

Model for predicting the nitrogen content of rice at panicle initiation stage using data from airborne hyperspectral remote sensing

Chanseok Ryu^{a*}, Masahiko Suguri^a, and Mikio Umeda^a

^a*Environmental Science and Technology, Graduate School of Agriculture,
Kyoto University, Sakyo-ku, Kitashirakawa Oiwake-cho, Kyoto 606-8502, Japan*

Abstract

Airborne hyperspectral remote sensing was used to provide data for a general-purpose model for predicting the nitrogen content of rice at panicle initiation stage using three years of data. There were significant differences between the vegetation data which were affected by the uptake of nitrogen from the soil depending on weather conditions. Therefore, the reflectance values obtained for one year may exhibit a different trend, due to the lack of vegetation. When the partial least squares regression (PLSR) models were estimated using all combination of three-year data, except for the model incorporating the data from 2005, correlation coefficients (r) were greater than 0.758, and the root mean squared error (RMSE) of prediction of the full-cross validation was less than 0.876 g m^{-2} . The accuracy of the 2003-2004-2005 model was determined using five latent variables (PCs), with $r = 0.938$ and $\text{RMSEP} = 0.774 \text{ g m}^{-2}$. There were two different patterns for the regression coefficients associated with the NIR or red-edge regions. When the 2003-2004 model was validated using the data from 2005, the prediction error of the PLSR model was 1.050 g m^{-2} . This became 2.378 g m^{-2} for the 2003-2005 model using the data from 2004 and 5.061 g m^{-2} for the 2004-2005 model with the data from 2003. There were similarities and differences for each latent variable between the 2003-2004 model and the 2003-2004-2005 model. The 2003-2004-2005 model might be more suitable for use as a general-purpose model, because it is possible to consider and validate all of the three years data.

Keywords: airborne hyperspectral remote sensing, nitrogen content, rice, partial least square regression (PLSR)

*Corresponding author. Tel.: +81-75-753-6317; Fax: +81-75-753-6167; E-mail: ryu@elam.kais.kyoto-u.ac.jp

1. Introduction

Precision farming or precision agriculture (PA) is a management practise made possible by the advent of suitable information technologies. It provides a framework within which agricultural managers can more accurately understand and control what happens in their farms (Blackmore, 1994). It requires cost-effective and time-effective methods of identifying various crop types to evaluate the crop conditions and recommend suitable measures. Information on the type of crop cultivar is particularly essential in the proper implementation of precision farming methods (Rama Rao et al, 2007). Thus, uncertainties in soil nitrogen supply and crop nitrogen demand, as cost methods, present challenges to farmers selecting nitrogen fertiliser rates (Lobell, 2007) and land levelling based on nitrogen fertiliser rates and application timing (Walker et al, 2003). Regarding time-effective methods, the practise of PA requires the development of accurate and reliable crop monitoring techniques to provide information on the spatial and temporal variations in key agronomic parameters (Steven, 2004). Remote sensing has great potential for several applications because it enables wide-area, non-destructive, and real-time acquisition of information on ecophysiological plant conditions (Inoue, 2001). Remotely sensed data, obtained either by satellite or aircraft, can provide a set of detailed and spatially distributed data on plant growth and development (Plant et al, 2000).

When remote sensing technology is applied in PA multi-spectral sensors are mainly used. Broadband data from multi-spectral sensors that use average spectral information over broad bandwidths often loose of critical information that is available in specific narrow bands (Blackburn, 1998). Thus, more research is being conducted using hyperspectral remote sensing that presents greater spectral information. Such research has suggested and identified the potential utility of hyperspectral remote sensing, especially for upland crops. There are two approaches for use of hyperspectral remote sensing in PA, via vegetation indices using narrow bands and via modelling using regression analysis.

Narrow-band vegetation indices for water stress in potatoes (Ray et al, 2006), for nitrogen and water stress in corn (Strachan et al, 2002), for chlorophyll a and b contents of vineyards (Zarco-Tejada et al, 2005) and for crops such as cotton, potatoes, soybeans, corn, and sunflower (Thenkabail et al, 2000) have been evaluated using hyperspectral reflectance. Regression models have been developed from hyperspectral reflectance measurement, to determine canopy biomass and nitrogen status in wheat (Hansen and Schjoerring, 2003), grassland yield and quality (Schut et al, 2006), water stress in cotton (Detar et al, 2006), and chemical fertiliser stress in barley (Jorgensen et al, 2007), among others.

In the case of rice, because the amount of nitrogen at heading stage is mainly affected by the amount of nitrogen fertiliser at the topdressing and the uptake nitrogen from the soil (Miyama, 1988), it is necessary to investigate vegetation growth at panicle initiation and heading stages as well as grain yield and quality (Umeda et al, 1999). The nitrogen content of rice at panicle initiation stage is an important consideration in deciding the amount of nitrogen fertiliser at topdressing (Ryu et al, 2004). However, research work has focused on special reflectance or vegetation indices using narrow bands of hyperspectral reflectance using ground-based remote sensing. Tian et al, (2005) evaluated the use of R_{680} (the reflectance at 680 nm) and the ratio of R_{810} and R_{680} , the water content, nitrogen rate, and leaf net photosynthetic rate. To assess the nitrogen status in a rice canopy, Lee et al, (2008) suggested the use of R_{735} as a simple spectral index. Wang et al, (2008) used the reflectance of a spectra-radiometer over the range 350-2500 nm to identify the biophysical parameters of rice. Within-field statistical variations in crop growth and nitrogen status were established via partial least squares regression (PLSR) by Nguyen et al, (2006).

Ryu et al, (2005) also evaluated a nitrogen content prediction model using airborne hyperspectral remote sensing at panicle initiation stage via multiple linear regression. The nitrogen content prediction models at panicle initiation and heading stages were also estimated via PLSR by Ryu et al, (2007a). The quality of grain, such as the protein and amylose content at the time of harvesting, was predicted via PLSR using the reflectance at heading stage (Ryu et al, 2007b). However, these models were estimated using data only for the year in which the studies were carried out. Thus, these could not be considered general-purpose models.

To confirm the usefulness and the possibility of practical airborne hyperspectral remote sensing it is necessary to establish a general-purpose nitrogen content prediction model for rice at panicle initiation stage.

In this study, airborne hyperspectral remote sensing was applied to the analysis of the nitrogen content of rice at panicle initiation stage using three years of data. The objectives of this research were: (1) to investigate the vegetation growth of rice at panicle initiation stage, considering weather conditions and the uptake nitrogen from the soil; (2) to establish the nitrogen content prediction model using any combination of the three-year data via multiple linear and PLSR; and (3) to compare such models and establish the general-purpose nitrogen content prediction model at panicle initiation stage.

2. Materials and Methods

2.1. Experimental field and management

A field of the Takatsuki Experimental Farm of Kyoto University Graduate School of Agriculture in Takatsuki City, Osaka Prefecture, Japan was used for the experiments (135°38'E, 34°51'N, 10 m above sea level). The soil in the field is classified as *Typical Fluvaquent* or *Gray Lowland* soil, and the ground was almost flat (maximum gradient 0.07%) (Yanai et al, 2000). The test crop was *Oryza sativa* L., cv. *HINO-HIKARI*, which is a late maturing cultivar in this region. Over three years, pre-germinated seeds were sown on seeding trays to produce uniform seedlings. Thirty to thirty-three day-old seedlings were transplanted into the field with space intervals of 0.30 m × 0.33 m (Inamura et al, 2003).

