Predicting nitrogen status of oilseed rape based on Vis-NIR spectroscopy

HUANG Min[1], HE Yong[1], ZHOU Weijun[2], CEN Haiyan[1]

[1]*College of Biosystems Engineering and Food Science, Zhejiang University, Hang zhou 310029, China*
[2] *College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310029, China*

Abstract

Visible and near infrared spectroscopy had been used to develop a model, which was combined with artificial neural network (ANN) and partial least square regression (PLS) method, to predict the nitrogen status of oilseed rape. Spectra tests were performed on the canopies of 150 rapeseed samples in the field using a spectrophotometer (325-1075nm). 5 optimal PLS principal components were determined by PLS analysis with cross-validation. They were selected as the input of BP neural network to establish the prediction model. The node number of input layer, hidden layer, and output layer was 5, 5, and 1, respectively. 110 samples were used as training set and the left 40 samples formed prediction set. The result showed that the prediction performance was excellent with the correlation value of 0.95405, higher than the result (0.8764) obtained only by using PLS method. Most of the relative standard deviation (RSD) was under 5% and the accuracy of prediction reached 95%. Thus, it is concluded that the proposed PLS-ANN model for the spectroscopic estimation of oilseed rape N status is superior to other existing spectroscopic methods based on Vis/NIRS.

Key words: Oilseed rape, nitrogen, Vis-NIR spectrum, PLS, ANN

Introduction

One of the most important aims of precision farming is for a decrease in fertilizer required for the same yield and/or higher yields with the use of the same amounts of fertilizer (Auernhammer, 2001). So precision farming requires close monitoring crop nutrition and determining fertilizer requirements to enable crops to grow faster (Chang & Robison, 2003). Nitrogen is an essential element for plant growth and is frequently the major limiting nutrient in most agricultural soil. Estimating the crop nitrogen status is an important research in precision farming. Traditionally, soil testing, plant tissue analysis, and long-term field trials have been used for assessing N availability for crops (Daughtry et al., 2000). But these methods are relatively expensive and time consuming, and would damage the crop. These can't satisfy the requirements of precision farming.

A chlorophyll meter (Minolta SPAD-502) has been developed to estimate the nitrogen status of crops quickly and nondestructively. The SPAD meter was initially developed in Japan to diagnose foliar N status and determine N fertilizer requirements of rice (Chubachi et al., 1986). Since then, several publications reported the application of SPAD meter to measure leaf greenness or chlorophyll content in fruit trees (Idso et al., 1996). Chlorophyll concentrations have been assessed with the SPAD meter in red maple (Van den Berg & Perkins, 2004). The SPAD meter can be used as a rapid diagnostic tool to chlorophyll concentration. The disadvantage of SPAD meter is that it requires taking 30 or more readings from representative plants in each area, which can be time-consuming for large fields that have spatial variations in N. And the result of estimating the N status in a large field is not accurate enough. So it is unpractical to monitor plant N status in a large field by using SPAD meter.

Nitrogen deficiency will affect spectral reflectance in both visible and near infrared (NIR) regions. Graeff reported that reflectance near 510, 516 and 546 nm was good for detecting corn plant N deficiencies (Graeff & Claupein, 2003). N deficiency usually decreases leaf Chl concentration resulting in increase of leaf reflectance in both green and red edge ranges (Daughtry et al., 2000; Zhao et al., 2003). Although many studies have found that nondestructive measurements of leaf or canopy reflectance can be used for detecting N-deficient stress in corn, rice, wheat, and sorghum (Wang et al., 2006; Zhao et al., 2005), few studies have reported in estimation N status in oilseed rape by Vis-NIR spectra.

The objectives of this study were to estimate the N status in oilseed rape by using the visible and near infrared spectroscopy and Minolta SPAD-502, and develop algorithms or quantitative relationship between canopy spectral reflectance and leaf Chl or N concentrations for oilseed rape plant N monitoring.

Materials and Methods

Plant material: Oilseed rape was planted in the field of Zhejiang University on November 18, 2005, and was grown in a sandy loam soil devoted to long-term N application rate studies. The field was maintained at high-test levels of P and K. Treatments included six rates of 0, 45, 90, 135, 180, 270 kg N/ha, and the recommended rate was 180 kg N/ha. Each N treatment was replicated twice. There were 60 to 70 plants in each plot. The examination was done on March 15, 2006, during the growing season under clear sky, in the fields. 13 plants in each plot were chosen as the samples, totally 150 samples were measured in this experiment.

