

Quantitative Analysis of Fat Content in Rice by Near-Infrared Spectroscopy Technique

H. L. Wang,¹ X. Y. Wan,¹ J. C. Bi,¹ J. K. Wang,² L. Jiang,¹ L. M. Chen,¹ H. Q. Zhai,² and J. M. Wan¹⁻³

ABSTRACT

Cereal Chem. 83(4):402–406

Fat content in rice is one of the most important nutritional quality properties. But the chemical analysis of fat content is time-consuming and costly and could result in poor reproduction between replicates. Near-infrared spectroscopy (NIRS) can solve those problems by providing a rapid, nondestructive, and quantitative analysis. Based on the NIRS technique and partial least squares (PLS) algorithm, four calibration models were established to quantitatively analyze fat content in brown rice grain and flour and milled rice grain and flour with 248 representative samples. The determination coefficients (R^2) of these calibration models were 0.79,

0.84, 0.89, and 0.91, respectively, with the corresponding root mean square errors 0.16, 0.14, 0.09, and 0.08%. The R^2 were 0.73, 0.81, 0.81, and 0.89 with the corresponding root mean square errors 0.17, 0.15, 0.12, and 0.09%, respectively, in cross validation. The R^2 were 0.62, 0.80, 0.81, and 0.87, respectively, with the root mean square errors 0.25, 0.31, 0.28, and 0.30% in external validation. These results indicate that the method of NIRS has relatively high accuracy in the prediction of rice fat content. The four calibration models established in the present study should be useful for nutrient quality improvement in rice breeding.

Rice is a staple cereal food for more than half of the world population. As one of the important components of rice nutrition quality, rice fat content is genetically controlled by multiple genes (Hyeon et al 1998; Hu et al 2004) and has a relatively higher heritability (60.9–68.25%) (Qi et al 1983). Fat content has great influence on rice appearance and eating quality although it is relatively low in rice cultivars (Qi et al 1983; Wu et al 1985; Chen et al 1998). High fat content is one of the characteristics of high quality rice and has become an important target trait in breeding programs. In previous studies, several traditional methods such as Soxhlet, Goldfish, and Soxtec with similar procedures of extracting fat (Matsler et al 2005) were widely used to analyze the fat content. However, based on these conventional methods, the samples need to be destroyed to powder, and much time is required for solubilizing and extracting the fat. Therefore, it is very difficult to determine the fat content of thousands of samples using these methods within a brief time period. Analyzing a large number of samples with these methods is sensitive to interlaboratory variability because of the complexity of these procedures.

With the advantages of nondestructive samples, high sensitivity, and fast speed, the NIRS technique has been widely used for qualitative and quantitative measurement in agriculture, food, and chemical industries for several decades. Near-infrared radiation is absorbed mainly by C-H, N-H, and O-H, which are the primary constituents of the organic compounds in plant tissues. The chemical constituents of the specific tissue determine the wavelength and amount of light absorbed. Every substance has its own proper infrared absorbance. Therefore, the absorbed spectrum of each sample can be used to evaluate the chemical composition of that material. Based on this characteristic, quantitative and qualitative analysis could be conducted. To improve the efficiency of rice quality breeding, the NIRS technique has been widely applied to predict many quality traits of rice, including amylose content (Bao et al 2001; Miryeong et al 2004; Wu et al 2004), protein content (Shuso et al 2003; Miryeong et al 2004), amino acid content (Wu et al 2002), gelatinization temperature (Delwiche et al 1996; Bao et al 2001), gel consistency (Bao et al 2001), and RVA parameters (Frederick et al 2002). Most of the models had

higher precision and accuracy and indicated that NIRS was an ideal technique for screening genetic resources in rice quality breeding programs

Calibration models have been reported for determining fat content in oilseed crops such as rapeseed (Velasco et al 1999) and soybean (Li et al 2004). Meanwhile, NIRS was used to predict fat content of cereal crops such as sorghum (Hicks et al 2002) and maize (Lu et al 2004; Wei et al 2005). Compared with other crops, fat content in rice is relatively low at 1.72–3.37% in brown rice and 0.09–1.52% in milled rice. So, it is difficult to measure the fat content using the chemical methods mentioned above. To our knowledge, there has been no report on the application of NIRS in the prediction of fat content in rice. In this study, calibration models combining NIRS and chemical analysis were established to determine fat content in brown and milled rice. The efficiency of these models with different spectra was compared. The spectral forms included brown rice grain and flour and milled rice grain and flour. The purpose of this study was to find whether the NIRS technique could replace chemical analysis to increase the efficiency of rice quality improvement.