To provide a range of plants with varying nitrogen contents, the 50 m × 100 m experimental field was divided into 80 plots (5 m × 10 m). There were no plots at both sides of the field (5 m × 50 m) where the fertiliser applicator turned. The amounts of nitrogen and potassium fertiliser (17% N and 17% K₂O, NK fertiliser) applied in the basal dressing were in the range 0 - 65 kg ha⁻¹. In 2003, the half of plots (40) was uniformly applied with 30 kg ha⁻¹ NK fertiliser as a basal dressing. In order to determine the amount of uptake nitrogen from the soil, fertilizer was not applied to several plots; designated as zero plots (Cassman et al, 1997).

In the case of phosphorus (17.5% P₂O₅), 30 kg ha⁻¹ of the fertiliser was also applied uniformly to each plot as a basal dressing. The weather conditions (temperature, accumulated radiation, and rainfall) were also measured using a weather station that was set up at the Takatsuki Experimental Farm.

2.2. Hyperspectral remote sensing

Images at each panicle initiation stage were taken with the airborne hyperspectral sensor (AISA Eagle, Spectral Image Ltd., Oulu, Finland) (Ye et al, 2007) on 1st August 2003, 2nd August 2004, and 5th August 2005. There were 68 bands from 400 to 1000 nm, and the half-bandwidth was changed from 8.15 to 9.19 nm depending on the wavelength. The images were taken at an altitude of approximately 1400 m. The ground spatial resolution at that time was 2.25 m². The reflectance panels were set on the sides of the field to compensate for the differences in the amount of sunlight for each measurement.

The images were analysed using the Environment for Visualizing Images software (ENVI version 4.2, ITT Visual Information Solutions, Boulder, USA). The reflectance of each plot was calculated using not the total area of each plot (around 24 pixels with 54 m²), but the inner part (8 pixels with 18 m²) in order to exclude the reflectance of boundary for the targeted plot and surrounding plots. Rice reflectance was calculated by dividing the reflectance of each plot by the reflectance of the reflectance panels.

2.3. Measurement of vegetation data

Each year, after the images were obtained, the vegetation data at the panicle initiation stage, consisting of the dry mass, the nitrogen concentration, and the nitrogen content of the leaves and stems, were quantified. Vegetation data was sampled at 27 plots with two zero plots in 2003, forty plots with five zero plots in 2004 and forty plots with eight zero plots in 2005. Because the half of plots (40 plots) was uniformly applied, the number of samples in 2003 was reduced.

At each plot, there were two sub-sampling points, each of which consisted of four stocks of rice. After the airborne hyperspectral images were obtained (Pasco Co., Ltd., Tokyo, Japan), the rice plants, without their roots, were picked at each sub-sampling point.

The plants were well-mixed and partitioned into parts of one-quarter and three-quarters. The one-quarter sample was separated into leaves and stems to calculate their ratio. Both sub-samples were dried in a circulation oven drier at 80°C for more than 48 h and then weighed. The dry mass of the leaves or stems for the three-quarters sub-sample was calculated using the ratio of leaves or stems from the one-quarter sub-sample. The leaves and stems of one-quarter samples were finely grounded using a pulverising mill. The nitrogen concentration of each part was measured three times using gas chromatography NC-900 (Sumica Bunseki Center Co., Ltd., Tokyo, Japan). The nitrogen content of the sub-sampling points was calculated as product of the dry mass per unit area and the mean of the nitrogen concentration. Finally, the mean of the two sub-samples was calculated as the nitrogen content of each plot.

2.4. Statistical analysis

Descriptive statistics were calculated over the three years of vegetation growth and reflectance data. The results of Duncan's multiple comparisons for vegetation growth were analysed with SPSS version 13 (SPSS Inc., Chicago, USA). PLSR analysis for reflectance was also performed using the full-cross validation method with Unscrambler version 9.6 (CAMO software AS, Oslo, Norway), which is an efficient way of utilising the samples and improving their relevance. The quality of the model was quantified with the coefficient of determination (R^2) and root mean squared error (RMSE) of calibration or RMSE of prediction (Roger et al, 2003), as shown in Eq. (1) and Eq. (2). The bias in RMSEP was considered a systematic error component, as shown in Eq. (3).

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{I_c} (\hat{y}_i - y)^2}{I_c - 1}} \quad (1)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{I_p} (\hat{y}_i - y_i - bias)^2}{I_p - 1}} \quad (2)$$

$$bias = \frac{\sum_{i=1}^{I_p} (\hat{y}_i - y_i)}{I_p} \quad (3)$$

where \hat{y}_i and y_i are the predicted and measured value of the I th sample; I_c is the number of samples in the calibration set; I_p is the number of samples in the prediction set.

3. Results and Discussion

3.1. Field management

Table 1 shows the crop management and weather conditions for the experimental field during the three years of the study. Crop management was similar over the three years, except for the date of the application of the basal dressing in 2005 and the amount of the nitrogen fertiliser used. However, as expected, weather conditions such as the accumulated temperature and daily radiation changed every year. The accumulated temperature between the transplanting and sampling periods in 2003 was three-quarters of that in the other years. The accumulated daily radiation in 2003 was almost half of that in 2004. Thus, weather conditions may have affected vegetation growth in each year.

Table 1 - Crop management and weather conditions at the experimental field from 2003 to 2005.

Management and weather condition	2003	2004	2005
Date of basal dressing	7 June	4 June	27 May
Date of transplanting	12 June	14 June	15 June
Date of airborne hyperspectral remote sensing	1 August	3 August	5 August
Date of rice plant sampling	4 August	4 August	8 August
Accumulate temperature ^{*1} [°C]	491	645	652
Accumulate daily radiation [MJ/m ²]	397	792	535
Accumulate vegetation day ^{*2}	53 days	51 days	54 days

(*¹: sum of average temperature minus 15°C from transplanting to sampling, *²: day from transplanting to sampling)

3.2. Vegetation data

The vegetation data at panicle initiation stage were quantified after the hyperspectral airborne images were taken during the three years. Table 2 shows the descriptive statistics for the vegetation data of rice at panicle initiation stage.

Table 2 - Descriptive statistics for vegetation data on rice plants at the panicle initiation stage from 2003 to 2005.

		2003 (n=27)	2004 (n=40)	2005 (n=40)
Dry mass [g/m ²]	Leaf	141±18.2*	226±35.0	272±25.9
	Stem	203±30.6	364±39.5	309±33.1
	Total	343±47.6	589±72.5	581±54.5
Nitrogen concentration [%]	Leaf	2.18±0.13	2.06±0.14	2.30±0.18
	Stem	0.76±0.06	0.74±0.07	0.77±0.11
Nitrogen contents [g/m ²]	Leaf	3.07±0.44	4.69±1.04	6.26±0.87
	Stem	1.53±0.24	2.70±0.46	2.36±0.34
	Total	4.60±0.66	7.39±1.47	8.63±1.13

(* : mean ± standard deviation)

Dry mass and nitrogen content in 2003 were different from those in 2004 and 2005. However, nitrogen concentration was similar. The dry mass in 2003 was 52-66% of that in the other years. The nitrogen content in 2003 compared was less than 65% of that in previous years. Nevertheless, the average amount of nitrogen fertiliser applied at each sampling place in 2003 was more than 1.5 times that in the other years, and the average vegetation growth in 2003 was less than 66% of that in the other years. In general, the amount of nitrogen accumulated by plants increases depending on the temperature and radiation intensity (Ta and Ohira, 1982). This depends on the amount of applied nitrogen fertiliser and the uptake nitrogen from the soil. In 2003, solar radiation in Kyoto City, which is near to the experimental field, was also extremely low compared to that in normal years (Katsura et al, 2008).

