Integrate Growing Temperature to Estimate the Nitrogen Content of Rice Plants at the Heading Stage Using Hyperspectral Imagery

Hiroyuki Onoyama, Chanseok Ryu, Masahiko Suguri, and Michihisa Iida

***Abstract*—Ground-based hyperspectral imaging was used for estimating the nitrogen content of rice plants at the heading stage. The images were separated into two parts: 1) the rice plant; and 2) other elements using the equation of "*GreenNDVI–NDVI*." Ref_{RICE} was calculated as the ratio of the reflectance of the rice plant to that of a reference board. Partial least square (PLS) model using reflectance data (R-PLS model) and PLS model using reflectance and temperature data (RT-PLS) was constructed to compare the accuracy between them. RT-PLS model was developed to improve the accuracy of R-PLS model by considering the differences of weather condition among years. When the precision (R^2) and accuracy [root-mean-square error (RMSE) and relative error (RE)] of each R-PLS model were evaluated for each year using twofold cross-validation, R^2 ranged from 0.42 to 0.81, RMSE ranged from 0.81 to $1.13\ \mathrm{gm}^{-2}$, and RE ranged from 10.1% to 11.8%. When R-PLS model of each year was used to predict the other years' data to determine the predictive power, RMSE values were higher (ranging from 1.40 to $5.82\ \mathrm{gm}^{-2}$) than those in each year's validation value due to over- or underestimation. When an R-PLS model based on the data of 2 years was fitted, RMSE ranged from 1.11 to $4.15\ \mathrm{gm}^{-2}$ and RE ranged from 13.7% to 42.8%. By contrast, in RT-PLS models, RMSE and RE fell to less than $1.21\ \mathrm{gm}^{-2}$ and 12.3%, respectively. Thus, a combination of reflectance and temperature data was useful for constructing a model of rice plant at the heading stage.**

***Index Terms*—Ground-based hyperspectral imaging, heading stage, model considered growing temperature, nitrogen content, paddy rice, partial least square (PLS) regression.**

I. Introduction

SIGNIFICANT agricultural production increase needed to satisfy future global food needs. Productivity improvements will be a key factor in reducing global food insecurity [1]. One approach to this problem is large-scale farming using labor-saving technologies. Large-scale farming, though cost-effective, has the disadvantages of increasing agricultural work, while risking poor control of delicate cropping operations such as fertilizer application. Automation, such as the operation of combine harvesters [2] or the guidance systems for them [3], reduces the agricultural workload.

Crop monitoring reduces the risk of poor control, because it is critical for the estimation of crop growing status, quality, and yield. Thus, it is important to study the growth and nitrogen content status at the panicle initiation, heading, and maturity stages [4]. The quality and quantity of rice grains are closely related to the growth and nitrogen content status at the heading stage [5], [6]. Thus, it is important to identify the growth and nitrogen content status of rice plants at the heading stage. Moreover, it has also been confirmed that variable rate fertilizer application is adoptable for controlling the growth and nitrogen content status of rice plants uniformly. However, the accurate quantification of crop growth status by chemical analysis is destructive, laborious, and time-consuming.

Remote sensing has been investigated by many researchers and its ability to predict crop biochemical or physiological properties nondestructively using crop canopy reflectance has been demonstrated [7]–[13]. Several remote sensing applications, such as satellite-based, aerial, and ground-based applications, have proven to be potential sources of reflectance data for the extraction of various spectral features and/or indices of vegetation [14]. Hyperspectral sensor on the Earth Observing-1 was applied to map forest nutrient [13]. Several vegetation indices were proposed for estimating canopy chlorophyll and nitrogen content using hyperspectral remote sensing [12]. In the case of rice plant, $\mathrm{R}_{830}/\mathrm{R}_{550}$ ratio index [ratio of near-infrared (NIR) to green reflectance] is effective for the estimation of leaf nitrogen content [15]. $\mathrm{R}_{810}/\mathrm{R}_{560}$ ratio index is also developed to nondestructively monitor the nitrogen status of rice plants [10]. Rice leaf nitrogen content would be estimated by high-resolution measurements in visible and NIR region [16]. The spatial resolution of ground-based remote sensing is high and prediction capabilities may be improved, although the area of application is smaller than that of satellite-based or aerial methods. In ground-based remote sensing, spectrometers have been widely used, but the data obtained are one-dimensional. Consequently, the data include the reflectance not only of the desired target but also of nontargets [17], such as soil background and irrigation water in the case of rice paddies. A hyperspectral image, in contrast, provides a three-dimensional (3-D) spatial and spectral data cube and can extract the reflectance of only the target.

Recently, hyperspectral imaging systems have been studied. The potential of hyperspectral imaging for the classification and characterization of turkey hams were investigated [18]. It is also

Manuscript received June 04, 2013; revised May 06, 2014; accepted May 26, 2014. Date of publication July 07, 2014; date of current version August 01, 2014. This work was supported in part by Nantan City and the Research Grants for Japan Society for the Promotion of Science Postdoctoral Fellows under Grant 21-09333.

H. Onoyama, M. Suguri, and M. Iida are with the Division of Environmental Science and Technology, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan.

C. Ryu is with the Department of Bioindustrial Machinery Engineering, College of Agriculture and Life Sciences, Institute of Agricultural and Life Sciences, Gyeongsang National University, Jinju 660-701, Korea (e-mail: ryucs@gnu.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2014.2329474

possible to detect various common defects on orange surface [19]. Hyperspectral imaging system was effective for the nondestructive measurement of localized nitrate concentration in a leaf of *Brassica rapa* var. *peruviridis* [20]. In addition to these laboratory-based studies, a system to map the herbage mass of perennial ryegrass using field-based hyperspectral imaging [21]. Hyperspectral sensor system loaded on unmanned autonomous vehicle also developed [22].

Partial least square (PLS) regression modeling has desirable properties that improve the precision of the modeling parameters with an increasing number of relevant variables and observations [23]. PLS regression analysis is one of the most efficient methods for extracting and creating reliable models in a wide range of fields [7] and when applied to hyperspectral reflectance data, it may be useful as an exploratory and predictive tool [24]. From a practical standpoint, a prediction model should have the ability to estimate crop status from new remote-sensing data with a desired precision or accuracy. For the purpose of developing a model considered growing condition, a mutual prediction was calculated by use of a single year's data for the training set and the remaining years' data for the test set [25]. Remote sensing models may be affected by differences in environmental variables, such as weather conditions [25], [26]. Previous studies have identified the important spectral bands that is related to the crop vegetation using PLS regression analysis and stepwise multiple linear regression using hyperspectral remote sensing, and they have revealed that the red edge is one of the critical band for estimation of crop vegetation [7], [24]–[26]. However, the red edge position moved to longer wavelength until close to heading stage and otherwise shifted to shorter wavelength after heading stage [27]. In the case of hyperspectral remote sensing, the effect of the red edge shift might have a significant impact on estimation modeling. Thus, we proposed a construction method of the hyperspectral imaging model that integrate growing temperature data as the phenological timing. Growing degree-day (GDD) is calculated from air temperature measurements and represents a weather condition that is frequently used to describe the timing of biological processes [28]. Thus, we proposed a model based on a combination of both reflectance and GDD. To our knowledge, previous studies have not used hyperspectral imaging with a GDD to develop an estimation model of nitrogen content in rice at the heading initiation stage. Therefore, this study aimed to construct a model that considered GDD over a 3-year period. The specific research objectives were 1) to establish a PLS regression model to determine nitrogen content at the heading stage using ground-based hyperspectral imaging; 2) to suggest an appropriate GDD-integrated model that worked across years; and 3) to compare the accuracy of conventional reflectance models and the suggested GDD-integrated model.

