Available online at www.sciencedirect.com
SCIENCE DIRECT®
Remote Sensing of Environment 86 (2003) 542–553
Remote Sensing of Environment
www.elsevier.com/locate/rse

Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression

P.M. Hansen[a], J.K. Schjoerring[b,*]

[a] Risoe National Laboratory, Plant Research Department, Frederiksborgvej 399, Box-49, DK-4000 Roskilde, Denmark
[b] Plant Nutrition Laboratory, Department of Agricultural Sciences, Royal Veterinary and Agricultural University, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark

Received 6 November 2002; received in revised form 23 April 2003; accepted 3 May 2003

Abstract

Hyperspectral reflectance (438 to 884 nm) data were recorded at five different growth stages of winter wheat in a field experiment including two cultivars, three plant densities, and four levels of N application. All two-band combinations in the normalized difference vegetation index $(\lambda 1 - \lambda 2)/(\lambda 1 + \lambda 2)$ were subsequently used in a linear regression analysis against green biomass (GBM, g fresh weight m^{-2} soil), leaf area index (LAI, m^2 green leaf m^{-2} soil), leaf chlorophyll concentration (Chl_{conc}, mg chlorophyll g^{-1} leaf fresh weight), leaf chlorophyll density ($Chl_{density}$, mg chlorophyll m^{-2} soil), leaf nitrogen concentration (N_{conc}, mg nitrogen g^{-1} leaf dry weight), and leaf nitrogen density ($N_{density}$, g nitrogen m^{-2} soil). A number of grouped wavebands with high correlation (R^2>95%) were revealed. For the crop variables based on quantity per unit surface area, i.e. GBM, LAI, $Chl_{density}$, and $N_{density}$, these wavebands had in the majority (87%) of the cases a center wavelength in the red edge spectral region from 680 to 750 nm and the band combinations were often paired so that both bands were closely spaced in the steep linear shift between R_{red} and R_{nir}. The red edge region was almost absent for bands related to Chl_{conc} and N_{conc}, where the visible spectral range, mainly in the blue region, proved to be better. The selected narrow-band indices improved the description of the influence of all six-crop variables compared to the traditional broad- and short-band indices normally applied on data from satellite, aerial photos, and field spectroradiometers. For variables expressed on the basis of soil or canopy surface area, the relationship was further improved when exponential curve fitting was used instead of linear regression. The best of the selected narrow-band indices was compared to the results of a partial least square regression (PLS). This comparison showed that the narrow-band indices related to LAI and Chl_{conc}, and to some extent also $Chl_{density}$ and $N_{density}$, were optimal and could not be significantly improved by PLS using the information from all wavelengths in the hyperspectral region. However, PLS improved the prediction of GBM and N_{conc} by lowering the RMSE with 22% and 24%, respectively, compared to the best narrow-band indices. It is concluded that PLS regression analysis may provide a useful exploratory and predictive tool when applied on hyperspectral reflectance data.

Keywords: Remote sensing; Hyperspectral reflectance; Winter wheat; GBM; LAI; Chlorophyll; Nitrogen; NDVI; PLS; Wavelength band selection

1. Introduction

Measurement of various crop canopy variables during the growing season provides an opportunity for improving grain yields and quality by site-specific application of fertilizers. Important variables in this context are leaf area and total aboveground biomass (Asseng, van Keulen, & Stol, 2000; Jamieson et al., 1998). Also, leaf chlorophyll and nitrogen concentration in the leaf dry matter are indicators for crop nitrogen requirements. The spatial and temporal variations in the field of these variables must be determined in order to match the crop requirements as closely as possible.

Different remote sensing applications have proved to be a potential source of reflectance data for estimation of several canopy variables related to biophysical, physiological or biochemical characteristics (Ahlrichs & Bauer, 1983; Fernández, Vidal, Simón, & Solé-Sugrañes, 1994; Serrano, Filella, & Penuelas, 2000; Thenkabail, Smith, & de Pauw, 2001; Wiegand et al., 1992; Yoder & Pettigrew-Crosby, 1995). Hyperspectral remote sensing, or imaging spectroscopy, consists of acquiring images in many (<10 nm)

* Corresponding author. Tel.: +45-35283495; fax: +45-35283460.
E-mail address: jks@kvl.dk (J.K. Schjoerring).

0034-4257/03/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S0034-4257(03)00131-7

narrow, contiguous spectral bands, thus providing a continuous spectrum for each pixel, unlike multi-spectral systems that acquire images in a few (>50 nm) broad spectral bands. Hyperspectral imaging is a powerful and versatile tool for continuous sampling and for selection of narrow wavebands, which are sensitive to specific crop variables. The possibilities for using the technique have increased over recent years, because the required spectrometers have become cheaper.