From the same experimental field, soil chemical properties contributed significantly to the plant yield and can explain as much as 65% of the spatially structured or non-random variations of the yield (Yanai et al, 2001). Therefore, it is necessary to measure the nitrogen content at the plots without nitrogen fertiliser to compare it with differences in the amount of nitrogen obtained from the soil. Table 3 shows the descriptive statistics for the vegetation data at panicle initiation stage without nitrogen fertiliser in the basal dressing for three years. There were significant differences in the dry mass and nitrogen content between 2003 and the other years with 5% of significant level. The results show a similar tendency in Table 2, especially with respect to the dry mass of the leaves, the total nitrogen concentration in the stems, the nitrogen content of the leaves, and the total nitrogen content.

Table 3 - Descriptive statistics for vegetation data on rice at the panicle initiation stage without nitrogen fertiliser in the basal dressing from 2003 to 2005.

		2003 (n=2)	2004 (n=5)	2005 (n=8)
Dry mass [g/m ²]	Leaf	111 ± 2.19 ^a	195 ± 15.9 ^b	256 ± 27.0 ^c
	Stem	164 ± 6.51 ^a	324 ± 14.5 ^b	293 ± 49.3 ^b
	Total	274 ± 4.32 ^a	519 ± 29.4 ^b	548 ± 75.7 ^b
Nitrogen concentration [%]	Leaf	2.22 ± 0.11 ^a	1.94 ± 0.12 ^b	2.33 ± 0.17 ^a
	Stem	0.76 ± 0.04 ^a	0.70 ± 0.06 ^a	0.82 ± 0.11 ^a
Nitrogen contents [g/m ²]	Leaf	2.47 ± 0.17 ^a	3.79 ± 0.30 ^b	5.94 ± 0.57 ^c
	Stem	1.24 ± 0.02 ^a	2.27 ± 0.12 ^b	2.38 ± 0.41 ^b
	Total	3.72 ± 0.19 ^a	6.06 ± 0.35 ^b	8.32 ± 0.96 ^c

(*: mean ± standard deviation, ^{a, b, c}: same letter are not significantly difference at the 5% level as small letter based on Duncan's multiple range test)

The vegetation growth at each plot without nitrogen fertiliser depended only on the amount of uptake nitrogen from the soil. This means that the uptake nitrogen might have differed depending on the weather conditions in each year. The amount of mineralised nitrogen, which is the available nitrogen from the organic soil nitrogen, was linearly related to the temperature (Angus et al, 1994) and the moisture content before irrigation (Hasegawa and Horie, 1994). The difference in the vegetation growth in 2003 may have been affected by the amount of nitrogen available from the soil, which was influenced by the weather conditions, including temperature and daily radiation (Horie et al, 1997). Therefore, there is the possibility that the coefficient of variation (CV) for the amount of uptake nitrogen from soil in 2003 might be lower than those in 2004 and 2005. Although the values of CV for dry mass of leaves, total dry mass and nitrogen content of stems were less than 2% in 2003, these values exceeded 5% in 2004 and 2005. This means that the uptake nitrogen from the soil might have partially reduced the spatial variation in the vegetation growth when the weather conditions were not favourable for growth.

3.3. Analysis of the hyperspectral reflectance

Fig. 1 shows the hyperspectral reflectance of each plot and the rice reflectance at panicle initiation stage for three years. Rice reflectance in 2003 differed from that in the other years, although it is difficult to understand the reason for this difference.

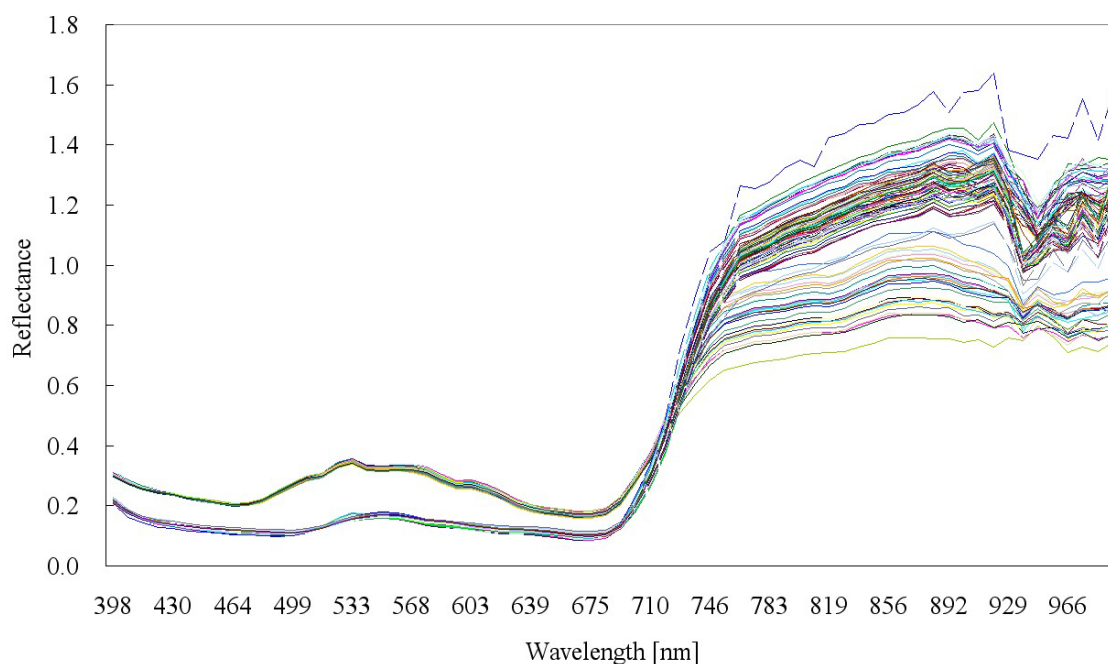


Fig. 1 - Hyperspectral reflectance of the rice plants at the panicle initiation stage for three years.

The dry mass, which is generally correlated to the vegetation coverage ratio (VCR) and the leaf area index (LAI) (Dobermann and Pampolino, 1995), differed from that in the other years as shown in Tables 2 and 3. The maximum of dry mass in 2003 was almost the same as or less than the minimum of vegetation data at the plots without nitrogen fertiliser in 2004 and in 2005. Because of these differences, the rice reflectance in 2003 might have included not only the rice plants but also the reflectance of other parts, such as irrigated water and soil background.

When the soil background affects the reflectance of the canopy, it biases results towards the increasing reflectance within visible regions and decreasing reflectance within the NIR regions (Kimura et al, 2004). Because the underlying water and soil probably contributed most to the measured canopy reflectance spectrum, the values of the reflectance in the visible region were only slightly lower than those in the NIR region (Chang et al, 2005). Because low temperature and radiation deficiency were the causes of the decreasing VCR, and the increasing reflectance of the background, the reflectance values for 2003 may have increased in the visible region and decreased in the NIR region.