II. Materials and Methods

A. Experimental Fields

The experimental fields were located in Yagi town, Nantan City, Kyoto Prefecture, Japan (135°55′E, 35°09′N; 125 m above sea level). Four fields were used over 3 years (2008–2010). The area of each field was approximately 0.25 ha and the soil type was silty clay. The tested rice species was *Oryza sativa L.* and the variety was *Kinu-Hikari*. There were 8–10 sampling points in each field, resulting in 34, 33, and 32 total sampling points in each year. Nitrogen fertilizer was applied 1–3 weeks before rice transplanting as basal dressing and at the panicle initiation stage as top dressing (Table I). The condition of nitrogen in the experimental fields was modeled on commercial base fields.

TABLE I
Nitrogen Treatment in Each Field for 3 Years

Year	Field number	Sampling points	Basal (kg N ha^{-1})	Top-dressing (kg N ha^{-1})	Total (kg N ha^{-1})
2008	1	10	65	68	133
2008	2	8	60	68	128
2008	3	8	45	76	121
2008	4	8	48	74	122
Total sampling points in 2008 = 34					
2009	1	8	63	78	141
2009	2	8	61	80	141
2009	3	8	65	83	148
2009	4	9	77	79	156
Total sampling points in 2009 = 33					
2010	1	8	72	55	127
2010	2	8	70	53	123
2010	3	8	73	55	128
2010	4	8	75	53	128
Total sampling points in 2010 = 32					

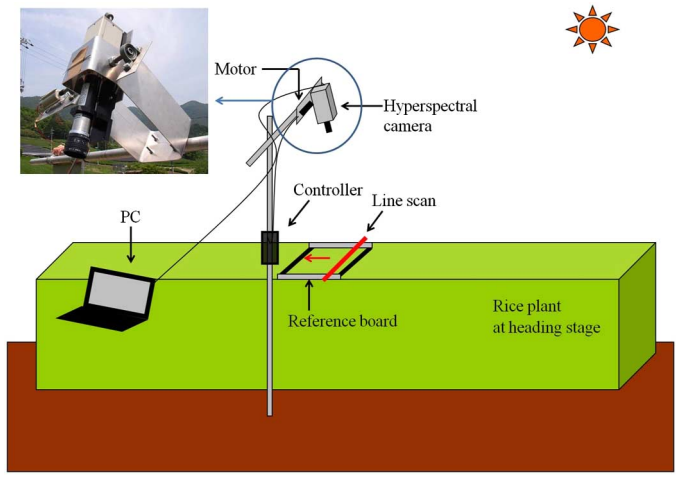


Fig. 1. Ground-based hyperspectral imaging system.

B. Ground-Based Hyperspectral Imaging System

Fig. 1 shows the experimental setup for a ground-based hyperspectral imaging system. All images were taken at the heading stage using a hyperspectral camera (ImSpector QE V10E; Specim, Oulu, Finland) placed approximately 2 m above the rice canopy.

The system consisted of a prism-grating-prism (PGP) component and a monochrome camera (C8484-05G; Hamamatsu Photonics K. K., Hamamatsu, Japan). The slit width of PGP was 50 μm and the slit length was 9.8 mm. The nominal spectral range was from 400 to 1000 nm with 5 nm nominal spectral resolution. The spatial and spectral pixels in the image frame were 1344 and 1024, respectively. A C-mount lens (Cosmicar, $f = 16$ mm, 1:1.4) was installed in the front of the camera. The aperture was

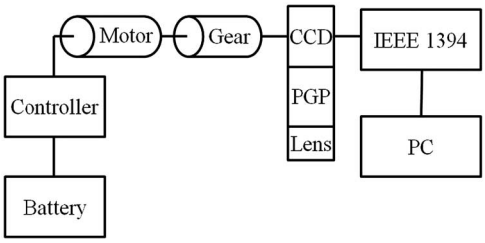


Fig. 2. Schematic representation of arrangement around hyperspectral camera.

set to F/2.8 and the focus was fixed at 2 m during image acquisition. The camera was connected to a planetary gear box and motor (TAM72001; Tamiya, Inc., Shizuoka, Japan) for constant control of camera angular speed at $0.004\ \mathrm{rad\ s^{-1}}$. Hyperspectral images were acquired from 10:00 A.M. to 2:00 P.M. to reduce the influence of the incident light angle and dew on the leaf surfaces. Reference boards (gray panel) were placed at the sides of the rice plant canopy to correct changes in the intensity of incoming light during image acquisition. Fig. 2 shows the schematic representation of the hyperspectral camera setup.

C. Image Acquisition

Images were acquired by a digital camera controller (Gnome IEEE-1394; Coriander version 1.0.1, Damien Douxchamps), a Linux open source program with a graphical user interface. Sensor sensitivity was manually controlled via Coriander software so as not to saturate during image acquisition under incoming light change. After scans were completed, the spatial–spectral matrices were combined to construct a 3-D spatial and spectral data cube. The spatial resolution of the images was approximately $2\ \mathrm{mm} \times 2\ \mathrm{mm}$ at 2 m above the rice canopy.

D. Rice Plant Sampling and Chemical Analysis

After the hyperspectral images were obtained, six rice plants without roots (aboveground biomass) were removed at each sampling point and the ground areas occupied by the plants were measured. The samples were then mixed and divided into two portions in a 1:2 proportion by weight. Samples from the smaller fraction were separated into leaves, stems, and heads and were analyzed to determine their ratio and nitrogen concentration. All samples were dried in a circulating drying oven at 80 °C for at least 72 h and then weighed [26]. The respective dry masses of the leaves, stems, and heads from the larger fraction were calculated from those of the smaller fraction. The leaves and stems of the smaller fraction were finely ground using a pulverizing mill (TI-100; Cosmic Mechanical Technology Co., Ltd., Iwaki, Japan). The nitrogen concentration of each fraction, except for the head fractions, was determined three times by gas chromatography (SUMIGRAPH NC-22F; Sumika Chemical Analysis Service, Ltd., Tokyo, Japan). The nitrogen content at each sampling point was calculated as a product of the dry mass per unit area and the mean nitrogen concentration. The experimental timetable is given in Table II. Daily average temperature data were obtained from the Japan Meteorological Agency.

TABLE II
FIELD MANAGEMENT AND WEATHER CONDITIONS

	2008	2009	2010
Management date			
Transplanting	25 May	24 May	27 May
Ground-based hyperspectral imaging	12 August	10 August	15 August
Sampling at the heading stage	13 August	11 August	16 August
Weather condition			
Accumulated vegetation day-period (day) [a]	80	79	81
Growing degree-days (°C) [b]	1117	1046	1155

[a]Day from transplanting to sampling.
[b]Sum of the daily average temperature minus $10^{\circ}\mathrm{C}$ from transplanting to sampling.