MATERIALS AND METHODS

Materials and Sample Preparation

A total of 248 representative cultivars were analyzed by chemical methods and scanned by near-infrared spectrometry, among which 216 cultivars were from 18 provinces in China, and 32 from the International Rice Research Institute in The Philippines. The 248 cultivars included 136 japonica rice, 88 indica rice, and 24 glutinous rice. The germinated seeds were sown in a seedling bed on May 14, 2003, and after 30 days, the seedlings were transplanted to experimental fields of Jiangsu Academy of Agricultural Sciences, Nanjing, China. Using a spacing pattern of 20 cm (between rows) \times 13.3 cm (within rows), each experimental plot consisted of two rows each of 10 plants, grown in a randomized block design.

Before the chemical analysis was conducted, all seeds were stored at room temperature for three months after harvesting. Paddy rice was dehulled with the JLGJ4.5 dehusker (Zhejiang Taizhou Food Instrument Co. Ltd, China) to obtain 40 g of brown rice, which was further milled for 1 min by the JNMJ3 miller (Zhejiang Taizhou) to remove the embryo and aleurone completely. Brown head grain (10 g) and milled head grain (10 g), selected from the total brown and milled samples obtained were ground by using the JFS-13A flour mill (Zhejiang Taizhou). The brown or milled rice flour that passed through the 80-mesh sieve was used for the maximum solvent contact in this study.

¹ National Key Laboratory for Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, 210095, China.

² Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, 100081, China.

³ Corresponding author. Phone: +86-25-84396516. Fax: +86-25-84396516. E-mail: wanjm@njau.edu.cn

NIRS Analysis

A Fourier transform infrared spectrometer (model Vector/22N, Bruker Co., Germany) was used to collect the absorbance spectra of all these samples over a wavelength range of 12000–4000 cm^{-1} , with the resolution of 8 cm^{-1} . For each sample, 64 repetitive scans were automatically performed in computer memory and averaged, and then stored as a computer file. It took ≈ 1 min to scan one sample, and each sample was measured three times under the same conditions. When the sample is placed manually, it is necessary to ensure that the sample to be scanned are spread evenly and filled consistently in the sample cup. These collected spectrum data of brown rice grain, brown rice flour, milled rice grain, and milled rice flour were used to develop different calibration models.

Chemical Analysis

Using the Soxtec-Avanti 2050 total fat system (Foss Co., Denmark), the fat in brown rice and milled rice was extracted in 40 mL of petroleum ether for each sample. The system contained a control unit and an extraction unit. The extraction procedure of the control unit included three steps: 1) boiling 30 min, 2) rinsing 1 hr, and 3) solvent recovery 10 min. The boiling temperature on the control unit was set at 135°C. In the extraction unit, six samples could be extracted simultaneously. After scanning, brown rice flour and milled rice flour samples were put into the aluminum box and kept in a 50°C oven for 48 hr to remove most of the water, which otherwise could interfere with the organic solvent. Individual samples were weighed on a balance with a precision of 0.1 mg: 2.0 g for brown rice flour (W1) and 2.5 g for milled rice flour (W1). Then the sample was placed into a thimble. A defatted cotton plug was set on top of the sample to prevent any sample from boiling out of the thimble during extraction. The clean extraction cups were placed into the oven at 105°C for 30 min to remove any moisture, transferred to a desiccator to cool for 30 min, and finally weighed (W2). The thimbles were placed into the