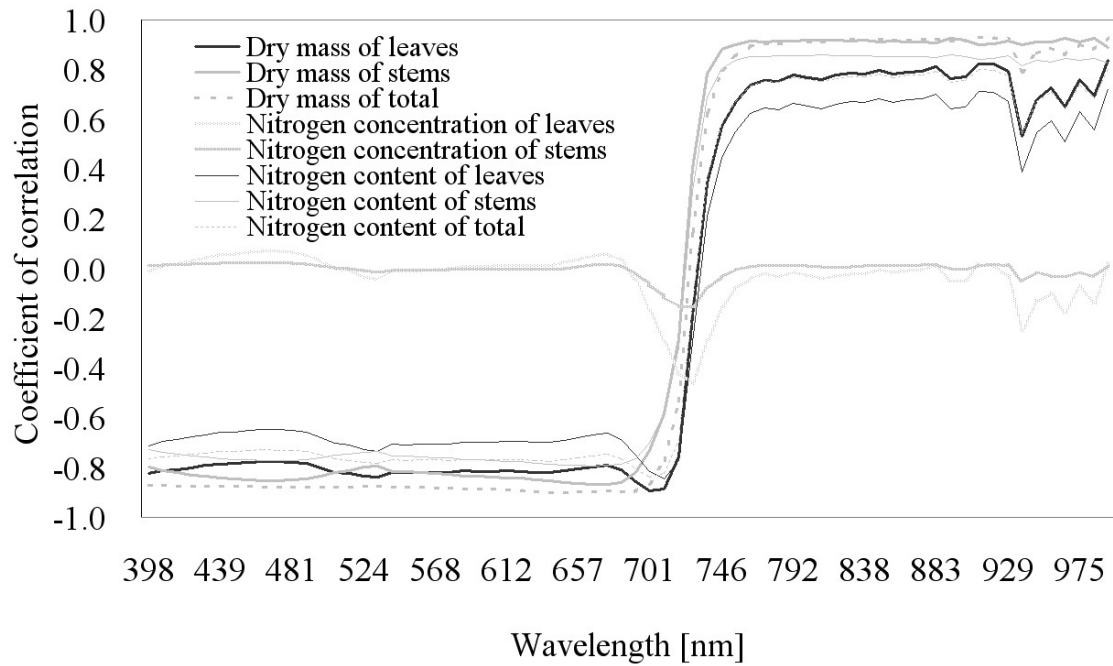


Fig. 2 - Patterns in coefficient of correlation between hyperspectral reflectance and vegetation growth using three years of data.

Fig. 2 shows the patterns in coefficient of correlation between hyperspectral reflectance and vegetation growth using three years of data. The patterns of dry masses and nitrogen content showed significant correlation at all the wavelength with a shift from negative to positive at red-edge (around 692 - 750 nm). The patterns, however, did not show any significant correlation except at the red-edge. The results indicate that the nitrogen content at panicle initiation stage is mainly influenced by dry mass than nitrogen concentration. Moreover, only at the red-edge do all the factors vary, as shown. Therefore, the red-edge could be an important region to predict the vegetation growth at panicle initiation stage.

3.4. Partial least square regression (PLSR)

3.4.1. Modelling

Table 4 shows the results of the PLSR analysis with the full-cross validation method using any combination of three-year data. The nitrogen content of rice at panicle initiation stage could be validated with full-cross validation of PLSR analysis.

The correlation coefficients of validation for each year model were 0.758, 0.847 and 0.443, respectively 2003, 2004 and 2005. The RMSEP of the 2005 model was more than 1 g m⁻². When the models were estimated using any combination of two-year data (2003-2004, 2003-2005 and 2004-2005), the coefficient of correlation was more than 0.792 and RMSEP was less than 0.876 g m⁻². Although the accuracy of RMSEP and SEP partially decreased, the coefficients of correlation for the full-cross validation increased to more than 0.792. The 2003-2004-2005 model was established with five latent variables (PCs), with r = 0.938, and RMSEC = 0.682, for 107 samples. When the accuracy was predicted via full-cross validation, the model was estimated by r = 0.919 and RMSEP = 0.771 g m⁻².

Table 4 - Results of the PLSR analysis with full-cross validation using any combination of the three-year data.

	2003 (n=27)		2004 (n=40)		2005 (n=40)		2003&2004 (n=67)		2003&2005 (n=67)		2004&2005 (n=80)		Total (n=107)	
	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.	Cal.	Val.
PCs ^{*1}	1		3		1		4		1		4		5	
Slope	0.627	0.602	0.868	0.795	0.257	0.225	0.931	0.895	0.867	0.863	0.711	0.656	0.879	0.856
Offset	1.715	1.826	0.978	1.547	6.406	6.687	0.434	0.675	0.933	0.956	2.316	2.747	0.865	1.031
R	0.792	0.758	0.932	0.847	0.507	0.443	0.965	0.934	0.931	0.927	0.843	0.792	0.938	0.919
RMSE ^{*2}	0.397	0.424	0.528	0.786	0.959	1.001	0.479	0.652	0.801	0.824	0.771	0.876	0.682	0.771
SE ^{*3}	0.405	0.433	0.535	0.793	0.972	1.013	0.438	0.657	0.807	0.830	0.775	0.882	0.685	0.774
Bias	-4.4e ⁻⁸	-0.003	6.0e ⁻⁸	0.035	-1.2e ⁻⁸	1.6e ⁻⁵	3.2e ⁻⁷	0.020	8.5e ⁻⁸	-0.001	3.3e ⁻⁷	-0.006	9.5e ⁻⁷	-4.0e ⁻⁴

(*¹: number of latent variable, *²: root mean square error, *³: standard of error)

Figs. 3 and 4 show two variation patterns of regression coefficients depending on the variable (wavelength) for the PLSR models using any combination of three-year data. There are two different variation groups of regression coefficients depending on the variables, yet the patterns of groups are similar. The regression coefficients of pattern 1 (Fig. 3), with any combination of the data in 2003 and 2005, were between -0.274 and 0.593. The models in the group of pattern 1 also have only one latent variable (PCs) as shown in the Table 4.

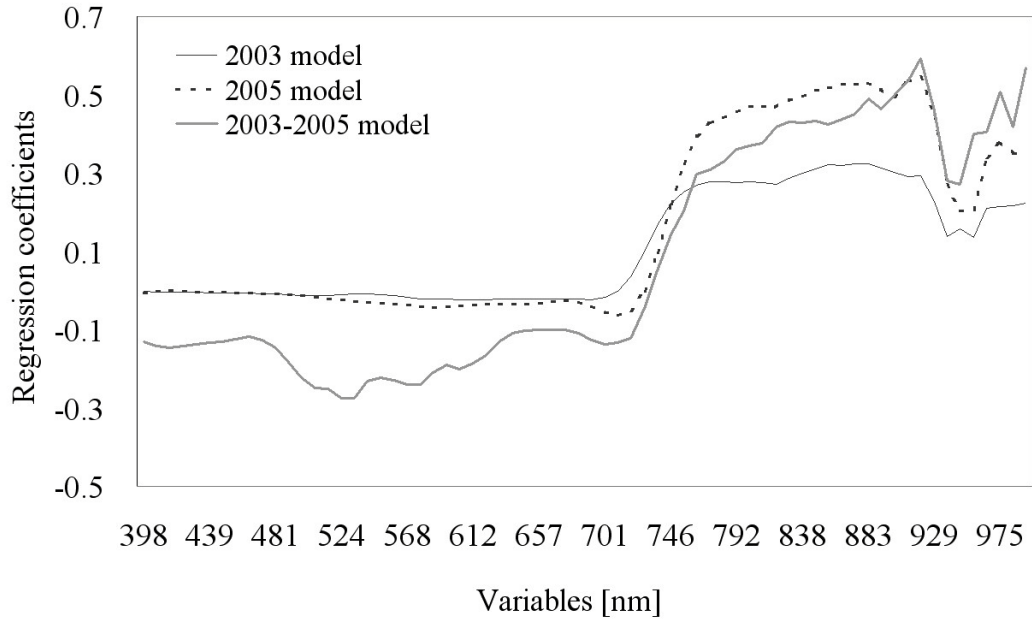


Fig. 3 – Pattern 1 variation of regression coefficients for PLSR models.

The regression coefficients of pattern 2 (Fig. 4), which were included in the 2004 model, the 2003-2004 model, and the 2003-2004-2005 model, existed more widely between -15.42 and 6.367, comparing simultaneously with Fig 3 and Fig. 4.

