academicJournals

Vol. 8(25), pp. 3334-3339, 4 July, 2013
DOI: 10.5897/AJAR12.281
ISSN 1991-637X ©2013 Academic Journals
http://www.academicjournals.org/AJAR

African Journal of Agricultural Research

Full Length Research Paper

Study temperature influence in the application of near-infrared reflectance spectroscopy (NIRS) to estimate amylose content of rice

Xingang Xie[1] and Lijuan Shi[2]*

[1]College of Engineering, Huazhong Agriculture University, 430070, Wuhan, China.
[2]College of Basic Sciences, Huazhong Agriculture University, 430070, Wuhan, China.

Accepted 22 April, 2013

In the application of near-infrared reflectance spectroscopy (NIRS) to estimate amylose content of rice, analysis was performed based on the relationship between absorbance and amylose content. But absorbance is sensitive to environmental temperature. In this study, characteristics of rice spectra collected at different temperatures were studied by comparing the average spectra and standard deviation spectra of rice with same amylose content at the temperature of 5, 10, 15 and 20°C. The results showed that spectra collected at different temperatures had significant difference in absorbance, and spectra acquired at different temperatures had different stabilities. When the environment temperature was 15°C, stability of spectra was better than when the temperature was 5, 10 and 20°C within the range consisting of abundant information. Accuracies of four linear models on calibration set acquired at four temperatures built by partial least square (PLS) were investigated by using validation sets got at four different temperatures to perform cross validation. The results showed that the relation between spectra and amylose content was not linear in the result of temperature's influence, and non-linear model methodology least square support vector machine (LS-SVM) method was used to establish calibration model on the calibration set acquired at 15°C. The validation results indicated that the non-linear model could predict the amylose content of samples collected at 15°C with high accuracy, and the root mean square error of prediction was 0.62. Meanwhile, non-linear model showed good prediction ability for samples whose spectra were collected at other temperatures.

Key words: Amylose content, near-infrared spectroscopy, temperature influence.

INTRODUCTION

Amylose content is an important characteristic in determining the eating and nutrient quality of rice, which made it necessary to know amylose content when breeding new varieties of rice and selecting raw materials in food processing (Wang et al., 2001). Traditional methods of determining amylose content cannot be applied to a mass of samples because of the lengthy operations. Due to mear-infrared reflectance (NIR), spectrum of material can provide both physical and chemical information about a given sample, it had become a useful analytical tool since the beginning of the 20th century. Near-infrared reflectance spectroscopy

*Corresponding author. E-mail: slj2002@mail.hzau.edu.cn

(NIRS) is not only a rapid and non-destructive technique, but also a kind of technique requiring minimal or no sample preparation. And it was extensively used in large-scale routine analysis of grain characters such as amylose, protein and amino acid (Christiane et al., 2004; Lee et al., 2007; Shao et al., 2009). NIRS, however, a kind of indirect quantitative analysis technique, performs analysis on the basis of the relationship between absorbance and the chemical component concentration. NIR spectrum is sensitive to environmental temperature in that the variance of temperature could cause the change of energy and absorbance (Hansen et at., 2000; Jaillais et al., 2005; Peinado et al., 2006; Wand et al., 2010). And It is difficult to find quantitative relationship from the spectrum directly. Therefore, the objective of this research was not only to explore characteristic of NIR spectra of rice collected at different temperatures, but also find the most appropriate temperature and the stable model to express the relationship between NIR spectrum and amylose content.

MATERIALS AND METHODS

A total of 140 rice accessions were selected, and their amylose content had been measured by iodine colorimetric method.

NIR measurements

A commercial Fourier NIR spectrometer “ANTARIS II” (Thermo Scientific, USA), fitted with gold-plated integral ball, PbS detector and quartz glass sample cup, was available to measure the NIR spectra. All spectra in this study were collected using the Result Integration software supplied by Thermo Scientific Company.

In order to get stable spectra, all samples were put in the laboratory for 24 h to balance the temperature and moisture. Prior to measurement, a system suitability test, consisting of checking the wavelength, absorbance scale and photometric noise, was performed. Samples were vigorously shaken before being placed on the same stage.