extraction unit and lowered to immerse the sample into the solvent; 40 mL of petroleum ether was measured for each extraction cup. After pressing the “run” key, the Soxtec-Avanti 2050 total fat system will perform the extraction and solvent recovery automatically. The total extraction required ≈ 2 hr. The extraction cups were dried at 105°C for 30 min to remove any residual petroleum ether and leave only dry material, which represented the extracted fat. After drying, the cups were transferred to a desiccator to cool for 30 min and weighed (W3). The formula used to calculate the fat content was

$$\text{Fat content (\%)} = \frac{W_3 - W_2}{W_1} \times 100 \%$$

where W1 is the sample weight (g), W2 is the net weight (g) of extraction cup, and W3 is the total weight (g) of the extraction cup plus the residue. Each sample was measured at least twice to ensure that the relative error of the parallel results was $< 5\%$.

Calibration Models and Standard of Evaluation

The Quant 2.0 software (v.3.0, OPUS-NT spectroscopic software) was used to perform the optimization and could automatically check common frequency ranges in combination with several data preprocessing methods. The results of spectral data preprocessing, range of wave number, and ranks of the models set up would be given. These were the optimal conditions. The calibration models were established on the basis of the partial least squares (PLS) algorithm, which was performed by combining the calibration set spectra with the Opus software package.

The calibration models and validation sets to which the models were applied to predict fat content were evaluated using several statistic parameters: coefficient of determination (R^2) between predicted and chemical values; root mean square errors of calibration (RMSEC); root mean square errors of cross validation (RMSECV); and root mean square errors (RMSEV) of external