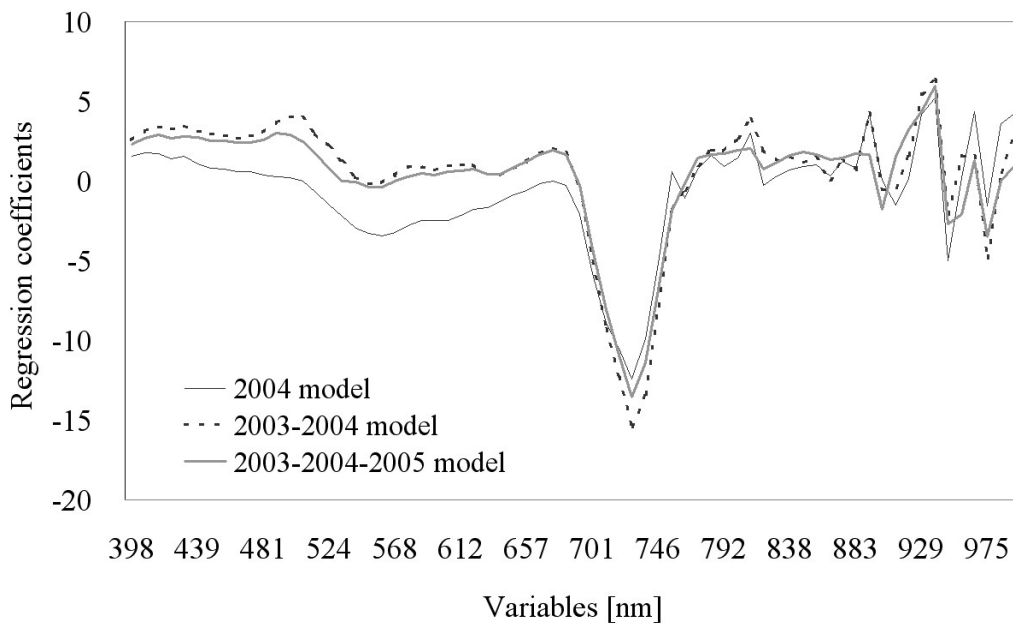


Fig. 4 – Pattern 2 variation of regression coefficients for PLSR models.

Moreover, the specific regions of pattern 1 were mainly in the NIR regions and partially in the visible regions only for the 2003-2005 model. The specific regions of pattern 2 were mainly in the red-edge regions and partially in the NIR regions. Especially in the red-edge region where the rice reflectance increased, all the regression coefficients changed from positive to negative values between R_{692} and R_{746} . It was also confirmed that depending on vegetation growth the variation of reflectance occurred at the red edge, as shown in Fig. 2. Several studies have been conducted using bands in the red-edge region, such as on RI_{1dB} (R_{735} and R_{720}) for the growth cycle of wheat (Gupta et al, 2003), and $mND705$ (R_{750} , R_{705} , and R_{445}) for the development of spectral indices to predict leaf pigment content, which is relatively insensitive to species and leaf structure variations. These results could therefore be applied to larger-scale remote-sensing studies (Sims and Gamon, 2002). Although it is difficult to identify the specific bands or the regions of reflectance, the red-edge is an important region in the estimation of the nitrogen content of rice at panicle initiation stage. Moreover, the 2004 data might play an important role in establishing the nitrogen content prediction model because the regression coefficients of the model are similar to those of the 2003-2004 model and the 2003-2004-2005 model.

3.4.2. Comparison of models

When the 2003-2004 model, which is established using the nitrogen contents in 2003 and 2004, was used to predict the 2005 data, the prediction error increased by 1.050 g m^{-2} . In the case of the 2003-2005 model, it also increased by 2.378 g m^{-2} . The prediction error of the 2004-2005 model to predict the data in 2003, was 5.061 g m^{-2} . Although the 2003-2004-2005 model may be used in predictions, the 2003-2004, the 2003-2005, and the 2003-2004-2005 models might more suitable to estimate the nitrogen content of rice at panicle initiation stage because the 2004-2005 model had low prediction accuracy. It is necessary to compare the regression coefficients of the selected models in order to choose a general-purpose prediction model.

Fig. 5 shows the predicted nitrogen contents of the data in 2005 using the 2003-2004 model and the 2003-2004-2005 model via the PLSR analysis.

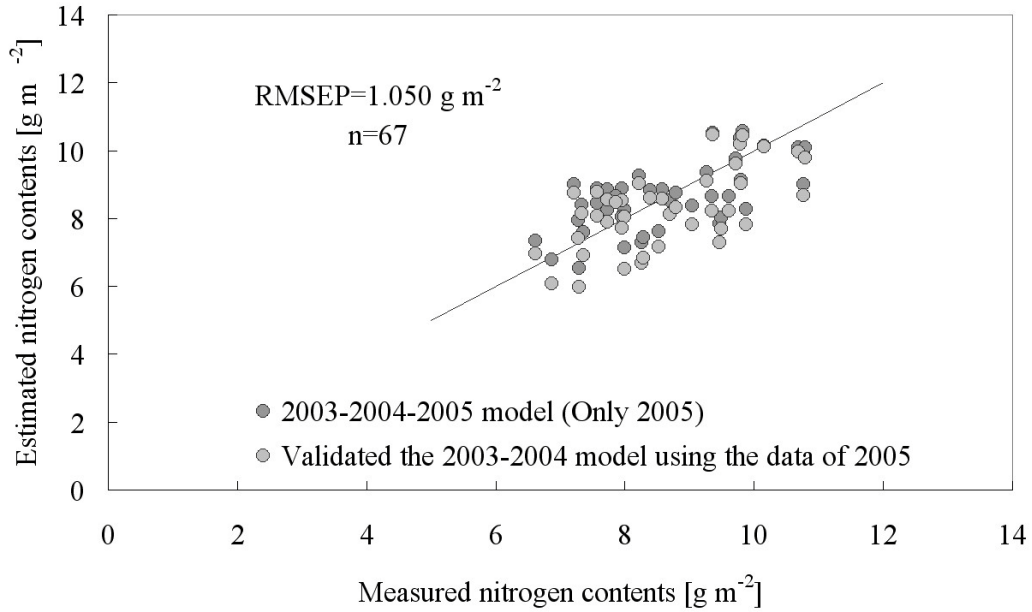


Fig. 5 - Predicted nitrogen contents of the data in 2005 using the 2003-2004 model and the 2003-2004-2005 model from PLSR analysis.

When the 2003-2004 model was used to predict the nitrogen contents in 2005, nitrogen content was underestimated compared with that predicted by the 2003-2004-2005 model or measured.

Fig. 6 shows the predicted nitrogen contents of the data in 2004 using the 2003-2005 model and the 2003-2004-2005 model via the PLSR analysis. In the case of the 2003-2005 model, nitrogen content in 2004 was always overestimated compared with that predicted by the 2003-2004-2005 model or measured. Because the accuracy of the predicted model increased to 2.378 g m⁻², the 2004 data did not fit the 2003-2005 model. The lower accuracy might be due to the greater sensitivity of the data in 2004, not in the NIR region but in the red-edge region, as shown in Fig. 4. As result, the 2003-2004 model and the 2003-2004-2005 model may be more suitable as general-purpose prediction models, because the 2003-2005 model is sensitive to the data obtained in 2004 and has only one PCs.

Moreover, the pattern of regression coefficients was different from those of 2003-2004 model and 2003-2004-2005 model.