E. Image Processing

Images were analyzed using the Environment for Visualizing Images (ENVI) software (version 4.7; ITT Visual Information Solutions, Boulder, CO, USA). Each image was separated into two parts: 1) the rice plant (leaves, stems, and heads); and 2) other elements (e.g., irrigation water and soil background) by the *GreenNDVI–NDVI* value as shown in (1). *GreenNDVI–NDVI* was suggested as a method to distinguish the area of new shoots from old leaves in a green tea canopy [29]. The *GreenNDVI–NDVI* might produce good results than the NDVI because of excluding the leaf area under leaf shadow or the leaf at lower part of rice plant

$$GreenNDVI - NDVI = \frac{DN_{NIR} - DN_{green}}{DN_{NIR} + DN_{green}} - \frac{DN_{NIR} - DN_{red}}{DN_{NIR} + DN_{red}} \quad (1)$$

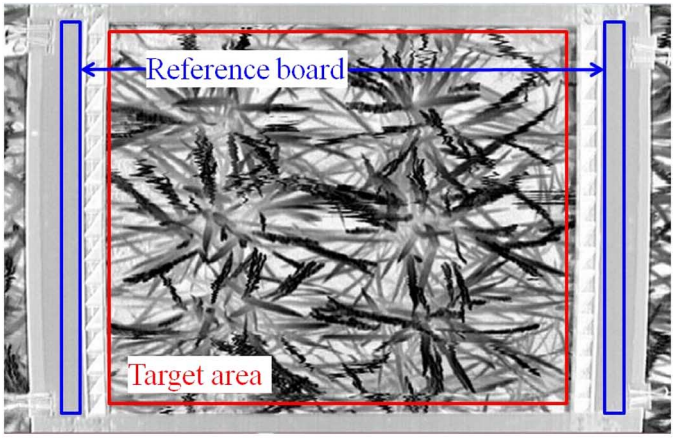
where DN represents the digital number of an NIR (845 nm), green (564 nm), or red (668 nm) spectrum, indicating the brightness of each spectrum.

Fig. 3(a) shows a *GreenNDVI–NDVI* image. The area of gray or black color corresponds to the rice plants and those of white color to the irrigation water or soil background. The threshold value was manually set to select only the rice plant area in the individual image.

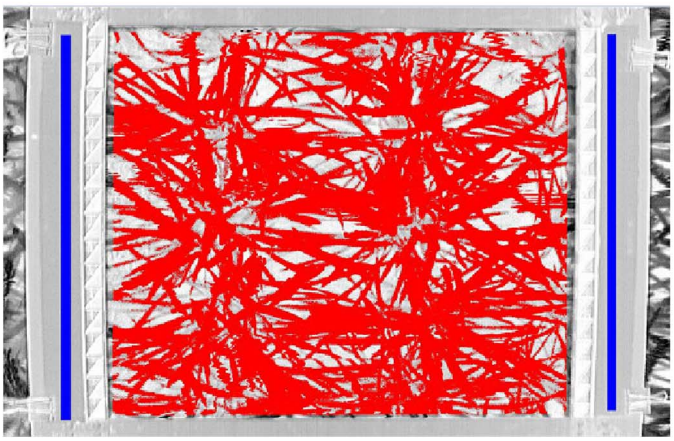
Fig. 3(b) shows the regions of interest (ROIs) in an image based on the threshold value of *GreenNDVI–NDVI*. A large part of the rice plant appears as the area of red color in this figure. The reflectance of the rice plant (ROI_RICE; red color) and the reflectance of the reference board area (ROI_BOARD; blue color) were extracted using the ROI function of the ENVI software. Then, we obtained 1024 spectral information of rice and reference board. Ref_{RICE} was calculated using the ratio of the mean value of the reflectance for the rice plant area to that of the reference board area, as shown as follows:

$$Ref_{RICE} = \frac{DN_{ROI\_RICE}}{DN_{ROI\_BOARD}} \quad (2)$$

where Ref_{RICE} represents the rice plant reflectance, and DN_{ROI_RICE} and DN_{ROI_BOARD} represent the mean values of the digital readings for rice plants and reference boards in the ROI.



(a)



(b)

Fig. 3. Composite grayscale image using the value of *GreenNDVI–NDVI* (a) and the region of interest (ROI) (b).

F. PLS Regression Analysis and Mutual Prediction

A PLS regression model for the relationship between the Ref_{RICE} and nitrogen content of the rice plant at the heading stage was generated using the statistics software R (version 2.13.0; R Foundation for Statistical Computing). The model was calculated using twofold cross-validation to evaluate the quality of the calibration model. In this method, the data is first partitioned into two equally (or nearly equally) sized segments or folds. The optimum number of principal components was determined by the method of Osten [30] given by (3). The criterion is based on testing the significance of incremental changes in predicted residual sum of squares (PRESS) with an F-test

$$F = \frac{PRESS_k - PRESS_{k+1}}{df_k - df_{k+1}} \Big/ \frac{PRESS_{k+1}}{df_{k+1}} \tag{3}$$

where $PRESS_k$ denotes the PRESS after including the first k principal components and df_k is the degree of freedom of $PRESS_k$. The criterion is compared against the $F(df_k - df_{k+1}, df_{k+1}; 0.95)$. This method may provide more robust performance than the local minimum in PRESS method. The criterion is based on testing the significance of incremental changes in PRESS with an F-test. The performance of the model was evaluated using the coefficient of determination (R^2) and the root-mean-square error (RMSE) of prediction as the indicators of the average error of analysis, as shown in (4). The predictive ability of the model was estimated using the relative error (RE), as shown in (5). RE represents the ratio of RMSE to the mean value for the nitrogen content. Mutual prediction was calculated by predicting the values of remaining years using the single- or multiyear (2-year) models. This procedure was performed to evaluate the predictive utility of a model considered growing condition.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$RE = \frac{100}{\bar{y}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5}$$

where y_i, $\hat{y}_i$, and $\bar{y}_i$ are the measured, predicted, and mean values of the nitrogen content, respectively, and n is the number of samples.

G. Analysis Models, R-PLS, and RT-PLS

A relationship between reflectance of rice plant and its nitrogen content was modeled with PLS regression (R-PLS model). This model was the ordinary regression one using remote-sensing data. For these variables, however, the accuracy of mutual prediction varied depending on the differences in several conditions among years [25]. R-PLS model incorporating temperature (RT-PLS) was constructed, based on the association between rice nitrogen content and cumulative temperature as well as reflectance. RT-PLS was expected to improve the accuracy of mutual prediction by reducing the influence of weather conditions. The equations of the R-PLS and RT-PLS are as follows:

$$\hat{y}_i = b_0 + \sum_{k=1}^{1024} b_k x_{ik} \tag{6}$$

$$\hat{y}_i = b_0 + \sum_{k=1}^{1024} b_k x_{ik} + b_{1025} g_{dd\_t} \tag{7}$$

where $\hat{y}_i$ is the predicted values of the nitrogen content; b_0 is the intercept; and b_k and b_{1025} are the PLS regression coefficients for a model with optimum number of principal components. x_{ik} is the reflectance of the kth spectral band and g_{dd-t} is GDDs under the base temperature t. Compared to the R-PLS model, the RT-PLS model has the growing temperature factor. We showed the flow of the experiment and statistical modeling in Fig. 4.

H. Temperature Data and GDDs

Daily average temperature data were obtained from the Japan Meteorological Agency. To constructing RT-PLS model, GDD was calculated as shown as follows [28]:

$$\mathrm{GDD} = \left(\frac{\mathrm{T_{MAX}} + \mathrm{T_{MIN}}}{2}\right) - \mathrm{T_{BASE}} \tag{8}$$

where $\mathrm{T_{MAX}}$ and $\mathrm{T_{MIN}}$ are daily maximum and minimum air temperature, respectively, and $\mathrm{T_{BASE}}$ is the base temperature. $\frac{\mathrm{T_{MAX}}+\mathrm{T_{MIN}}}{2}$ was replaced by daily average temperature.