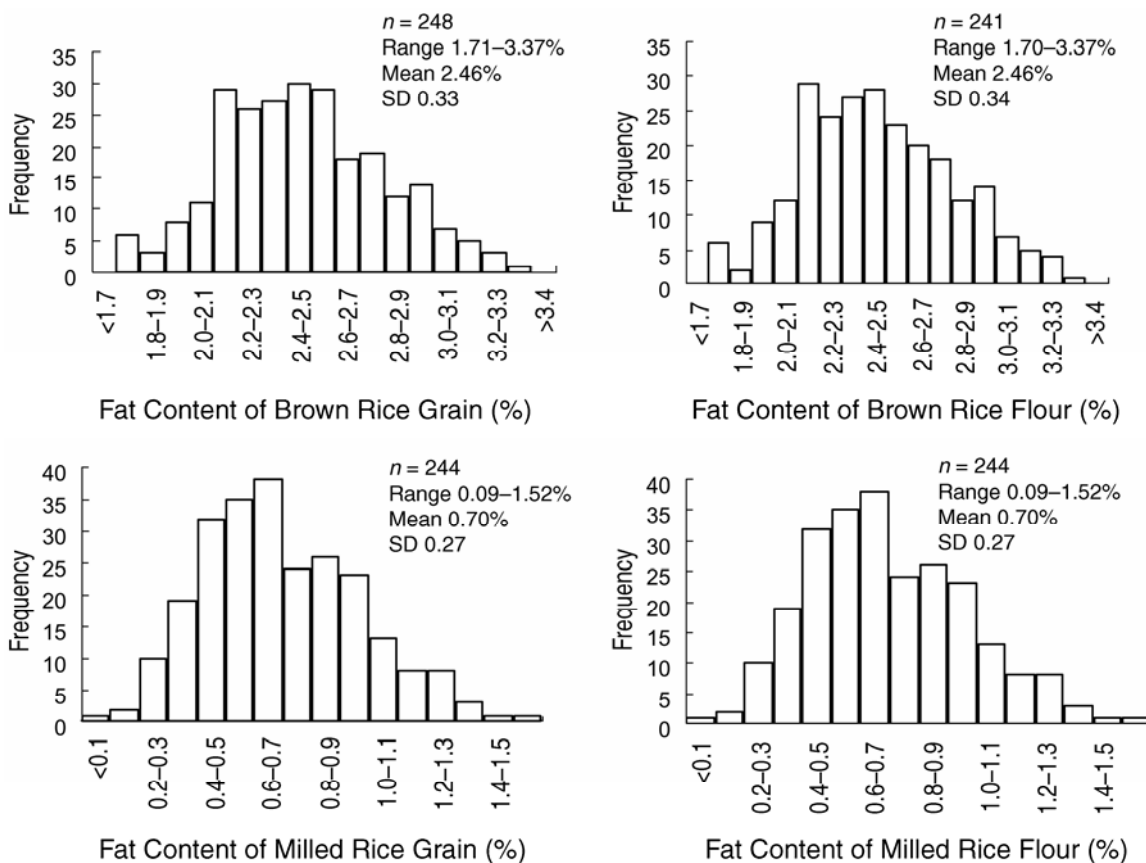


Fig. 1. Distribution and statistical parameters of chemical values of fat content in brown and milled rice.

validation. The criteria for the quality models are the higher value of R^2 and lower value of RMSE.

RESULTS

Chemical Analysis of Fat Content

Figure 1 showed the distributions and statistical parameters of chemical analysis for fat content in brown rice and milled rice. The fat contents were 1.70–3.37%, 1.71–3.37% with the same mean of 2.46% in the brown rice grain and flour sets, respectively, and both varied from 0.09 to 1.52% with the same mean of 0.70% in the milled rice grain and flour sets. The phenotypic distributions of the two traits showed continuous and approximately normal variations with wider ranges; they represented all cultivars presently used in rice production in China.

Optimal Conditions for NIR Calibration Models

PLS was used to develop calibration models from the original spectra sets. In general, the best model was found by trial and error. But the Quant software, in combination with several data preprocessing methods, could also automatically check common frequency ranges. Table I listed all the optimal parameters for establishing NIR calibration models.

Rank is defined as the number of PLS vectors. The criterion to determine the optimum rank is the RMSEP from analysis of the cross validation. The RMSEP is depicted against the rank used in each model, the minimum RMSEP implies the optimum rank. In this research, the optimum ranks were 9, 8, 10, and 9 for brown rice grain, brown rice flour, milled rice grain, and milled rice flour, respectively. Data preprocessing is an important calibration step wherein variations in offset or different linear baselines can

be eliminated. For the spectra of brown and milled rice grain samples, straight line subtraction was needed to optimize the spectrum. For the brown and milled rice flour samples, with sample particles <0.5 mm and relatively uniform form compared with the grain samples, there was no spectral data preprocessing.

Development of NIRS Calibration Models

The samples were randomly divided into two groups. Calibration sets consisted of 208 brown rice grain, 201 brown rice flour, 204 milled rice grain, and 204 milled rice flour samples. External validation sets contained 40 samples for each set. To recognize and cancel outliers, significant probability was 0.01. After the F test, concentration and spectral outliers were removed from the calibration set. As a result, the ultimate calibration sets were adjusted to 198, 195, 201, and 201 for brown rice grain, brown rice flour, milled rice grain, and milled rice flour, respectively.

As shown in Table II, the R^2 values of the calibration set were sufficiently high (0.91 and 0.89) for determination of fat content in milled rice. The corresponding RMSEC values were 0.09 and 0.08%, which were lower than those of brown rice. The R^2 of brown rice grain was the lowest and the corresponding RMSEC is the largest in the four models, possibly because the different appearance of the brown rice grain affected the spectral data.

Evaluation of Calibration Models

Two methods, cross validation and external validation, were used to evaluate the performance of the calibration models in this study. In cross validation, one sample was excluded from the entity of samples at the beginning of the calibration. This sample was used for the validation. The remaining samples were used to calibrate the system. This cycle was repeated until all samples had been used for validation once. In external validation, a set of independent samples was used to validate the models. Statistical parameters of cross validation and external validation are shown in Table III. Compared with the models of the grain spectra, the models of the flour spectra had better effects of prediction. For all the four models, the generally similar effects were observed in the cross validation and external validation, except the model set up with brown rice grain spectra. The range of samples in the cross validation were the same as those observed in the calibration set. In the external validation, the ranges were slightly narrower in all cases, except one sample in the brown rice flour; the chemical value of 1.70% in this sample was similar to the predicted value of 1.83%. Thus, the two validations were both suitable to examine the calibration models.