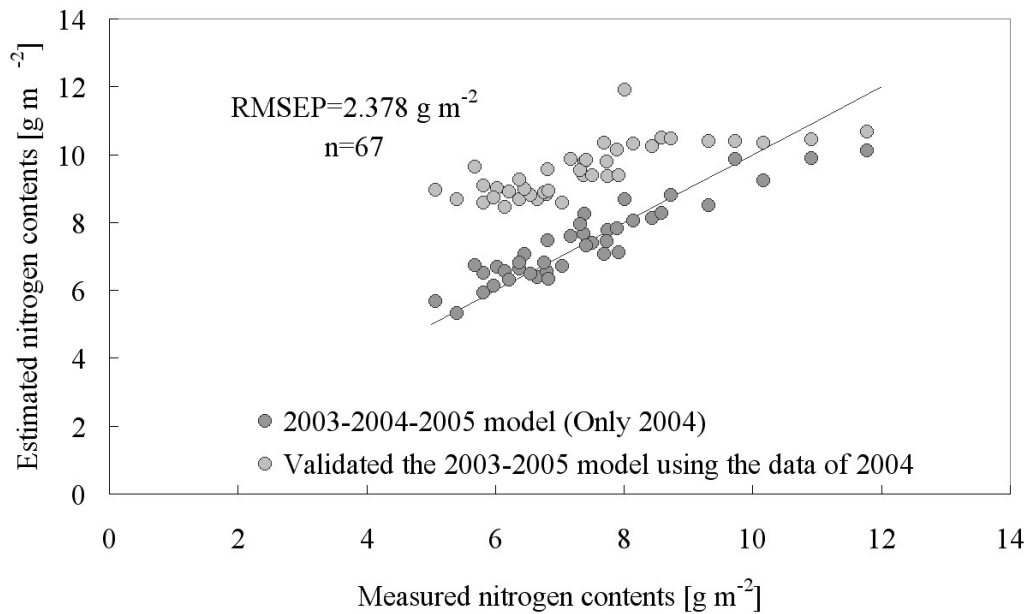


Fig. 6 - Predicted nitrogen contents of the data in 2004 using the 2003-2005 model and the 2003-2004-2005 model from PLSR analysis.

3.4.3. General-purpose prediction model

The similarities and the differences between the 2003-2004 model and the 2003-2004-2005 model were compared using the latent variables (PCs) in order to develop a robust general-purpose model. The four latent variables of the 2003-2004 model explained sequentially 98%, 2%, 0%, and 0% of the X variable (the reflectance of hyperspectral band), and are not shown in this paper. In the case of Y variable, the nitrogen contents were explained as 75%, 4%, 11%, and 4% by four latent variables. The five latent variables of the 2003-2004-2005 model explained 95%, 3%, 2%, 0%, and 0% of X variables, and 59%, 16%, 5%, 7%, and 0% of Y variable. This means that the latent variables PC5 of the 2003-2004-2005 model did not affect the variables and the prediction of the nitrogen content of rice. When the 2003-2004-2005 model was composed of four PCs, its accuracy slightly changed but by less than 0.03 g m⁻² with respect to RMSE and standard of error (SE), which are not shown in this

report. It is necessary to compare the X loading of each PC in the 2003-2004 model with that in the 2003-2004-2005 model.

Fig. 7 shows the similarities of X loading weights between the 2003-2004 model and the 2003-2004-2005 model depending on the variables. There were similar patterns of X loadings for latent variables between the 2003-2004 model and the 2003-2004-2005, as shown in Fig. 6: (1) PC1 of this model, (2) PC2 of the 2003-2004 and PC3 of 2003-2004-2005 model, and (3) PC3 of the 2003-2004 model and PC4 of the 2003-2004-2005 model. This means that the only different PCs were PC4 (4%) of the 2003-2004 model and PC2 (16%) of the 2003-2004-2005 model, because PC5 had 0% variable and nitrogen content, as shown in Fig. 8. These similarities and differences might have affected the regression coefficients and the accuracy of the prediction model. It is difficult to decide the general-purpose prediction model between the 2003-2004 model and the 2003-2004-2005 model, because the 2003-2004-2005 model must be checked for accuracy using other data. However, the 2003-2004-2005 model might be more suitable for the general-purpose model than the 2003-2004 model, because it is possible to consider and validate using three years data.

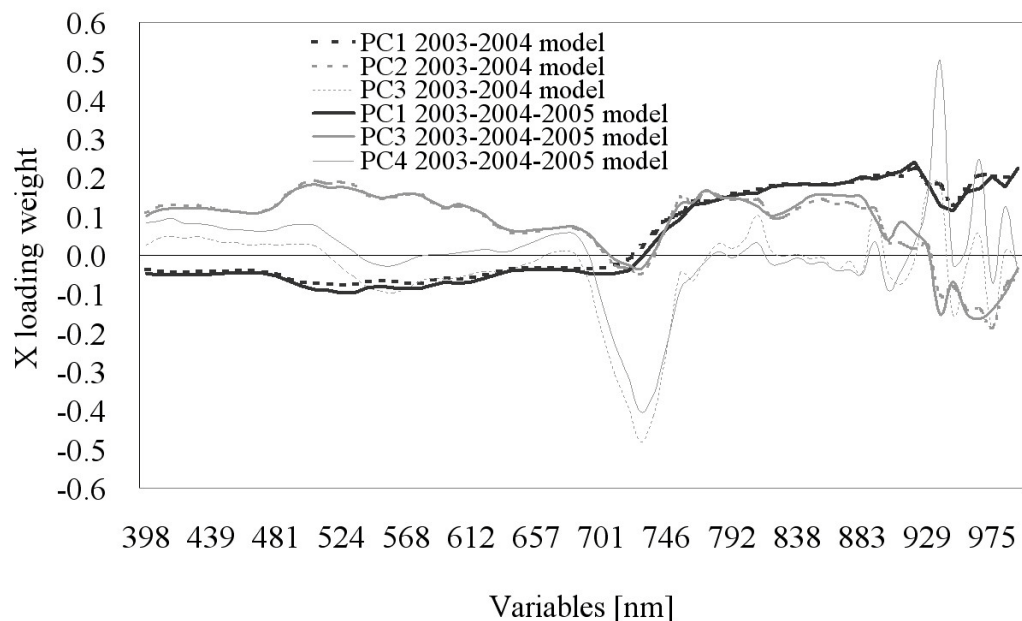


Fig. 7 - Similarities of X loading weights between the 2003-2004 model and the 2003-2004-2005 model depending on the variables.

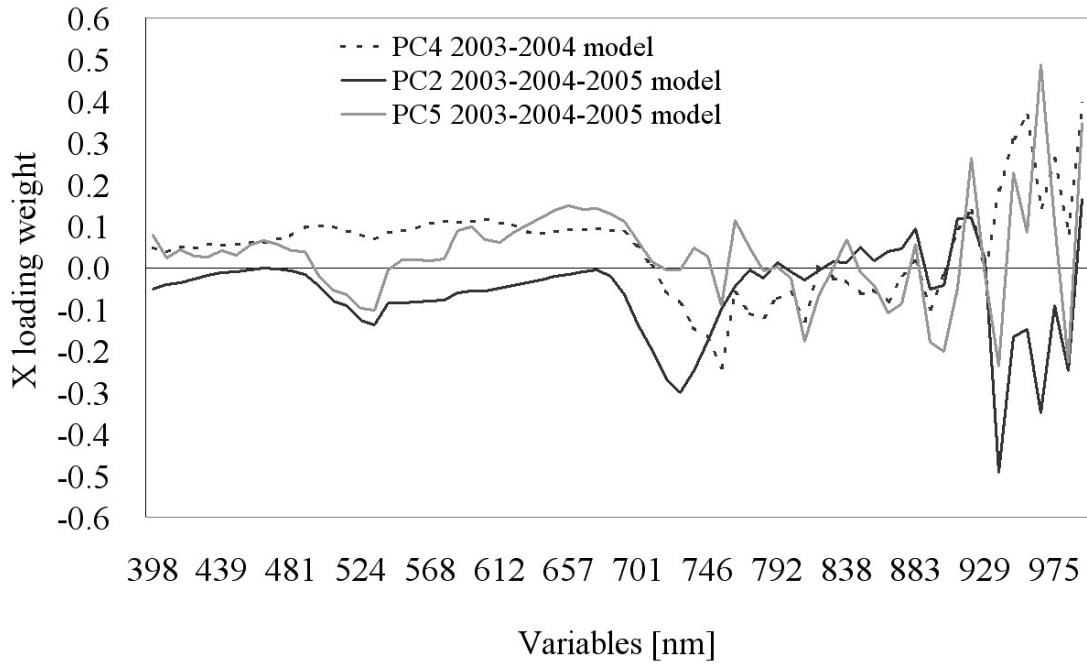


Fig. 8 - Difference of X loading weights between the 2003-2004 model and the 2003-2004-2005 model depending on the variables.