TABLE III
SUMMARY OF NITROGEN CONTENT IN RICE PLANT AT THE HEADING STAGE

Field	2008			2009			2010		
	n	Mean (g m^{-2})	SD	n	Mean (g m^{-2})	SD	n	Mean (g m^{-2})	SD
1	5	6.92	0.69	8	8.21	1.27	6	8.05	0.84
2	6	11.93	1.58	8	9.37	1.29	6	8.19	0.99
3	6	12.68	1.71	5	9.80	1.87	5	8.89	1.81
4	7	9.58	1.80	8	8.14	1.09	8	7.49	0.40
Total	24	10.39	2.63	29	8.79	1.45	25	8.07	1.09

n, number of samples; SD, standard deviation.

TABLE IV
RESULT OF SINGLE-YEAR R-PLS MODEL CALIBRATION AND TWOFOLD CROSS-VALIDATION

	n	PCs	Calibration			Validation		
			R^2	*RMSE*	*RE*	R^2	*RMSE*	*RE*
2008	24	2	0.88	0.88	8.5	0.81	1.13	10.9
2009	29	3	0.78	0.67	7.6	0.47	1.04	11.8
2010	25	3	0.65	0.63	7.8	0.42	0.81	10.1

n, number of samples; PCs, number of principal components; R^2, coefficient of determination; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

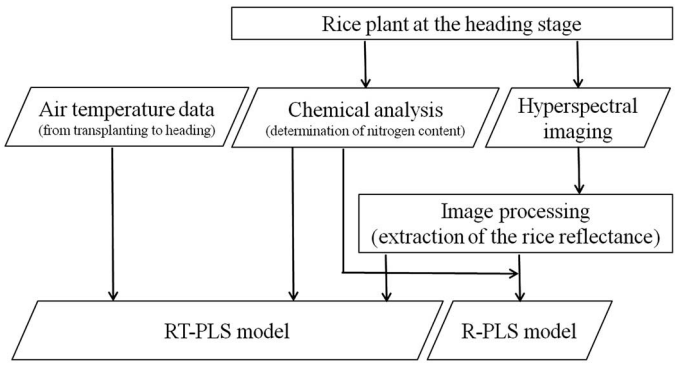


Fig. 4. Flowchart of the experiment and statistical modeling.

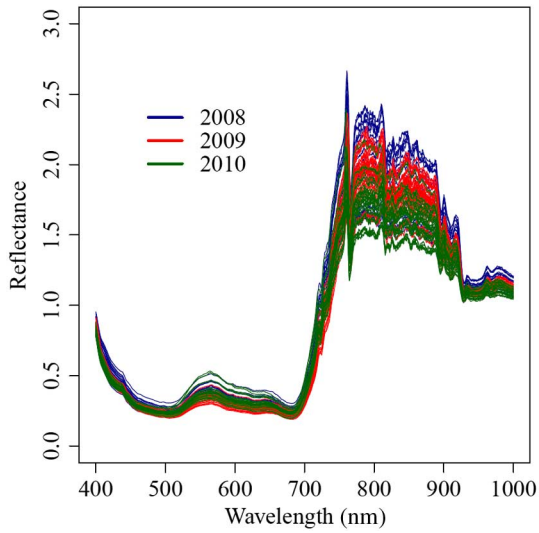


Fig. 5. Rice plant reflectance at the heading stage extracted by *GreenNDVI–NDVI*.

III. RESULTS AND DISCUSSION

A. Measurement of Vegetation Data and Canopy Reflectance

Table III shows the measured nitrogen content at the heading stage for 3 years. Due to strong wind during image acquisition, some images were defective and therefore were excluded from the analysis. Thus, the number of images was 24 in 2008, 29 in 2009, and 25 in 2010.

Mean nitrogen content ranged from 6.92 to 12.68 gm^{-2} in 2008, 8.14 to 9.80 gm^{-2} in 2009, and 7.49 to 8.89 gm^{-2} in 2010 among fields. The differences among years are significant ($p < 0.001$; ANOVA). Result of multiple comparison reveal presence of significant differences in the mean nitrogen content among years with a 1% level of significance.

Fig. 5 shows the mean reflectance of a rice plant (Ref_{RICE}), which was extracted by *GreenNDVI–NDVI*. The value of reflectance on y-axis is over 1 due to the use of gray color reference board. The number of lines was from 600 to 1000 and the image acquisition took from 2 to 3 min. The reflectance of a rice plant is moderately variable in the visible regions and variable in the NIR regions. This result is in accordance with previous observation [7], [14], [24].

B. Single-Year R-PLS Model Performance

The hyperspectral reflectance was related to the nitrogen content of rice plants at the heading stage through PLS analysis. Table IV indicates that single-year R-PLS models may be considered to have good practical accuracy.

The range in validation precision (R^2) from 0.42 to 0.81 may be due to differences among years in the variances of nitrogen content (Table III). Compared to the calibration results, precision and accuracy (RMSE and RE) in validation are slightly lower, but these models may still be useful to predict the nitrogen content of rice plants at the heading stage, as shown in Fig. 6. Accordingly, the mutual predictions of each single-year R-PLS model were calculated using the data in other years as the test set to determine the predictive potential of a model considered growing condition.

C. Mutual Prediction With the Single-Year R-PLS Models

In mutual predictions with the single-year R-PLS models, RMSE range from 1.40 to 5.82 gm^{-2} and RE range from 17.3% to 62.7% (Table V). RMSE and RE for mutual prediction are much higher than those for validation. The best result in mutual prediction is that 2010 data is predicted by 2008 single-year R-PLS model; RMSE and RE are 1.40 gm^{-2} and 17.3%, respectively. However, when 2009 data are predicted with the 2008 single-year R-PLS model, RMSE and RE increase to 4.34 gm^{-2} and 49.4%, respectively.

Thus, as a model considered growing condition, the 2010 single-year R-PLS model might be better than the 2008 model. The reason is that RMSE is relatively low, and there are no significant differences between the RMSE when the data of 2008 and 2009 are predicted by the 2010 model.

Fig. 7 shows the relationships between measured and predicted nitrogen content for the mutual prediction. When the 2008 single-year R-PLS model is used to predict the 2009 and 2010

TABLE V
RESULT OF MUTUAL PREDICTION WITH SINGLE-YEAR R-PLS MODELS

Validation					Mutual prediction			
Model	*n*	PCs	*RMSE*	*RE*	Data	*n*	*RMSE*	*RE*
2008	24	2	1.13	10.9	2009	29	4.34	49.4
2008	24	2	1.13	10.9	2010	25	1.40	17.3
2009	29	3	1.04	11.8	2008	24	5.82	57.9
2009	29	3	1.04	11.8	2010	25	5.06	62.7
2010	25	3	0.81	10.1	2008	24	2.78	27.7
2010	25	3	0.81	10.1	2009	29	2.49	28.3

n, number of samples; PCs, number of principal components; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

data, the 2009 data are overestimated and the 2010 data are underestimated. With the 2009 model, both 2008 and 2010 data are underestimated. Similarly, the 2010 model prediction overestimated the 2009 data. It is desirable to narrow the distance between the over- and underestimation from mutual prediction.