TABLE I
Optimal Conditions for Establishing NIRS Calibration Models of All Samples

Sample Form	Optimal Conditions for Models		
	Spectral Data Preprocessing	Range of Wave No. (cm ⁻¹)	Rank
Brown rice grain	Straight line subtraction	7502.5–4246.9	9
Brown rice flour	No spectral data preprocessing	7502.5–5446.5	8
Milled rice grain	Straight line subtraction	7502.5–4597.9	10
Milled rice flour	No spectral data preprocessing	7502.5–5446.5 4601.8–4246.9	9

TABLE II
Calibration Statistics for Fat Content in Brown and Milled Rice

Spectrum Model	Samples	Mean (%)	Range (%)	R^2	RMSEC (%) ^a
Brown rice grain	198	2.47	1.72–3.37	0.79	0.16
Brown rice flour	195	2.47	1.72–3.37	0.84	0.14
Milled rice grain	201	0.69	0.09–1.52	0.89	0.09
Milled rice flour	201	0.69	0.09–1.52	0.91	0.08

^a Root mean square errors for calibration.

TABLE III
Validation Statistics for Fat Content in Rice

Spectrum Model	Cross Validation					External Validation				
	N	Mean (%)	Range (%)	R^2	RMSECV (%) ^a	N	Mean(%)	Range (%)	R^2	RMSEEV (%) ^b
Brown rice grain	198	2.47	1.72–3.37	0.73	0.17	40	2.45	1.86–3.21	0.62	0.25
Brown rice flour	195	2.47	1.72–3.37	0.81	0.15	40	2.42	1.70–3.29	0.80	0.31
Milled rice grain	201	0.69	0.09–1.52	0.81	0.12	40	0.72	0.11–1.30	0.81	0.30
Milled rice flour	201	0.69	0.09–1.52	0.89	0.09	40	0.73	0.11–1.30	0.87	0.30

^a Root mean square errors for cross validation.

^b Root mean square errors for external validation.

DISCUSSION

The NIRS analysis has the advantage of being fast, economical, and environmentally safe. Once the effective calibration model has been set up, it would be possible to perform the rapid and effective measurement of multiple traits simultaneously. So the NIRS technique is particularly useful for quality trait evaluation in crop breeding programs, in which thousands of samples need to be screened in time. So far, NIRS has been applied to analyze oil content in crops such as rapeseed (Velasco et al 1999), soybean (Li et al 2004), sorghum (Hicks et al 2002), and maize (Lu et al 2004; Wei et al 2005). The R^2 in these validations were all >85% and RMSECV were all >0.33% except the model constructed by Hicks et al (2002). However, there have been no reports on developing NIRS calibration models for predicting fat content in rice. In this study, 248 rice cultivars with fat content of brown rice at 1.72–3.37% and that of milled rice at 0.09–1.52% were used to set up four calibration models (Table I). In the calibration sets, the R^2 values were 79–91% and RMSEC values were 0.08–0.16%. In the external validation, the R^2 values were 62–87% and RMSECV values were 0.25–0.31. Compared with many previous studies, these results obtained here should be better due to the wider range of rice fat content and the larger number of samples.

Casler et al (1985) and Windham et al (1987) both proposed that particle size of samples used for establishing the NIRS models had an important effect on NIR absorbance spectrum. Increasing particle size could result in an increased apparent path length and therefore change the spectra of samples (Zhao et al 2004). Similarly, our results also demonstrated that the models developed by using flour sample forms were relatively more accurate than those using grain sample forms (Table II). In the present study, the particle size of flour samples is relatively uniform because its diameter is normally <0.5 mm. However, the grain samples contain short-, medium-, and long-grain rice, which lead to great differences on spectral forms among samples and ultimately reduce the accuracy of the calibration model. On the other hand, we found that the model set up with milled rice grain had better effects on evaluating the fat content than that with brown rice grain, possibly because the color of the brown rice grain affected the absorbance spectrum. In other words, the spectra data of milled rice grain were more representative of the true values.

