:10.1007/978-1-4419-0209-2_10]

Zhao, J., Vittayapadung, S., Chen, Q., Chaitep, S., Chuaviroj, R., 2009. Nondestructive measurement of sugar content of apple using hyperspectral imaging technique. *Maejo Int. J. Sci. Technol.*, **3**(1):130-142.

Zhu, D., Ji, B., Meng, C., Shi, B., Tu, Z., Qing, Z., 2009. A Comparison of Linear Regression Methods for the Detection of Apple Internal Quality by Near Infrared Spectroscopy. *In*: Li, D., Zhao, C. (Eds.), Computer and Computing Technologies in Agriculture II, Volume 3. IFIP Advances in Information and Communication Technology, Springer, Boston, Volume 295, p.1671-1680. [doi:10.1007/978-1-4419-0213-9_17]

Zou, X.B., Zhao, J.W., Huang, X.Y., Li, Y.X., 2007. Use of FT-NIR spectrometry in non-invasive measurements of soluble solid contents (SSC) of 'Fuji' apple based on different PLS models. *Chemom. Intell. Lab. Syst.*, **87**(1):43-51. [doi:10.1016/j.chemolab.2006.09.003]

Zude, M., Pflanz, M., Kaprielian, C., Aivazian, B.L., 2008. NIRS as a tool for precision horticulture in the citrus industry. *Biosyst. Eng.*, **99**(3):455-459. [doi:10.1016/j.biosystemseng.2007.10.016]