In model calibration, for the purpose of model adaptability, data collection is commonly conducted under varying nitrogen fertilization, irrigation, planting density, cultivar, and weather. Accordingly, an R-PLS model based on 2 years' data (the multiyear R-PLS model) was constructed to account for variation among years.

D. Multiyear R-PLS Model and Mutual Prediction

Table VI shows the result of PLS regression analysis for nitrogen content using multiyear R-PLS models. RMSE in validation ranges from 0.88 to 1.51 gm^{-2} and RE range from 9.9% to 15.9%. This result indicates that there are no significant differences in accuracy between the multi- and single-year R-PLS models.

Table VII and Fig. 8 show the result of mutual prediction with the multiyear R-PLS models. When the 2010 data are predicted by the multiyear R-PLS model based on the 2008 and 2009 data ($2008 + 2009$ multiyear R-PLS model), *RMSE* and *RE* are 1.11 gm^{-2} and 13.7%, respectively. Thus, the accuracy is superior to that of mutual prediction with the 2008 or 2009 single-year R-PLS models. When 2009 or 2008 data are predicted by the multiyear R-PLS model based on the remaining data, RMSE or RE is superior to the worst result of mutual prediction with the single-year R-PLS models. When the model constructed with big difference of weather condition among years, the accuracy of multiyear model is lower than the single-year.

The relationships between predicted and measured values for the multiyear R-PLS are approximately parallel to the 1:1 line [Fig. 8(b) and (c)], suggesting that the prediction ability might increase compared to the single-year R-PLS models. The multiyear model might thus be a potential model considered growing condition. However, the intercept, or distance between the 1:1 line and an approximate fitted line through the predicted values, is nonzero [Fig. 8(b) and (c)].

E. Development of the RT-PLS Model

Because all data tend to be overestimated or underestimated when nitrogen content is predicted by single-year R-PLS models in mutual predictions, prediction accuracy is reduced (Table V; Fig. 7). The large and small values of GDDs (Table II) follow this overestimation and underestimation. Specifically, single-year

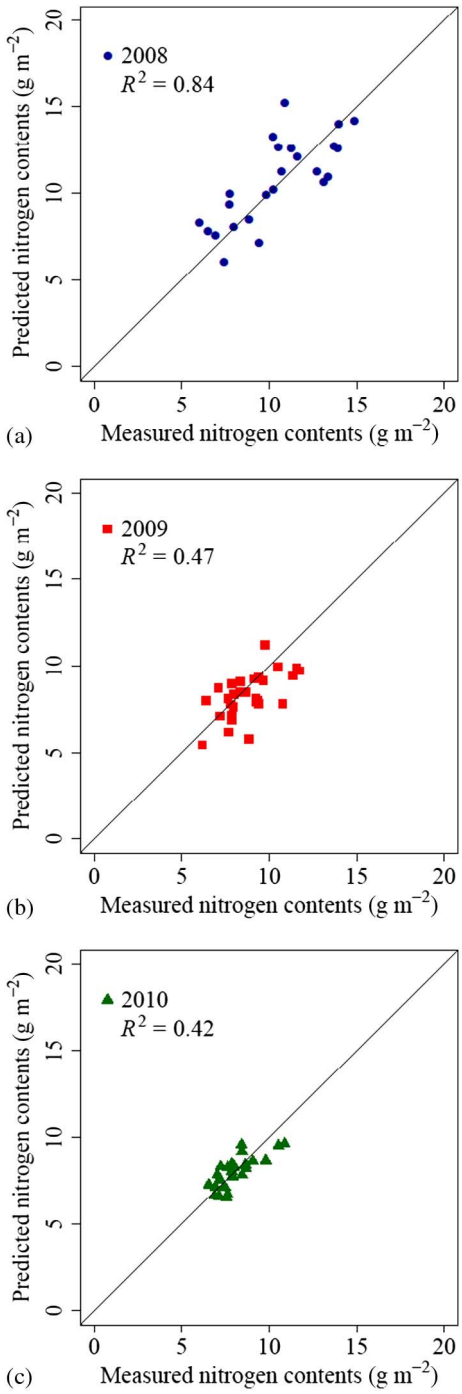


Fig. 6. Relationship between measured and predicted nitrogen content using twofold cross-validation for 2008 (a), 2009 (b), and 2010 (c).

R-PLS models constructed under relatively high GDD values tend to overestimate the relatively low GDD data. In contrast, single-year R-PLS models constructed under relatively low GDD values tend to underestimate relatively high GDD data. These observations suggest using the temperature and effective thermal data [31] to correct the influence of weather differences among years. GDD does not constrain the growth rate, but the developmental rate such as the rate of leaf emergence [32]. But the former rather than the latter corresponds to the amount of growth predicted by remote sensing. Although GDD mainly

TABLE VI
RESULT OF MULTIYEAR R-PLS MODEL CALIBRATION AND TWOFOLD CROSS-VALIDATION

Model	n	PCs	Calibration			Validation		
			R^2	*RMSE*	*RE*	R^2	*RMSE*	*RE*
2008+2009	53	3	0.71	1.17	12.7	0.52	1.51	15.9
2008+2010	49	4	0.91	0.69	7.5	0.84	0.91	9.9
2009+2010	54	4	0.74	0.67	7.9	0.55	0.88	10.4

$2008 + 2009$: 2008 and 2009 data for training set; $2008 + 2010$: 2008 and 2010 data for training set; $2009 + 2010$: 2009 and 2010 data for training set; n, number of samples; PCs, number of principal components; R^2, coefficient of determination; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

TABLE VII
RESULT OF MUTUAL PREDICTION WITH MULTIYEAR R-PLS MODELS

Calibration					Mutual prediction			
Model	n	PCs	*RMSE*	*RE*	Data	n	*RMSE*	*RE*
2008+2009	53	3	1.17	12.7	2010	25	1.11	13.7
2008+2010	49	4	0.69	7.5	2009	29	3.76	42.8
2009+2010	54	4	0.67	7.9	2008	24	4.15	40.0

$2008 + 2009$: 2008 and 2009 data for training set, and 2010 data for test set; $2008 + 2010$: 2008 and 2010 data for training set, and 2009 data for test set; $2009 + 2010$: 2009 and 2010 data for training set, and 2008 data for test set; n, number of samples; PCs, number of principal components; R^2, coefficient of determination; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

constrains the developmental rate, it also indirectly influences the amount of growth; the sooner the tillering stage begins due to relatively high GDD, the higher the increase in the number of leaves on the main axis (the amount of growth). In the results of mutual prediction, the amount of growth owing to the effect of GDD is over- or underestimated by ground-based hyperspectral imaging (400–1000 nm), despite the relationship between NIR reflectance and the amount of growth. The indirect influence of GDD on growth rate might be quantified to improve the accuracy of the mutual prediction. Accordingly, the number of PCs for multiyear RT-PLS models is fixed at the number used for multiyear R-PLS models in order to quantify variation in the amount of growth due to weather variation. Incorporation of GDD data in the R-PLS model required corresponding optimization.

GDD was calculated as the sum of average daily temperature subtracted from a base temperature in order to quantify the variation in the amount of growth due to weather differences among years. RMSE and standard error (SE) in the mutual prediction were calculated for base-temperature increments by 1 °C as shown in Fig. 9. Multiyear RT-PLS models were fitted and the data of remaining years were predicted. When the base temperature is 23 °C or 27 °C, RMSE and SE are at their minimum values (1.09 and 0.66 gm^{-2}, respectively).