Fig. 9 shows the measured and estimated nitrogen contents for the 2003-2004-2005 model from the PLSR analysis. Although vegetation data showed that the nitrogen content in 2003 was less than in the other years, few values of the nitrogen content in 2003 exceeded those in 2004, as shown at Fig. 9. The nitrogen concentration of the leaves might have caused this since the maximum dry mass in 2003 was less than the minimum in 2004, as shown in Tables 2 and 3. Therefore, some of the nitrogen content values in 2003 might have been greater than those in 2004. Furthermore, it can be seen that the 2004 data plays an important role; had the 2004 data not existed, the 2003 and 2004 data would have been completely separated into two groups.

Nevertheless, because there were differences between several environmental variables that affected the vegetation growth of the rice, it is possible to establish a general-purpose nitrogen prediction model using airborne hyperspectral remote sensing. This means that the general-purpose nitrogen prediction model might be able to estimate the nitrogen content of rice at panicle initiation stage regardless of changes in environmental variables.

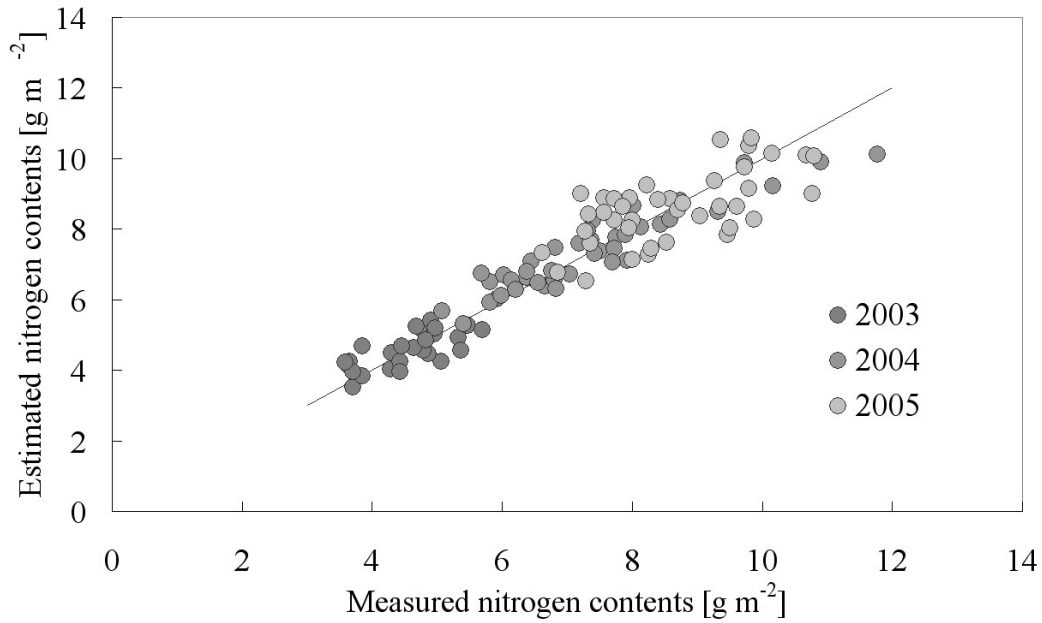


Fig. 9 - Measured and estimated nitrogen contents for the 2003-2004-2005 model from PLSR analysis.

Therefore, it might be possible to provide the accurate and reliable information in order to control vegetation growth using variable rate management for paddy field. However, there are still several problems with the precision of the model. The cause of the hyperspectral reflectance differences between 2003 and the other years need to be analysed. It might be necessary to apply a ground-based hyperspectral image to clearly analyse the absolute reflectance of rice and compare it with a mixed reflectance, such as of the rice and the background (Okamoto et al, 2007). It is also necessary to confirm the extensiveness of the general-purpose nitrogen content prediction model at panicle initiation stage using the data on the different locations and growth conditions, such as on the early- and medium-maturing cultivars. Furthermore, it is necessary to establish the general-purpose nitrogen content estimation model for other growth stages of rice in order to estimate vegetation growth and predict grain yield and the grain quality, such as the amylose and protein contents.

4. Conclusions

Airborne hyperspectral remote sensing was applied to the estimation and prediction of the nitrogen content of rice at panicle initiation stage using three years of data. There were significant differences in dry mass depending on the weather conditions each year and this, affected by the amount of uptake nitrogen from the soil. There was a similar tendency with respect to the vegetation data at without nitrogen fertiliser application in the basal dressing. The VCR in 2003 might have decreased, and the reflectance of the plots might have had different tendencies than in the other years.

In the PLSR analysis, the coefficient of correlation was more than 0.758 and RMSEP was less than 0.876 g m^{-2} among the models. When each model was predicted with the remaining data, RMSEP was 1.050 g m^{-2} for the 2003-2004 model, 2.378 g m^{-2} for the 2003-2005 model, and 5.061 g m^{-2} for the 2004-2005 model. The 2003-2004-2005 model, using the three-year data, was also estimated with five latent variables, with $r = 0.938$ and $\text{RMSEC} = 0.682$ for 107 samples. Although there were two different groups of regression coefficients according to the variables, the patterns of the groups were similar. The specific regions of any combination of models using the 2003 and 2005 data were mainly in the NIR regions with positive value; and in the other models, such as the 2004 model, the 2003-2004 model, and the 2003-2004-2005 model, mainly in the red-edge regions (from R_{692} to R_{746}) with a negative value. The 2003-2004 model and the 2003-2004-2005 model were more suitable than the other models. Several X loadings of latent variables between the 2003-2004 model and the 2003-2004-2005 model were similar to each other. The only different latent variables were PC4 of the 2003-2004 model and PC2 of the 2003-2004-2005 model. These similarities and differences might have affected the regression coefficients and the accuracy of the prediction model. From the regression coefficients, and the X loadings of those models, it is possible to identify the importance of the red-edge and to understand the tendency of the models to establish the general-purpose prediction model. It is difficult to directly compare the accuracy of the models, because the 2003-2004-2005 model does not have the data set required for validation. It is possible, however, to improve the accuracy of the models by accumulating data over several years.

Therefore, data over several years must be collected to increase the accuracy of the general-purpose model or to establish an improved general-purpose model.

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science KAKENHI (A: 13306019). The author wishes to thank Professor Tatsuya Inamura of the Kyoto University of Agriculture Sciences for assistance of rice management at Takatsuki Experimental farm of the Graduate School of Agriculture, Kyoto University.