Canopy Reflectance Measurement: In the research, a handheld field spectroradiometer (FieldSpec Pro FR (325-1075 nm)/A110070), Trademarks of Analytical Spectral Devices, Inc. (ASD) was used to take the canopy reflection spectra of rapeseed plants. The spectroradiometer had high sensitivity from 325 to 1075 nm. The Handheld FieldSpec was placed at a distance above the top of oilseed rape plant about 20cm. So the diameter of the detected oilseed rape canopy was about 9cm. For each sample, three canopy reflection spectra were taken with the spectroradiometer. The scan number for each spectrum was set 30 at exactly the same position. Other software was used, such as ASD View Spec Pro, Unscramble V9.2 (CAMO, PROCESS, AS, OSLO, Norway) and DPS (data procession system for practical statistics)

SPAD Reading: A Minolta SPAD-502 chlorophyll meter was used to measure chlorophyll concentration of the canopy leaves in the view area of the spectroradiometer. Measurements of the canopy consisted of 30 readings. According to the area of the leaf, the points were increased or decreased. The mean of the 30 readings was used as the chlorophyll concentration of that sample. The SPAD Values of the samples varied from 35.5 to 55.7.

PLS analysis: Vis-NIR spectra were difficult to interpret directly because of the overlap of weak overtone and combinations of fundamental vibration bands. As a result, multivariate calibration was required for qualitative analysis of sample varieties by Vis/NIRS. Various calibration methods had been used to relate near-infrared spectra with measured properties of materials. Principal Components Regression (PCR), Partial Least Squares (PLS), Stepwise Multiple Linear Regression (SMLR), artificial neural networks were the most used multivariate calibration techniques for NIRS (Annia et al., 2003; He et al., 2006). PLS was a bilinear modeling method where the original independent information (X-data) was projected onto a small number of latent variables (LV) to simplify the relationship between X and Y for predicting with the smallest number of LVs. The optimal number of PLS components that optimized the predictive ability of the model should be determined. This choice was typically made with the use of cross-validation. Prediction residual sum of squares (PRESS) or total residual variance (RV) for the test samples was determined as a function of the number of PLS components retained in the regression model that was formed with the training data. The value of PRESS with the principal component number from 1 to d was obtained. The number of PLS components which produced minimum PRESS error was used as the number of latent variables. Another, total residual variance was computed as the sum of squares of the residuals from all the variables, divided by the number of degrees of freedom.

BP neural network model: BP neural network was one of the most popular neural network topologies. It had the advantages of being easy to understand and easy to implement. It was a one-way multilayer feed-forward network (He et al., 2005). This neural network had one or more hidden layers nodes besides input node and output node, and there was no coupling in the same hidden layer. The input signals could be passed through hidden layer node from the input node to the output node. The output of on layer became the input to the following layer. Fig. 1 showed architecture of BP neural network with three layers.

In this study, the samples were randomly separated into two groups: 110 oilseed rape samples were used to develop the calibration models and the remaining 40 samples were used to prediction set. The architecture of BP neural network with three layers was constructed. The node of input layer, hidden layer, and output layer was 5, 5, and 1 (the SPAD value), respectively. The input values were 5 principal components from PLS analysis.

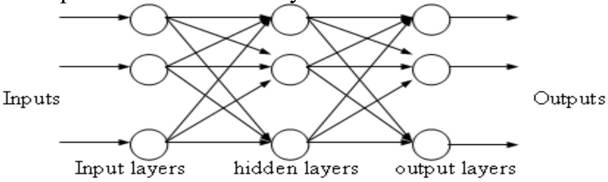

Fig. 1. Schematic of the BP neural network

Results and Discussion

The average reflectance spectra from 400 to 1000 nm were showed in Fig. 2 for five representational selected samples. Excluding the close overlapping of the wavebands at the 530-630nm wavelengths, especially near 700nm wavelengths, the others could be used in estimating N status via Vis-NIR spectroscopy.

PLS Used Only: PLS with cross-validation and Vis-NIR spectra were used to develop calibration models of the SPAD values of samples. The number of principle components was determined by the PRESS (predicted error sum of squares) function in order to avoid overfit of the models (Qi et al., 2003). As a result, 110 samples were used to develop calibration models and remaining 40 samples were used to validation set.

In this study, the calibration and validation results only established by PLS. In calibration part, SEC was 1.660104, R was 0.901010 and RMSEC was 1.654909. And in validation part, SEP was 1.705191, R was 0.895253 and RMSEP was 1.699854. PLS prediction result for SPAD value was presented in scatter plots in Fig. 3. The ordinate and abscissa axes represented the predicted and measured fitted values of the SPAD. When the model was used to predict the 40 unknown samples, prediction result was good (r= 0.876392), standard error of prediction (SEP) was 1.691751, and RMSEP was 1.727854.