When the base temperature is 25 °C, RMSE increases to 11 gm^{-2}. This is because the data in 2009 are predicted with high error (more than 18 gm^{-2} in RMSE) by the $2008 + 2010$ multiyear RT-PLS model.

F. Comparison of R-PLS and RT-PLS Models

In order to quantify the differences in weather conditions among years, PLS models were developed using reflectance data of 2 years combined with temperature data (RT-PLS model) and by comparing with PLS models established by only reflectance data (R-PLS model).

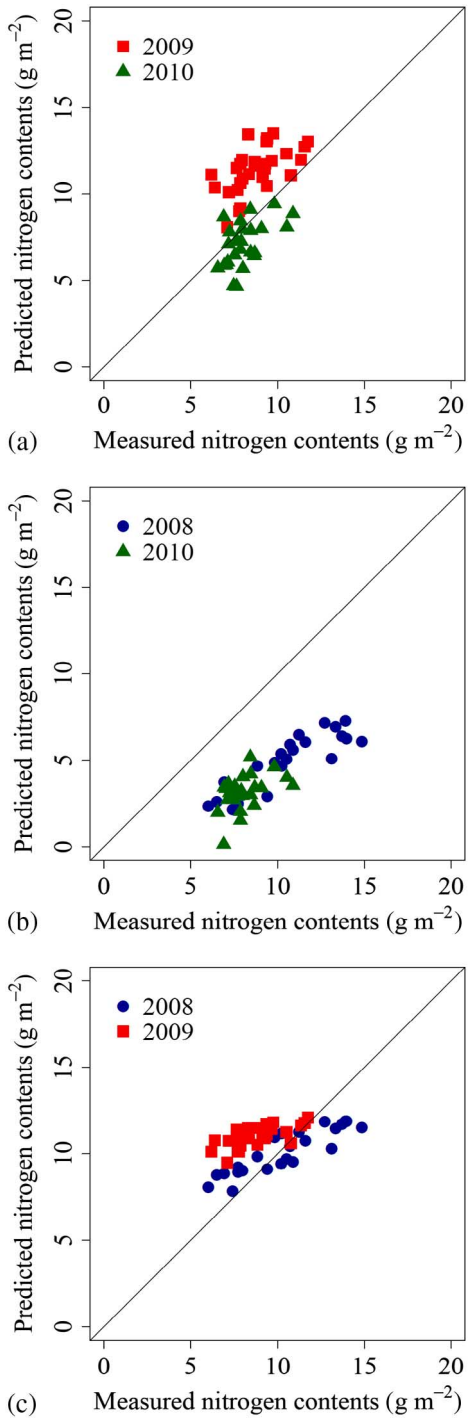


Fig. 7. Result of mutual prediction with the single-year R-PLS models: 2008 (a), 2009 (b), and 2010 (c). The single-year R-PLS model was used to predict the other years' data.

Table VIII shows the result of multiyear RT-PLS model calibration and validation. In validation, R^2 and RMSE range from 0.49 to 0.82 gm^{-2} and 0.95 to 1.10 gm^{-2}, respectively, while RE is approximately 11%. Thus, in terms of precision and accuracy in internal validation, there are no significant differences between the RT-PLS and R-PLS models.

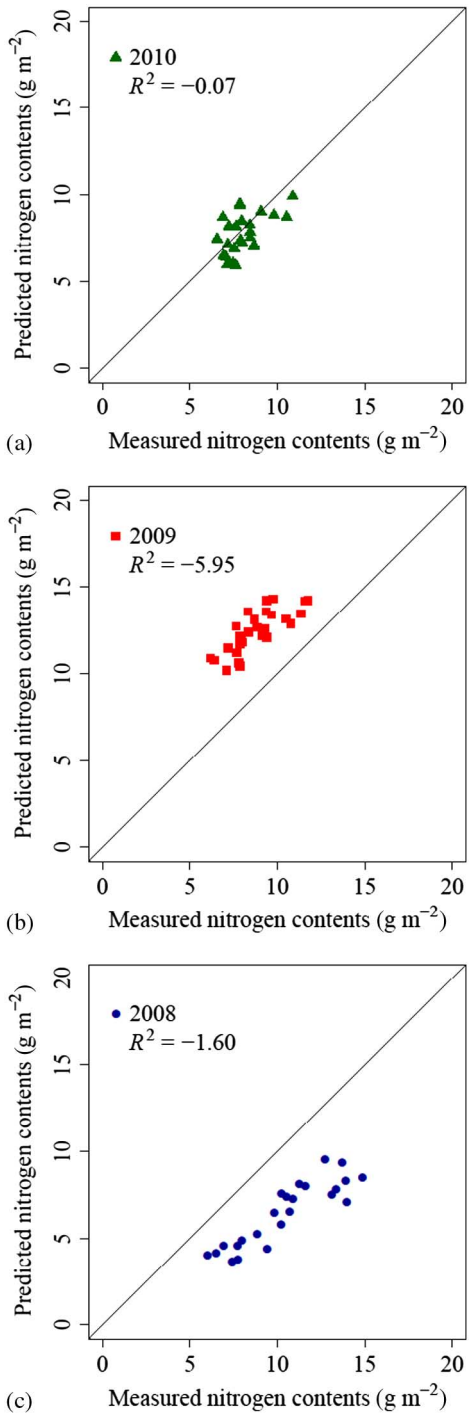


Fig. 8. Result of mutual prediction with multiyear R-PLS models: 2010 (a), 2009 (b), and 2008 (c). Data not used for the training set were predicted.

TABLE VIII
RESULT OF MULTIYEAR RT-PLS MODEL CALIBRATION AND TWOFOLD CROSS-VALIDATION

Model	n	PCs	Calibration			Validation		
			R^2	*RMSE*	*RE*	R^2	*RMSE*	*RE*
2008+2009	53	3	0.78	1.02	10.7	0.75	1.10	11.5
2008+2010	49	4	0.88	0.78	8.5	0.82	0.96	10.4
2009+2010	54	4	0.73	0.69	8.2	0.49	0.95	11.2

A: 2008 and 2009 data for training set; B: 2008 and 2010 data for training set; C: 2009 and 2010 data for training set; n, number of samples; PCs, number of principal components; R^2, coefficient of determination; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

TABLE IX
RESULT OF MUTUAL PREDICTION WITH MULTIYEAR RT-PLS MODELS

Calibration					Mutual prediction			
Model	n	PCs	*RMSE*	*RE*	Data	n	*RMSE*	*RE*
2008+2009	53	3	1.02	10.7	2010	25	0.99	12.3
2008+2010	49	4	0.78	8.5	2009	29	1.07	12.1
2009+2010	54	4	0.69	8.2	2008	24	1.21	11.6

$2008 + 2009$: 2008 and 2009 data for training set, and 2010 data for test set; $2008 + 2010$: 2008 and 2010 data for training set, and 2009 data for test set; $2009 + 2010$: 2009 and 2010 data for training set, and 2008 data for test set; n, number of samples; PCs, number of principal components; R^2, coefficient of determination; RMSE, root-mean-square error (gm^{-2}); RE, relative error (%).