References

- Angus J F; Ohnishi M; Horie T; Williams R L (1994). A preliminary-study to predict net nitrogen mineralization in a flooded rice soil using anaerobic incubation. *Australian Journal of Experimental agriculture*, 34(7), 995-999
- Blackburn G A (1998). Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches. *Remote Sensing of Environment*, 66(3), 273-285
- Blackmore S (1994). Precision farming: An introduction. *Outlook on Agriculture*, 23(4), 275-280
- Cassman K G; Peng S; Olk D C; Ladha J K; Reichardt W; Dobermann A; Singh U (1998). Opportunities for increased nitrogen-use efficiency from improved resource management in irrigated rice systems. *Field Crops Research*, 56, 7-39
- Chang K W; Shen Y; Lo J C (2005). Predicting rice yield using canopy reflectance measured at booting stage. *Agronomy Journal*, 97, 872-878
- Detar W R; Penner J V; Fuck H A (2006). Airborne remote sensing to detect plant water stress in full canopy cotton. *Transactions of the ASABE*, 49(3), 655-665
- Dobermann A; Pampolino M F (1995). Indirect leaf area index measurement as a tool for characterizing rice growth at the field scale. *Communications on Soil Science and Plant Analysis*, 26, 1507-1523

- Gupta R K; Vijayan D; Prasad T S (2003). Comparative analysis of red edge hyperspectral indices. *Adv. Space Res.*, 32, 2217-2222
- Hansen P M; Schjoerring J K (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86(4), 542-553
- Hasegawa T; Horie T (1994). A simplified model for estimating nitrogen mineralization in paddy soil. *Japanese Journal of Crop Science*, 63(3), 496-501
- Horie T; Ohnishi M; Angus J F; Lewin L G; Tsukaguchi T; Matano T (1997). Physiological characteristics of high-yielding rice inferred from cross-location experiments. *Field Crops Research*, 52, 55-67
- Inamura T; Hamada H; Iida K; Umeda M (2003). Correlation of the amount of nitrogen accumulated in the aboveground biomass at panicle initiation and nitrogen content of soil with the nitrogen uptake by lowland rice during the period from panicle initiation to heading. *Plant Prod. Sci.*, 6(4), 302-308
- Inoue Y (2003). Synergy of remote sensing and modeling for estimating ecophysiological processes in plant production. *Plant Prod. Sci.*, 6(1), 3-16
- Jorgensen R N; cristensen L K; Bro R (2007). Spectral reflectance at sub-leaf scale including the spatial distribution discriminating NPK stress characteristics in barley using multiway partial least squares regression. *International Journal of Remote Sensing*, 28(5), 943-962
- Katsura K; Maeda S; Lubis I; Horie T; Cao W; Shiraiwa T (2008). The high yield of irrigated rice in Yunnan, China: A cross-location analysis. *Field Crops Research*, 107, 1-11
- Kimura R; Okada S; Miura H; Kamichika M (2004). Relationships among the leaf area index, moisture availability, and spectral reflectance in an upland rice field. *Agricultural Water Management*, 69, 83-100
- Lee Y J; Yang C M; Chang K W; Shen Y (2008). A simple spectral index using reflectance of 735nm to assess nitrogen status of rice canopy. *Agronomy Journal*, 100, 205-212
- Lobell D B (2007). The cost of uncertainty for nitrogen fertilizer management: A sensitivity analysis. *Field Crops Research*, 100, 210-217

- Miyama M (1988). A new method of standardization of fertilizer application for rice plant based on the optimum nitrogen contents (in Japanese with English abstract). *Spec. Bull. Chiba Agric. Exp. Stn*, 15, 1-92
- Nguyen H T; Kim H J; Nguyen A T; Nguyen L T; Shin J C; Lee B W (2006). Using canopy reflectance and partial least squares regression to calculate within-field statistical variation in crop growth and nitrogen status of rice. *Precision Agriculture*, 7(4), 249-264
- Okamoto H; Murata T; Kataoka T; Hata S (2007). Plant classification for weed detection using hyperspectral imaging with wavelet analysis. *Weed biology and Management*, 7, 31-37
- Plant R E; Munk D S; Roberts B R; Vargas R L; Rains D W; Travis R L; Hutmacher R B (2000). Relationships between remotely sensed reflectance data and cotton growth and yield. *Transaction of the ASAE*, 43(3), 535-546
- Ray S S; Das G; Singh J P; Panigrahy S (2006). Evaluation of hyperspectral indices for LAI estimation and discrimination of potato crop under different irrigation. *International Journal of Remote Sensing*, 27(23-24), 5373-5387
- Rama Rao N; Garg P K; Ghosh S K (2007). Development of an agricultural crops spectral library and classification of crops at cultivar level using hyperspectral data. *Precision Agriculture*, 8, 173-185
- Roger J M; Chauchard F; Bellon Maurel V (2003). EPO-PLS external parameter orthogonalisation of PLS application to temperature independent measurement of sugar content of intact fruits. *Chemometrics and Intelligent Laboratory System*, 66(2), 191-204
- Ryu C S; Iida M; Suguri M; Umeda M; Inamura T; Inoue H; Shinjo H; Moritsuka N (2004). Effect of variable rate fertilizer application aimed at reducing the spatial variability of grain yield on rice taste (in Japanese with English abstract). *Journal of JSAM*, 66(5), 49-62
- Ryu C S; Suguri M; Nishiike Y; Umeda M (2005). Making nitrogen contents model using hyperspectral remote sensing and estimation nitrogen contents by nitrogen content model (in Japanese with English abstract). *Journal of JSAM*, 67(6), 47-54
- Ryu C S; Suguri M; Iida M; Umeda M (2007). Validation of rice taste elements influenced by amount of nitrogen fertilizer and estimation using remote sensing (in Japanese with English abstract). *Journal of JSAM*, 69(1), 52-58

- Ryu C S; Suguri M; Umeda M (2007). Estimation and comparison of nitrogen contents model of rice plant using remote sensing (in Japanese with English abstract). *Journal of JSAM*, 69(3), 43-50
- Schut A G T; Van Der Heijden G W A M; Hoving I; Stienezen M W J; Van Evert E K; Meuleman J (2006). Imaging spectroscopy for on-farm measurement of grassland yield and quality. *Agronomy Journal*, 98(5), 1318-1325
- Sims D A; Gamon J A (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81(2-3), 337-354
- Steven M C (2004). Correcting the effects of field of view and varying illumination in spectral measurements of crops. *Precision Agriculture*, 5, 55-72
- Strachan I B; pattey E; Boisvert J B (2002). Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. *Remote Sensing of Environment*, 80(2), 213-224
- Ta T; Ohira K (1982). Effects of temperature and light intensity on the uptake and assimilation of ¹⁵N-labeled ammonium and nitrate in Indica and Japonica rice plants. *Soil Sci. Plant Nutr.*, 28(1), 79-90
- Thenkabail P S; Smith R B; De Pauw E (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71(2), 158-182
- Tian Y; Zhu Y; Cao W (2005). Monitoring leaf photosynthesis with canopy spectral reflectance in rice. *Photosynthetica*, 43(4), 481-489
- Umeda M; Inamura T; Kosaki T; Shibata Y; Hara Y (1999). Technology in the 21st century: Precision agriculture (2) (Special edition, in Japanese). *Journal of JSAM*, 61(4), 3-23
- Walker T W; Kingery W L; Street J E; Cox M S; Larry Oldham J; Gerard P D; Han F X (2003). Rice yield and soil chemical properties as affected by precision land leveling in alluvial soils. *Agronomy Journal*, 95, 1483-1488
- Wang F M; Huang J F Wang X Z (2008). Identification of optimal hyperspectral bands for estimation of rice biophysical parameters. *Journal of integrative plant biology*, 50(3), 291-299
- Yanai J; Lee C K; Umeda M; Kosaki T (2000). Spatial variability of soil chemical properties in a paddy field. *Soil Sci. Plant Nutr.*, 46(2), 473-482

- Yanai J; Lee C K; Kaho T; Iida M; Matsui T; Umda M; Kosaki T (2001). Geostatistical analysis of soil chemical properties and rice yield in a paddy field and application to the analysis of yield-determining factors. *Soil Sci. Plant Nutr.*, 47(2), 291-301
- Ye X; Sakai K; Mango M; Asada S; Sasao A (2007). Prediction of citrus yield from airborne hyperspectral imagery. *Precision Agriculture*, 8(3), 111-125
- Zarco-Tejada P J; Berjon A; Lopez-Lozano R; Miller J R; Martin P; Cachorro V; Gonzalez M R; De Frutos A (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99(3), 271-287