ANN Combined with PLS: PLS was performed on the 150 oilseed rape samples in the training set with the whole wavelengths from 400nm to 1000nm. Table 1 showed the explained rate plot, standard error of calibration (SEC) and correlation (R) for the first eight principal components (PCs). It was clear that the first five PCs could explain over 98% of the total population variance and the remainders could account for a little. The model with five principal components had a high correlation coefficient (R) 0.903474. At the same time, the standard error of calibration (SEC) 1.640374 was also very low.

Therefore, the first five PCs were appropriate for characteristic description of the oilseed rape canopy spectral curve.

Thus, 5 principal components were set as input neurons of ANN to establish the prediction model. The node of input layer, hidden layer, and output layer was 5, 5 (chose by compared the prediction results with different integral value from 3 to 8), and 1 (the value of SPAD). 110 samples were used as training set and reminder samples (total 40 samples) formed prediction set. The goal error was set as 0.0001, and the speed of learning was 0.2, the time of training was 3000. Prediction result for SPAD value was presented in scatter plots as Fig. 4 shown. The ordinate and abscissa axes represented the predicted and measured fitted values of the SPAD. The correlation between the measured and predicted values showed an excellent prediction performance with the value of 0.95405, higher than the result (0.8764) obtained only by used the algorithm of PLS. The regressed residual error was 0.00013567 and most of the relative standard deviation (RSD) was under 5%. Generally, the accuracy of prediction reached 95%.

Table 1. Main parameters in PLS calibration model

PLS	Explained variance	R	SEC
1	72%	0.8466	2.0366
2	89%	0.9010	1.6601
3	92%	0.9019	1.6529
4	94%	0.9029	1.6443
5	98%	0.9034	1.6403
6	99%	0.9073	1.60855
7	99%	0.9113	1.5756
8	100%	0.9127	1.5634

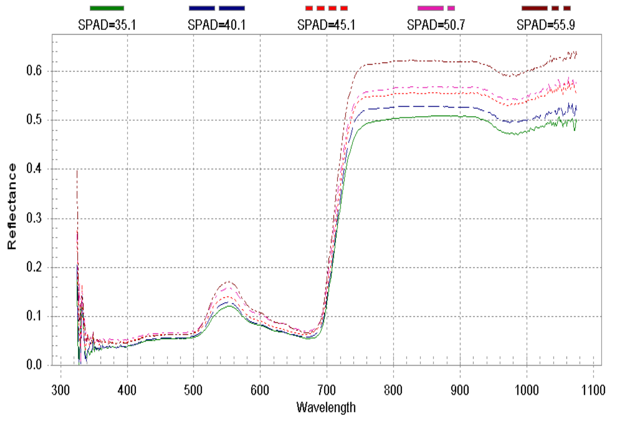

Fig. 2. Vis/ NIR canopy reflectance spectra of the oilseed rape

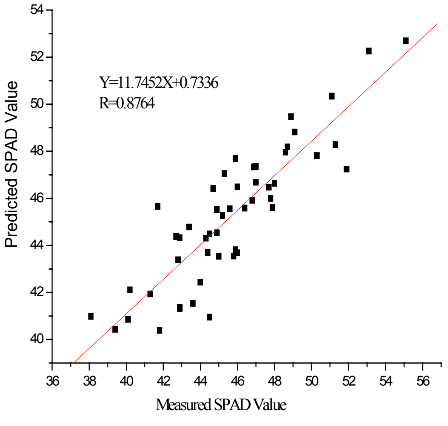

Fig. 3. Vis/NIR prediction results from the PLS models for SPAD value

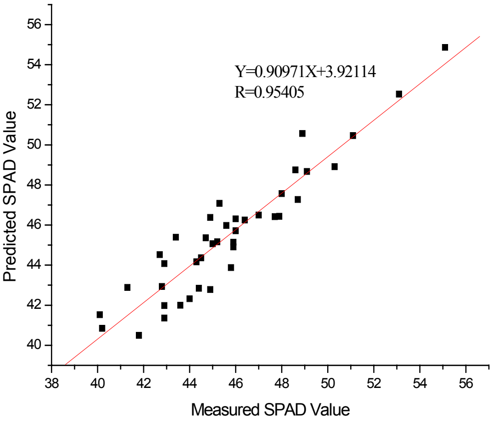

Fig. 4. Correlation coefficient between measured and predicted SPAD value by using PLS-ANN