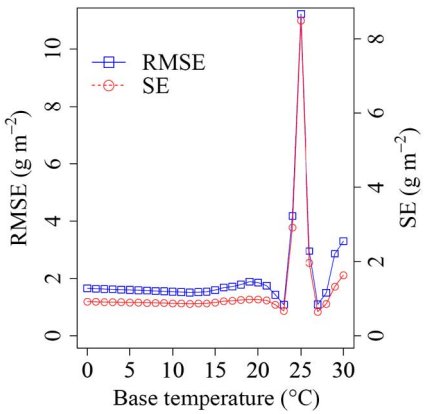


Fig. 9. Root-mean-square error (RMSE) and standard error (SE) in mutual prediction using multiyear RT-PLS models as a function of base temperature.

Table IX shows the result of mutual prediction using the multiyear RT-PLS models. When the 2010 data are predicted by the $2008 + 2009$ multiyear RT-PLS model, RMSE and RE decreased to $0.99\ \mathrm{gm}^{-2}$ and 12.3%, respectively. When the 2009 data are predicted by the $2008 + 2010$ multiyear RT-PLS model, RMSE and RE decreased to $1.07\ \mathrm{gm}^{-2}$ and 12.1%, respectively. Similarly, when the 2008 data are predicted by the $2009 + 2010$ multiyear RT-PLS model, RMSE and RE decrease to $1.21\ \mathrm{gm}^{-2}$ and 11.6%, respectively. Differences of the environmental factor such as weather conditions may not be considered by a PLS model derived from only hyperspectral data because in the result of mutual prediction, bias is nonzero as shown in Figs. 6 and 7. In the areas of crop phenology and development, the concept of heat units, expressed in GDD, is used to describe phenological events [28]. The red edge position of rice canopy reflectance moves to longer wavelength until close to heading stage and otherwise shifted to shorter wavelength after heading stage [27]. Hyperspectral sensor is more sensitive to the shift. The PLS regression coefficients of red edge region were selected as important variables [25]. The bias in the mutual prediction may be caused by phenological shift. Therefore, by incorporating optimized GDD, the accuracy of the RT-PLS model may be higher than that of the R-PLS model.

Fig. 10 shows the result of mutual prediction with the multiyear RT-PLS models. Compared with the result of the multiyear R-PLS model shown in Fig. 8, the intercepts are clearly reduced in magnitude. RE was decreased by 28.4% in 2008, 30.7% in 2009, and 1.4% in 2010. It means that the

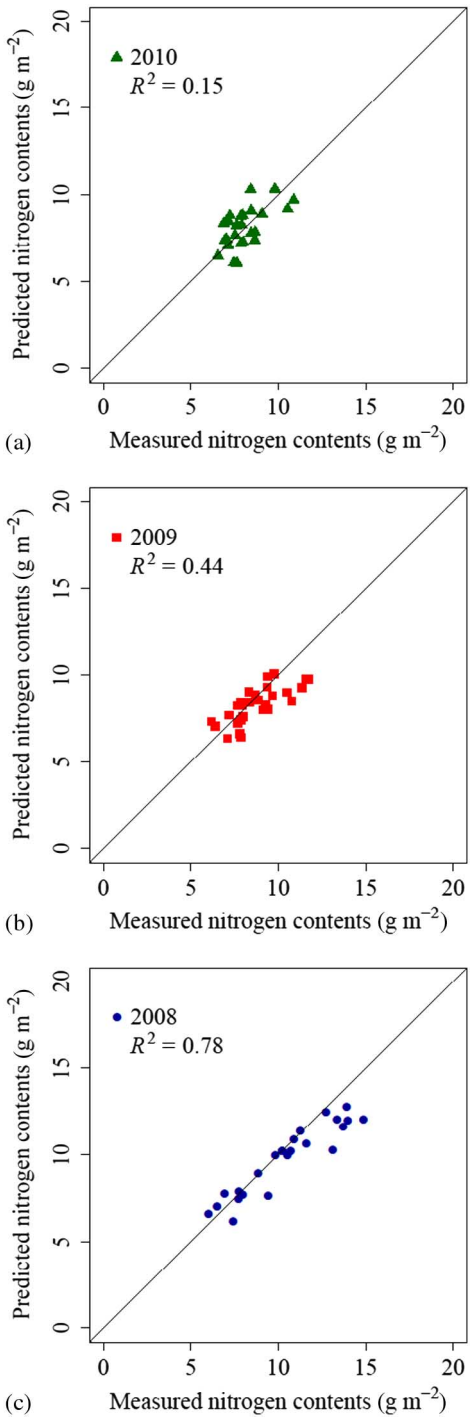


Fig. 10. Result of mutual prediction with multiyear RT-PLS models: 2010 (a), 2009 (b), and 2008 (c). Data not used for the training set were predicted.

temperature data are useful for establishment of a model-considered growing condition.

Further studies should investigate whether the RT-PLS model is useful for predicting the nitrogen content of rice plants at the panicle initiation stage and whether it can be applied to other crops. It should also be determined whether the inclusion of other factors such as daily range of temperature or radiation can improve the accuracy of the model.

IV. Conclusion

Ground-based hyperspectral imaging was applied to rice plants at the heading stage in order to estimate nitrogen content. The images were divided into two parts: 1) the rice plant; and 2) other elements (e.g., irrigation water and soil background) using the equation of *GreenNDVI–NDVI*. Ref_{RICE} was calculated as the ratio of the reflectance for the rice plant to that for a reference board. PLS regression was used to construct estimation models.

Single-year R-PLS models were established for each year. In validation, RMSE ranged from 0.81 to 1.13 gm^{-2} and RE were approximately 11%, indicating that accuracy was high and consistent among years. Accordingly, mutual prediction was performed to determine the predictive potential of a model considered growing condition. RMSE and RE values in the mutual prediction were higher (ranging from 1.40 to 5.82 gm^{-2}) than those in each year's validation. A multiyear R-PLS model based on the data of 2 years was then constructed to account for variation among years. Although the accuracy of the single-year R-PLS models did not significantly differ, the approximate 1:1 slopes of the relationships between predicted and measured values suggested that the predictive ability might be improved by fitting models incorporating data of 2 years.

Examination aimed at discovering a more accurate estimation model suggested a correspondence between GDD variation and N content over- and underestimation. It was inferred that hyperspectral imaging at 400–1000 nm could not predict variation in the amount of growth caused by weather variation expressed as GDD. Accordingly, the cumulative temperature data were included in the R-PLS model, leading to a model (RT-PLS) incorporating reflectance and temperature data of 2 years. The combined reflectance and temperature data were useful for constructing a model considered growing condition. The accuracy of mutual prediction was improved, with RMSE and RE less than 1.21 gm^{-2} and 12.3%, respectively.

Acknowledgment

The authors would like to thank Prof. T. Inamura of Kyoto University of Agriculture Sciences for assistance in experiment.

References

[1] OECD/FAO, *OECD-FAO Agricultural Outlook 2012-2021*, 2012.
[2] J.-I. Sato, K. Shigeta, and Y. Nagasaka, "Automatic operation of a combined harvester in a rice field," in *Proc. IEEE/SICE/RSJ Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, 1996, pp. 86–92.
[3] T. Coen *et al.*, "Autopilot for a combine harvester," *Comput. Electron. Agric.*, vol. 63, no. 1, pp. 57–64, 2008.
[4] S. Blackmore, "Precdision farming: An introduction," *Outlook Agric.*, vol. 23, no. 4, pp. 275–280, 1994.
[5] B.-W. Lee, "Spikelet number estimation model using nitrogen nutrition status and biomass at panicle initiation and heading stage of rice," *Korean J. Crop Sci.*, vol. 47, no. 5, pp. 390–394, 2002.
[6] D. Ntanos and S. Koutroubas, "Dry matter and N accumulation and translocation for Indica and Japonica rice under Mediterranean conditions," *Field Crops Res.*, vol. 74, no. 1, pp. 93–101, 2002.
[7] H. T. Nguyen and B.-W. Lee, "Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least square regression," *Eur. J. Agron.*, vol. 24, no. 4, pp. 349–356, 2006.