When samples were scanned with resolution of 8 cm^{-1}, the Result Integration software controlled instrument to collect reflected strength every 3.86 cm^{-1} and translated it into log (1/R) with the average value stored in the form of file in computer. Each recorded spectrum was the average of 64 scans over the wave number range of 3800 to 12000 cm^{-1}. Spectra were collected at the temperature of 5, 10, 15 and 20°C, respectively.

Least square support vector machine (LS-SVM) method

In recent years, SVM had become a popular machine learning method based on statistical learning theory and the structural risk minimization principle. It can efficiently deal with the problem in small sample size, restrain overfitting, and improve the generalization capability (Vapnik, 1999). LS-SVM, a modified SVM algorithm, expresses the training process by solving a set of linear equations instead of a quadratic programming problem, which reduces computational cost greatly (Suykens and Vandewalle 1999).

The principle of LS-SVM for function fitting can be simply described below. Given a training set including N data, $\{x_k, y_k\}_{k=1}^{N}$, where the input data is x_k ($x_k \in R^m$), and the output data is y_k ($y_k \in R^m$). The problem of function fitting can be described as the following optimal problem.

$$\min_{w,e} J(w,e) = \frac{1}{2} w^T w + \frac{1}{2}\gamma \sum_{k=1}^{N} e_k^2 \tag{1}$$

$$s.t. \qquad y_k = w^T \varphi(x_k) + b + e_k, \qquad k = 1,\ldots,N$$

Where $\varphi(\cdot): R^m \to R^{m_k}$ is the function which maps the input data to high dimension characteristics space, w ($w \in R^{m_k}$) is weight vector, e_k ($e_k \in R$) is error variables, and b ($b \in R$) is the offset value. γ ($\gamma > 0$) is punishment coefficient. The optimal problem of Equation 1 can be transformed to dual space to solve, and Lagrange function is got:

$$L(w,b,e,a) = J(w,e) - \sum_{k=1}^{N} a_k \{w^T \varphi(x_k) + b + e_k - y_k\} \tag{2}$$

Where a_k ($a_k \in R$) is called support value. Partial derivative was calculated on each variable, and the following conditional equalities are got:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \to w = \sum_{k=1}^{N} a_k \varphi(x_k), \\ \dfrac{\partial L}{\partial b} = 0 \to \sum_{k=1}^{N} a_k = 0, \\ \dfrac{\partial L}{\partial e_k} = 0 \to a_k = \gamma e_k, & k = 1,\ldots,N, \\ \dfrac{\partial L}{\partial a_k} = 0 \to w^T \varphi(x_k) + b + e_k - y_k = 0, & k = 1,\ldots,N. \end{cases} \tag{3}$$

When variables w and e are eliminated, linear equations are available.

$$\begin{bmatrix} 0 & I^T \\ I & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{4}$$

Where,

$y = [y_1, \ldots, y_N]^T, I = [1, \ldots, 1]^T, a = [a_1, \ldots, a_N]^T, \Omega = \{\Omega_{kl} | k, l = 1, \ldots, N\}$, $\Omega_{kl} = \varphi(x_k)^T \varphi(x_l) = K(x_k, x_l), \quad k, l = 1, \ldots, N$. $K(x_k, x_l)$ is the core function. In this study, Radial basis function was adopted as the core function.

$$K(x_k, x_l) = \exp(-\|x_k - x_l\|^2 / \sigma^2) \tag{5}$$

Therefore, the LS-SVM fitting model is:

$$y(x) = \sum_{k=1}^{N} a_k K(x_k, x) + b \tag{6}$$

Where, a and b are the solutions of Equation 4 (Suykens et al., 2002).