Conclusions

This study was conducted to find relationships between the oilseed rape N status and their canopyspectral characteristics, as a preliminary step to develop a N status estimation sensor. The major findings from this research were: Vis/NIR Spectrum technique with the advantage of low cost, high efficiency, fast analytical speed, ease of operation, non-destruction and limited preparation, had been widely used in many fields. The oilseed rape samples at different N status showed significant differences in their spectra. The result demonstrated that it was possible to develop a non-destructive technique for estimation of N status based on canopy reflectance spectra of oilseed rape. A new combined method that artificial neural network (ANN) combined with partial least square regression (PLS) method, was used to establish the estimating model and showed an excellent prediction performance. After the PLS process carried out with all 150 samples, the optimal number of PLS components was determined as 5 by cross-validation. The 5 PLS principal components could explain most of the spectra information from

400nm to 1000nm of the oilseed rape N status, and they were selected as the input of BP neural network to establish the prediction model. The node of input layer, hidden layer, and output layer was 5, 5, and 1. 110 samples were used as training set and reminder samples formed prediction set. After the training, the result showed that the correlation between the measured and predicted values showed an excellent prediction performance with the value of 0.95405, higher than the result (0.8764) obtained only by using PLS method. Thus, we could conclude that the proposed PLS-ANN model for the spectroscopic estimation of oilseed rape N status was superior to other existing spectroscopic methods based on Vis/NIRS.

Acknowledgements

This study was supported by the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE, P. R. China, Natural Science Foundation of China (Project No: 30270773), Specialized Research Fund for the Doctoral Program of Higher Education (Project No: 20040335034), Natural Science Foundation of Zhejiang (Project No: RC02067) and Science and Technology Department of Zhejiang Province (Project No. 2005C21094).

References

Annia G.P., Antihus H.G., He Y. (2003). Advances in measurement and application of physical properties of agricultural products. Transactions of the ASAE **19**, 7-11.

Auernhammer H. (2001). Precision farming - the environmental challenge. Computers and Electronics in Agriculture **30**, 31-43.

Chang S.X., Robison D. J. (2003). Nondestructive and rapid estimation of hardwood foliar nitrogen status using the SPAD-502 chlorophyll meter. Forest Ecology and Management **181**, 331-338.

Chubachi T., Asano I., Oikawa T. (1986). The diagnosis of nitrogen nutrition of rice plants (Sasanishiki) using chlorophyll meter. Soil Sci. Plant Nutr. **57**, 190-193.

Daughtry C.S.T., Walthall C.L., Kim M.S., Colstoun E.B., Mcmurtrey J.E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment **74**, 229-239.

Graeff S., Claupein W. (2003). Quantifying nitrogen status of corn (Zea mays L.) in the field by reflectance measurements. European Journal of Agronomy **19**, 611-618.

He Y., Feng S.J., Deng X.F., Li X.L. (2006). Study on lossless discrimination of varieties of yogurt using the Visible/NIR-spectroscopy. Food Research International **39**, 645-650.

He Y., Zhang Y., Xiang L.G. (2005). Study of application model on BP neural network optimized by fuzzy clustering. Lecture Notes in Artificial Intelligence **3789**, 712-720.

Idso S.B., Kimball B.A., Hendrix D. L (1996). Effects of atmospheric CO2 enrichment on chlorophyll and nitrogen concentrations of sour orange tree leaves. Environmental and Experimental Botany **36**, 323-331.

Qi X.M., Zhang L.D., Du Z.L. (2003). Quantitative analysis using NIR by buiding PLS-BP model. Spectroscopy and Spectral Analysis **23**, 870-872.

Van den Berg A.K., Perkins T.D. (2004). Evaluation of a portable chlorophyll meter to estimate chlorophyll and nitrogen contents in sugar maple leaves. Forest Ecology and Management **200**, 113-117.

Wang S.H., Zhu Y., Jiang H.D., Cao W.X. (2006). Positional differences in nitrogen and sugar concentrations of upper leaves relate to plant N status in rice under different N rates. Field Crops Research **96**, 224-234.

Zhao D.L., Reddy K.R., Kakani V.G., Read J.J., Carter G.A. (2003). Corn (Zea mays L.) growth, leaf pigment concentration, photosynthesis and leaf hyperspectral reflectance properties as affected by nitrogen supply. Plant and Soil **257**, 205-217.

Zhao D.L., Reddy K.R., Kakani V.G., Reddy V.R. (2005). Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorhum. European Journal of Agronomy **22**, 391-403.

AGRONOMY

Crop Physiology