[8] N. Oppelt and W. Mauser, "Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data," *Int. J. Remote Sens.*, vol. 25, no. 1, pp. 145–159, 2004.
[9] F. M. Wang, J. F. Huang, and X. Z. Wang, "Identification of optimal hyperspectral bands for estimation of rice biophysical parameters," *J. Integr. Plant Biol.*, vol. 50, no. 3, pp. 291–299, 2008.
[10] L. Xue *et al.*, "Monitoring leaf nitrogen status in rice with canopy spectral reflectance," *Agron. J.*, vol. 96, no. 1, pp. 135–142, 2004.
[11] P. Chen *et al.*, "Critical nitrogen curve and remote detection of nitrogen nutrition index for corn in the northwestern plain of Shandong province, China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 682–689, Apr. 2013.
[12] J. G. P. W. Clevers and L. Kooistra, "Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 574–583, Apr. 2012.
[13] N. C. Sims *et al.*, "Towards the operational use of satellite hyperspectral image data for mapping nutrient status and fertilizer requirements in Australian plantation forests," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 320–328, Apr. 2013.
[14] X. Ye *et al.*, "A ground-based hyperspectral imaging system for characterizing vegetation spectral features," *Comput. Electron. Agric.*, vol. 63, no. 1, pp. 13–21, 2008.
[15] Y. Inoue, M. S. Moran, and T. Horie, "Analysis of spectral measurements in paddy field for predicting rice growth and yield based on a simple crop simulation model," *Plant Prod. Sci.*, vol. 1, pp. 269–279, 1998.
[16] Y. Inoue *et al.*, "Hyperspectral and directional remote sensing measurements of rice canopies for estimation of plant growth variables," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2000, pp. 1471–1473.
[17] J. White *et al.*, "Nitrogen concentration in New Zealand vegetation foliage derived from laboratory and field spectrometry," *Int. J. Remote Sens.*, vol. 21, no. 12, pp. 2525–2531, 2000.
[18] G. ElMasry *et al.*, "Quality classification of cooked, sliced turkey Hams using NIR hyperspectral imaging system," *J. Food Eng.*, vol. 103, no. 3, pp. 333–344, 2011.
[19] J. Li, X. Rao, and Y. Ying, "Detection of common defects on oranges using hyperspectral reflectance imaging," *Comput. Electron. Agric.*, vol. 78, no. 1, pp. 38–48, 2011.
[20] H. Itoh *et al.*, "Measurement of nitrate concentration distribution in vegetables by near-infrared hyperspectral imaging," *Environ. Control Biol.*, vol. 48, no. 2, pp. 37–49, 2010.
[21] Y. Suzuki *et al.*, "Mapping the spatial distribution of botanical composition and herbage mass in pastures using hyperspectral imaging," *Grassl. Sci.*, vol. 58, no. 1, pp. 1–7, 2012.
[22] K. Uto *et al.*, "Characterization of rice paddies by a UAV-mounted miniature hyperspectral sensor system," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 851–860, Apr. 2013.
[23] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: A basic tool of chemometrics," *Chemom. Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, 2001.
[24] P. Hansen and J. Schjoerring, "Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression," *Remote Sens. Environ.*, vol. 86, no. 4, pp. 542–553, 2003.
[25] C. Ryu, M. Suguri, and M. Umeda, "Multivariate analysis of nitrogen content for rice at the heading stage using reflectance of airborne hyperspectral remote sensing," *Field Crops Res.*, vol. 122, no. 3, pp. 214–224, 2011.
[26] C. Ryu, M. Suguri, and M. Umeda, "Model for predicting the nitrogen content of rice at panicle initiation stage using data from airborne hyperspectral remote sensing," *Biosyst. Eng.*, vol. 104, no. 4, pp. 465–475, 2009.
[27] M. Evri, T. Akiyama, and K. Kawamura, "Spectrum analysis of hyperspectral red edge position to predict rice biophysical parameters and grain weight," *Jpn. Soc. Photogramm. Remote Sens.*, vol. 47, no. 2, pp. 4–15, 2008.
[28] G. S. McMaster and W. Wilhelm, "Growing degree-days: One equation, two interpretations," *Agric. Forest Meteorol.*, vol. 87, no. 4, pp. 291–300, 1997.
[29] C. Ryu, M. Suguri, and M. Umeda, "Estimation of the quantity and quality of green tea using hyperspectral sensing," *J. Soc. Agric. Mach.*, vol. 72, pp. 46–53, 2010.
[30] D. W. Osten, "Selection of optimal regression models via cross validation," *J. Chemom.*, vol. 2, no. 1, pp. 39–48, 1988.
[31] M. Ebata, "Effective heat unit summation and base temperature on the development of rice [*Oryza sativa*] plant, 2: On heading, flowering, kernel development and maturing of rice," *Jpn. J. Crop Sci.*, vol. 59, pp. 233–238, 1990.
[32] F. W. Went, *The Experimental Control of Plant Growth: With Special Reference to the Earhart Plant Research Laboratory at the California Institute of Technology*. New York, NY, USA: Ronald, 1957.

Hiroyuki Onoyama received the B.S. and M.S. degrees in agriculture from Kyoto University, Kyoto, Japan, in 2009 and 2011, respectively. Currently, he is pursuing the Ph.D. degree at the Graduate School of Agriculture, Kyoto University.

Since 2012, he has received Research Grants for Japan Society for the Promotion of Science Fellows. His research interests include field-based remote sensing for crop management and the development of autonomous unmanned aerial vehicle as remote sensing platforms.

Chanseok Ryu received the B.S. and M.S. degrees in agricultural engineering from Gyeonsang National University, South Korea, and from Seoul National University, in 1998 and 2000, respectively, and the Ph.D. degree in agriculture from Kyoto University, Kyoto, Japan, in 2004.

Since 2012, he is an Assistant Professor with the Gyeongsang National University. His research interests include the quantitative of phytochemicals in the field using hyperspectral imagery and the management control using GIS database in agriculture.

Masahiko Suguri received the B.S. and M.S. degrees in agricultural engineering from Kyoto University, Kyoto, Japan, in 1988 and 1990, respectively, and the Ph.D. degree in agriculture from Kyoto University in 2005.

Currently, he is an Assistant Professor with the Division of Environmental Science and Engineering, Graduate School of Agriculture, Kyoto University. His research interests include machine vision, precision agriculture, and remote sensing.

Michihisa Iida received the B.S. and M.S. degrees in agricultural engineering from Kyoto University, Kyoto, Japan, in 1989 and 1991, respectively, and the Ph.D. degree in agriculture from Kyoto University in 1997.

Currently, he is a Professor with the Division of Environmental Science and Engineering, Graduate School of Agriculture, Kyoto University. His research interests include autonomous navigation and mapping, mobile robotics, and informatics in agriculture.