Selection of new wavebands in hyperspectral imaging has been performed in a number of cases, mainly focusing on how to increase the sensitivity of the vegetation indices to chlorophyll and other pigments (Blackburn, 1998b). These investigations have mainly been performed on leaf level, on canopies grown under controlled conditions or on vegetation very distinct from winter wheat with respect to leaf structure, canopy geometry, and background due to senescent vegetation (Broge & Leblanc, 2001; Gamon, Penuelas, & Field, 1992; Gitelson, Merzlyak, & Lichtenthaler, 1996; Penuelas, Camon, Griffin, & Field, 1993; Penuelas, Filella, & Gamon, 1995; Penuelas, Gamon, Fredeen, Merino, & Field, 1994; Yoder & Waring, 1994). As a consequence of the different measurement conditions, some degree of disagreement exists in the selection of wavebands. A limited number of results have been published where hyperspectral reflectance was recorded under natural conditions in high input and high yielding crops (Broge & Mortensen, 2002; Filella, Serrano, Serra, & Penuelas, 1995; Serrano et al., 2000).

The normalized difference vegetation index (NDVI) is the classical index, where red reflectance (R_{red}) and near-infrared reflectance (R_{nir}) is used ($NDVI = R_{nir} - R_{red}/R_{nir} + R_{red}$). This vegetation index has been related to crop variables such as biomass, leaf area, plant cover, leaf gap fraction, nitrogen, and chlorophyll in cereals (Aparicio, Villegas, Casadesus, Araus, & Royo, 2000; Best & Harlan, 1985; Christensen & Goudriaan, 1993; Lukina, Stone, & Raun, 1999). NDVI have during the past decades been based on either broad wavebands (50–100 nm scale) from, e.g. the satellite-based Landsat Thematic Mapper using the TM-spectrometer (TM), or short wavebands (10 nm scale) from field-based spectroradiometers like Skye Sensors (Skye Instruments, Scotland), CropScan MSR87 and MSR16 (CropScan, USA). The broadband VIs use, in principle, average spectral information over a wide range resulting in loss of critical spectral information available in specific narrow (hyperspectral) bands. Further improvement in indices is generally obtained through the use of spectral data from distinct short bands. Some studies on hyperspectral reflectance data from field based spectrometers have been conducted on both leaf and canopy level in various plants (Blackburn, 1998a,b; Thenkabail et al., 2001; White, Trotter, Brown, & Scott, 2000; Yoder & Pettigrew-Crosby, 1995). The results obtained in these studies indicate that knowledge of the connection between the investigated variable and the spectral data can improve the performance of vegetation indices (Gitelson, Kaufman, & Merzlyak, 1996; Gitelson & Merzlyak, 1997; Maccioni, Agat, & Mazzinghi, 2001; Thenkabail et al., 2001).

The objective of the present investigation was to compare the predictive power of (i) models based on predefined short and broadbands for a normalized difference type of index, (ii) the best combination of narrow wavebands for a normalized difference type of vegetation index, and (iii) partial least squares regression (PLS) using all available wavebands. The models were fitted to six different biophysical/biochemical variables; green biomass (GBM), leaf area index (LAI), leaf chlorophyll concentration (Chl_{conc}), leaf chlorophyll density ($Chl_{density}$), leaf nitrogen concentration (N_{conc}), and leaf nitrogen density ($N_{density}$) in winter wheat. The registrations of GBM, LAI, leaf chlorophyll, and nitrogen content were performed in a plot experiment with different cultivars, plant densities, and nitrogen supplies. The spectral recordings and measurements of crop variables were performed five times from early stem elongation until heading to include the variation during vegetative growth. This procedure ensured that the normally occurring variation due to canopy growth stage and management factors was included in the models, giving a more realistic basis for model development.

2. Materials and methods

An experiment was carried out 2000/2001 at Risø National Laboratory, Roskilde, located at 55°41′N, 12°05′E. The crop rotation on the experimental field included winter oilseed rape, winter wheat (1st year), winter wheat (2nd year), spring barley, and winter barley. The straw was usually removed from the field except that of winter oilseed rape. No manure or slurry had been supplied to the field for centuries. The soil was typical for the region, i.e. a glacially created sandy loam with 10–15% clay and 2–5% organic matter in the topsoil. Average annual precipitation in the area is 650 mm. Average monthly rainfall in the growth period ranges from 40 mm in April to 65 mm in September. The distribution of the rainfall in the experimental area in year 2001 was close to normal except for a relatively dry July (25 mm) and a wet August (108 mm). The experimental area was managed as the surrounding field except for the experimental treatments, implying that suitable pesticides were used preventing weed and fungal disease problems.