The quality of models could be influenced by many factors. In this study, there were at least three factors playing important roles in calibration. The first was the accuracy and precision of the data from chemical and spectra analysis. The validity of chemical and spectra data could be improved by strictly performing the operation procedures. The second was having optimal conditions for the calibration model. The best conditions included spectral data preprocessing, range of wave number, and ranks (Chu et al 2004; Ma et al 2004; Wu et al 2004). In those studies, the best conditions were empirically confirmed by trial and error. In this study, the Quant 2 software was used to perform the optimization and it could automatically check for the best conditions (Table I), which ensured the efficiency of the models. The last factor was the quality of samples used in the calibration set. To optimize the model, chemical and spectral outliers in the calibration set were removed from the study. Outliers would result from measurement errors due to the instruments, the special impurities in the sample not found in other samples, or the perturbations caused by additional unknown components.

Based on the above concept, the two more accurate calibration models established by using flour samples could be used with a sufficient number of samples for highly accurate results. The calibration model set up with brown grain was more suitable in rice breeding programs. One reason was that the data demand was less stringent because it screened the extreme materials from a large number of rice germplasm resources. Also, the savings in time and costs, together with a reasonable reliability, were more

important than a very high accuracy in early stages of rice breeding programs. In this case, the R^2 and RMSEC of the model based on brown rice grain are enough to predict fat content. Thus brown rice grain was considered as an ideal sample to conveniently screen rice germplasm resources in rice breeding programs. Even when a more accurate selection is desired, a combination of NIRS models of brown rice grain and flour or milled rice grain and flour can save considerable time and costs. For the most important breeding materials, chemical methods could be further used to confirm fat content after prescreening with NIRS. Additionally, because milled rice is widely used in the food industry, it is necessary to grade rice cultivars according to the quality index. The precision and accuracy of the two models of milled rice grain and flour for determining fat content are sufficiently high to enable classification of rice cultivars.

CONCLUSIONS

Calibration models based on rice flour samples were more accurate than those based on grain samples. Assuming that grinding cost is low and the sample size is sufficient, flour samples are recommended in prediction of fat content. Though the model set up with brown rice grain was relatively poor, it should be suitable for prescreening the cultivars in early generation in rice breeding programs and large screening of rice germplasm resources. Final screening with flour samples or chemical analysis could be done when higher accuracy is needed.

ACKNOWLEDGMENTS

We are very grateful for comments and suggestions on the manuscript from the editors and the anonymous reviewers. This research is supported by the 973 Program of China (2004CB1172004), the National Natural Science Foundation of China (No.30500315), the National High Technology Research and Development Program of China, Jiangsu Province High Technology Program (BG2004303), and Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT).