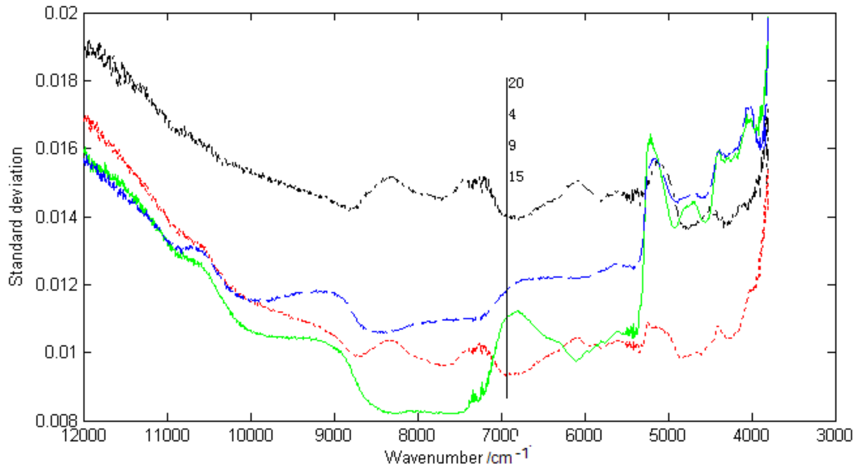

Figure 1. Standard deviation spectra of different temperatures.

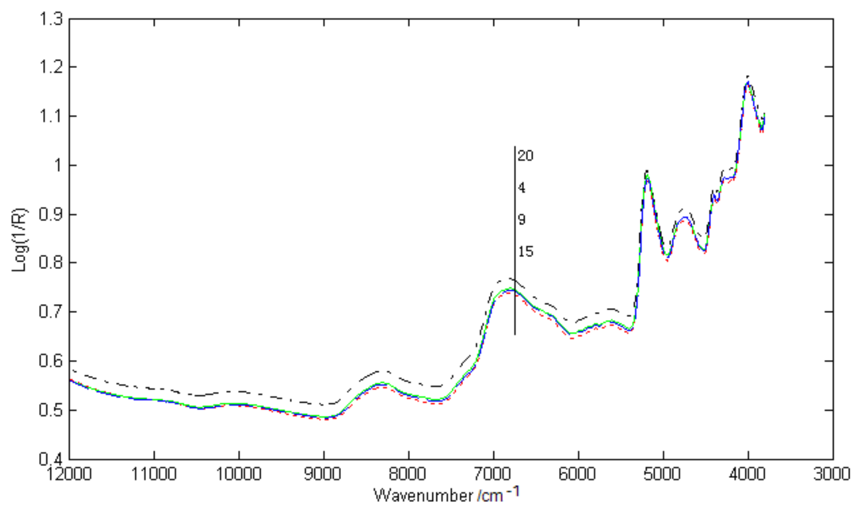

Figure 2. Average spectra of different temperatures.

RESULTS

Influence of environmental temperature on the spectrum of rice

Rice with amylose content of 27.1% was scanned for 10 times at four different temperatures. Four groups of spectra were got with ten spectra in each group and 2217 data points in each spectrum. The standard deviation of the absorbance of spectrum acquired at the same temperature was calculated at each wave number, and four standard deviation spectra are showed in Figure 1. The absorbance of each wave number of spectra acquired at the same temperature was averaged, and four average spectra are showed in Figure 2.

Further studies were performed on absorbance at three sharp peaks of 5179.86, 4740.17, and 4393.05 cm^{-1} within the range of 3800 to 5200 cm^{-1}. The basic statistics of absorbance value of three peaks are showed respectively in Tables 1, 3 and 5, and the results of single factor analysis of variance on temperature are showed, respectively in Tables 2, 4 and 6.

Influence of temperature on the quantitative analysis model

To investigate the influence of temperature on the quantitative analysis model, four independent models were established. Prior to calibration, baseline shift was performed using multiplicative signal correction (MSC), and spectra were smoothed using Savizky-Golay convolution, and finally partial least square (PLS) method was executed over the wave number rang of 4000 to

Table 1. Statistics of absorbance of rice from different temperatures at 5179.86 cm^{-1}.

Temperature (°C)	Samples number	Mean	Standard deviation
5	10	0.9727	0.0157
10	10	0.9793	0.0161
15	10	0.9683	0.0107
20	10	0.9886	0.0154

Table 2. Single factor analysis of variance on temperature at 5179.86 cm^{-1}.

Source	Sum of squares	DF	Mean square	F-value	Pr > F
Model	0.00233922	3	0.00077974	3.64	0.0216
Error	0.00771267	36	0.00021424		
Corrected total	0.01005189	39			

Table 3. Statistics of absorbance of rice from different temperatures at 4740.17 cm^{-1}.

Temperature (°C)	Samples number	Mean	Standard deviation
5	10	0.8948	0.0146
10	10	0.8948	0.0144
15	10	0.8877	0.0099
20	10	0.9123	0.0137

Table 4. Single factor analysis of variance on temperature at 4740.17 cm^{-1}.

Source	Sum of squares	DF	Mean square	F value	Pr > F
Model	0.00328284	3	0.00109428	6.19	0.0017
Error	0.00636005	36	0.00017667		
Corrected total	0.00964289	39			

Table 5. Statistics of absorbance of rice from different temperatures at 4393.05 cm^{-1}.

Temperature (°C)	Samples number	Mean	Standard deviation
5	10	0.8948	0.0146
10	10	0.8948	0.0144
15	10	0.8877	0.0099
20	10	0.9123	0.0137

Table 6. Single analysis of variance on temperature at 4393.05 cm^{-1}.

Source	Sum of squares	DF	Mean square	F value	Pr > F
Model	0.00351312	3	0.00117104	5.74	0.0026
Error	0.00784480	36	0.00020402		
Corrected total	0.01085793	39			

6000 cm^{-1}. Sample of each temperature were divided into calibration subset and external validation subset. The calibration subset was used to develop the calibration equation, and the external validation subset was used to

Table 7. RMSEP result of cross validation for four models.

Model (°C)	RMSEP				
	5°C validation set	10°C validation set	15°C validation set	20°C validation set	Mean
5	2.86	3.25	3.81	4.25	3.54
10	4.31	2.69	2.98	3.27	3.31
15	3.41	2.99	2.54	3.07	3.00
20	4.25	3.65	3.18	2.60	3.42

Table 8. RMSEP of LS-SVM for validation sets acquired at different temperatures.

Validation set (°C)	RMSEP
5	2.53
10	1.19
15	0.62
20	1.52

evaluate the calibration equation. These four models were called 5, 10, 15 and 20°C models, respectively. Validation set of different temperature were called 5, 10, 15, and 20°C validation sets. And cross validation was conducted on the four models using four validation sets to evaluate the prediction capacity of each model, which was measured by root mean square error of prediction (RMSEP).

DISCUSSION

Influence of environmental temperature on the spectrum of rice

Four standard deviation spectra are showed in Figure 1. In the range of 3800 to 12000 cm^{-1}, the standard deviation of spectra acquired at the temperature of 20°C was almost the biggest. And in the range of 7000 to 12000 cm^{-1}, spectra acquired at the temperature of 9°C had the minimal standard deviation. While in the range of 3800 to 7000cm^{-1}, spectra acquired at the temperature of 15°C had almost the smallest standard deviation. The smaller the standard deviation was, the better stability spectra collected at this temperature had.

It can be seen from Figure 2 that the average spectra of 20°C deviated greatly from other three spectra. Compared with the spectral region of 5200 to 12000 cm^{-1}, the spectral region of 3800 to 5200 cm^{-1} showed much more abundant spectral information, but differences were not invisible between four spectra.

In Tables 2, 4 and 6, the analysis of variance on temperature factor at three peaks indicated that significant difference existed in absorbance value at the same wave number of rice with same amylose content. That is to say, temperature led to substantial effect on NIR spectrum of rice. It was found that the minimal standard deviation values of absorbance were always gained at 15°C when the data in the last column of Tables 1, 3 and 5 were compared.

It can be concluded that spectra acquired at different temperatures had different stability. When the environment temperature was 15°C, stability of spectra was better than when the temperature was 5, 10 and 20°C within the range consisting of abundant information.