LITERATURE CITED

- Bao, J. S., Cai, Y. Z., and Corke, H. 2001. Prediction of rice starch quality parameters by near-infrared reflectance spectroscopy. *J. Food Sci.* 66:936-939.
- Casler, M. D., and Shenk, J. S. 1985. Effect of sample grinding size on forage quality estimates of smooth brome grass clones. *Crop Sci.* 25:167-170.
- Chen, J. G., and Zhu, J. 1998. Genetic analysis of fat content in *indica-japonica* intersubspecific hybrid rice. *J. Trop. Subtrop. Bot.* 6:347-351.
- Chu, X. L., Yuan, H. F., Wang, Y. B., and Lu, W. Z. 2004. Developing robust near-infrared calibration models. *Spectrosc. Spect. Anal.* 24:666-671.
- Frederick, M., and Franklin, E. 2002. Determination of Rapid Visco Analyser parameters in rice by near-infrared spectroscopy. *Cereal Chem.* 79:563-566.
- Hicks, C., Tuinstra, M. R., Pedersen, J. F., Dowell, F. E., and Kofoid, K. D. 2002. Genetic analysis of feed quality and seed weight of sorghum inbred lines and hybrids using analytical methods and NIRS. *Euphytica* 127:31-40.
- Hu, Z. L., Li, P., Zhou, M. Q., Zhang, Z. H., Wang, L. X., Zhu, L. H., and Zhu, Y. G. 2004. Mapping of quantitative trait loci (QTLs) for rice protein and fat content using doubled haploid lines. *Euphytica* 135:47-54.
- Hyeon, J. K., Yong, G. C., Young, T. L., Young, D. K., and Moo, Y. E. 1998. QTL mapping of genes related with grain chemical properties based on molecular map of rice. *Kor. J. Crop Sci.* 43:199-204.
- Li, N., Min, S. G., Qin, F. L., Zhang, M. X., and Ye, S. F. 2004. Nondestructive analysis of protein and fat in whole kernel soybean by NIR. *Spectrosc. Spect. Anal.* 24:45-49.
- Lu, Y. J., Qu, Y. L., Pu, R. G., and Zhang, J. 2004. The algorithm of eliminating the similar samples in the process of calibration and prediction. *Spectrosc. Spect. Anal.* 24:158-161.
- Ma, X., Wan, G. Y., Wen, Y. D., Xie, L. H., Cui, Y. H., Zhang, J., and Li, H. B. 2004. Determination of chemical components in tobacco leaves by FT-NIR spectroscopy: Study of influence of spectral ranges on PLS

- modeling. *Spectrosc. Spect. Anal.* 124:444-446.
- Matsler, A. L., and Siebenmorgen, T. J. 2005. Evaluation of operating conditions for surface lipid extraction from rice using a Soxtec system. *Cereal Chem.* 82:282-286.
- Sohn, M., Barton, F. E., McClung, A. M., and Champagne, E. T. 2004. Near-infrared spectroscopy for determination of protein and amylose in rice flour through use of derivatives. *Cereal Chem.* 81:341-344.
- Qi, Z. B., Li, B. J., Yang, W. G., and Wu, X. F. 1983. A study on the genetic of exterior quality and fat of the rice grains. *Acta Genet. Sinica* 10:452-458.
- Shuso, K., Motoyasu, N., Kazuhiro, T., and Kazuhiko, I. 2003. Development of an automatic rice-quality inspection system. *Comput. Electron. Agric.* 40:115-126.
- Velasco, L., Möllers, C., and Becker, H. C. 1999. Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (*Brassica napus* L.) by near-infrared reflectance spectroscopy. *Euphytica* 106:79-85.
- Wei, L. M., Jiang, H. Y., Li, J. H., Yan, Y. L., and Dai, J. R. 2005. Predicting the chemical composition of intact kernels in maize hybrids by near infrared reflectance spectroscopy. *Spectrosc. Spect. Anal.* 25:1404-1407.
- Windham, W. R. 1987. Influence of grind and gravimetric technique on dry matter determination of forages intended for analysis by near infrared reflectance spectroscopy. *Crop Sci.* 27:773-776.
- Wu, J. G., and Shi, C. H. 2004. Prediction of grain weight, brown rice weight and amylose content in single rice grains using near-infrared reflectance spectroscopy. *Field Crop Res.* 87:13-21.
- Wu, S. Z., Huang, C. W., Ou, L. C., and Liu, J. Z. 1985. Studies on varietal characteristics in culture of *Oryza Sativa*. III. A study on grain quality character of rice varieties. *Sci. Agric. Sinica* 18:1-7.
- Wu, J. G., Shi, C. H., and Zhang, X. M. 2002. Estimating the amino acid composition in the milled rice powder by near-infrared reflectance spectroscopy. *Field Crop Res.* 75:1-7.
- Zhao, L. L., Li, J. H., Zhao, L. D., and Yan, Y. L. 2004. Influence of FT-NIR spectrometer scanning requirements on the math model's precision. *Spectrosc. Spect. Anal.* 24:12-44.

[Received August 22, 2005. Accepted April 10, 2006.]