Influence of temperature on the quantitative analysis model

Table 7 showed that the model whose calibration set and validation set were acquired at the same temperature had the smallest RMSEP, and the larger the temperature varied between the calibration set and validation set, the bigger the RMSEP is. The comparison of the mean of RMSEP in Table 7 indicated that 15°C model had the best prediction ability. However, the prediction accuracy of each model was not satisfactory on the whole. Because temperature is a part of information included in NIR of rice, the relation between amylose content and NIR spectra is not linear. Certainly, PLS which is a linear model cannot reflect this kind of non-linear relation.

In order to get model with high accuracy, spectra acquired at 15°C were taken as the calibration set, and LS-SVM was taken as the method to build model. Other three validation sets were used again to validate this model, and the RMSEP for each validation set is shown in Table 8.

Table 8 shows that not only the amylose content of samples collected at 15°C could be predicted with high accuracy, but also prediction accuracies for amylose content of samples collected at other temperatures were improved greatly in comparison with PLS model.

It can be concluded that quantitative analysis model was also subject to the influence of temperature. The non-linear model was able to reflect the relation between amylose content and NIR spectra instead of linear model.

Conclusions

Application of NIRS to the estimation of amylose content is an alternative choice in favor of fast measurement. It was demonstrated in this study that NIR of rice was subjected to the influence of environmental temperature. Spectra acquired at different temperatures had different stability. When the environment temperature was 15°C, stability of spectra was better than when the temperature was 5, 10 and 20°C. NIR calibration model built with calibration set collected at a certain temperature had better prediction capacity on the samples whose spectra were collected at the same temperature than on the samples whose spectra were collected at other temperatures. The relation between spectra and amylose content was not linear in the result of temperature’s influence, therefore, non-linear model methodology LS-SVM was used to establish calibration model on the calibration set acquired at 15°C. The validation results indicated that the non-linear model could not only predict the amylose content of samples collected at 15°C with high accuracy, but also showed good prediction ability for samples whose spectra collected at other temperatures.

ACKNOWLEDGEMENTS

This work was supported by the Fundamental Research Funds for the Central Universities of PR China under Grant No. 2009JC006 and No. 2011PY038.

REFERENCES

Christiane C, Fertig, Fridrun P, Roger D (2004). Feasibility study for the rapid determination of the amylose content in starch by near-infrared spectroscopy. Eur. J. Pharm. Sci. 21:155-159.

Hansen WG, Wiedemann SCC, Snieder M, Wortel VAL (2000). Tolerance of near infrared calibrations to temperature variations: a practical evaluation. J. Near Infrared Spectrosc. (8):125-132.

Jaillais B, Pinto R, Barros AS, Rutledge DN (2005). Outer-product analysis (OPA) using PCA to study the influence of temperature on NIR spectra of water. Vibrat. Spectrosc. (39):50-58.

Lee JC, Yoon YH, Kim SM, Pyo BS, Hsieh FH, Kim HJ, Eun JB (2007). Rapid prediction of amylose content of polished rice by Fourier transform near-infrared spectroscopy. Food Sci. Biotech. 16(3):477-481.

Peinado AC, van den Berg F, Blanco M, Bro R (2006). Temperature-induced variation for NIR tensor-based calibration. Chemom. Intell. Lab. Syst. (83):75-82.

Shao Y, Zhao C, He Y, Bao Y (2009). Application of infrared spectroscopy technique and chemometrics for measurement of components in rice after radiation. Trans. ASABE. 52(1):187-192.

Suykens JAK, Vandewalle J (1999). Least squares support vector machine classifiers. Neural Processing Lett. 9(3):293-300.

Suykens JAK, Brabanter J, Lukas L, Vandewalle J (2002). Weighted least squares support vector machines: robustness and sparse approximation. Neurocomputing 48(1):85-105.

Vapnik V (1999). An overview of statistical learning theory. IEEE Trans. Neural Networks 10(5):988-999.

Wand D, Xiong YM, Huang R, Wu HB, MIN SG (2010). Influence of temperature on near-infrared spectroscopic quantitative analysis models of compound emulsifiable concentrate. Chin. J. Analyt. Chem. 38(9):1311-1315.

Wang YC, Zhang XM, Shi CH (2001). Studies on the analyzing method of amylose content in rice. Chin. Agri. Sci. Bull. 17(5):30-32.