2.1. Experimental design

The experiment was planned as a three-factor split–split–plot design with three complete randomized blocks as replicates. Two cultivars of winter wheat, i.e. Ritmo (erectophile) and Gunbo (planophile) were used. The cultivars were supposed to differ mainly in leaf orientation. Each cultivar was seeded to obtain three different plant densities 75, 150,

and 225 plants/m^2. Each of the plant density levels was furthermore split in four and each of these subplots received 30, 80, 130, and 190 kg nitrogen/ha. The number of plots was 72 and the plot size was 3 × 12 m. The wheat was drilled 23 September 2000 in well-cultivated soil with a seeding depth of 3 cm. Phosphorus (P), Potassium (K), and Sulfur (S) was supplied in adequate amounts according to the general nutrient status of the field as determined by soil samples: 20 P/ha, 60 kg K/ha, and 20 kg S/ha, respectively. The nitrogen was supplied as granulated NH_4NO_3 (Kemira Danmark, Denmark) in a split strategy with 30 kg N/ha supplied to all plots at 10 March 2001 and the rest of the nitrogen at 10 April 2001. The nitrogen was spread with a 1.5-m broad man-pulled plot spreader (Fiona Probe, Denmark).

2.2. Spectral measurements

Canopy reflectance was measured using a dual spectrometer SD2000 (Ocean Optics, USA). The measurement range was 400–900 nm, but due to low signal-to-noise ratio in both ends of the spectrum, the effective range was set to 438–883 nm and the resolution was 1 nm. The spectrometer was fitted with a 50- and a 100-μm optic fiber for simultaneous upward and downward detection of light. Pythagoras, Finland developed the Software 'NMES' controlling the spectrometers. A special feature of the system was dynamic integration time (DIT) for both channels. DIT of the spectrometer was automatically changed according to the light conditions, so that both channels always were delivering a sufficient amount of light to fill the CCD array between 75% and 95%. This gave an optimal use of the spectrometer and the highest signal-to-noise ratio.

The sensor heads were placed approximately 2-m above ground on a tractor-mounted boom at nadir position. The system was fitted with a 20.8° field-of-view optics. The measurements were performed between 1100 and 1400 local time (GMT + 1) at 7 May, 14 May, 22 May, 29 May, and 5 June, which in the experimental year corresponded to BBHC growth stage 30, 32, 42, 48, and 51, respectively. The solar angle compared to nadir was for all measurements less than 45° and no disturbing clouds were observed. The crop reflectance was calculated as the ratio between energy reflected by the crop and energy incident on the crop (solar irradiation). Thus, all reflectance data were corrected for noise related to differences between the two channels. This correction was made by the use of an average channel1/channel2 ratio spectrum using measurements of a white Spectralon ($BaSO_4$) reflectance panel (Labsphere, USA) in clear sunshine. The same reference spectrum made by the average of 10 repeated scans for each channel was used for all measurements.

2.3. Plant sampling and harvest procedure

Green plant sampling was performed five times during the vegetative growth stages from early stem elongation until heading. The actual dates were 7 May, 14 May, 22 May, 29 May, and 5 June 2001. The developmental stage of the wheat crop corresponded to BBHC growth stage 30, 32, 42, 48, and 51, respectively (Lancashire et al., 1991). An area of 0.375 m^2 was cut with a pair of scissors just above ground and brought to the lab for weight determination and calculation of the amount of green biomass (GBM). A representative sample was taken and the weight was determined. Leaves and stems were separated by excision at the leaf base. The surface area of the leaf sample was determined by the use of a leaf area meter (LICOR 3100, USA). Based on these measurements, leaf area index (LAI, m^2/m^2) was calculated. The leaf and stem samples were dried at 80 °C for 36 h and their dry weight subsequently measured.

Two sub-samples consisting of 10 leaves each were randomly selected among the youngest fully developed leaves. One sub-sample was dried at 80 °C for 36 h, ground and analyzed for the concentration of total nitrogen (N_{conc}; % nitrogen of dry weight) using an EA 1110 Automatic Elemental Analyzer (CE Elantech, USA). The other sub-sample was stored at −80 °C for less than 2 months and analyzed for total chlorophyll concentration (Chl_{conc}; mg chlorophyll g^{-1} fresh weight) according to Lichtenthaler (1987). The calculations of leaf chlorophyll density ($Chl_{density}$; mg chlorophyll m^{-2} soil), and leaf nitrogen density ($N_{density}$; g nitrogen m^{-2} soil) were based on the assumption that all green leaves from a sample contained the same amount of chlorophyll or nitrogen, respectively, as the selected leaves used for the chemical analysis. So, $Chl_{density}$ was calculated on the basis of Chl_{conc} and the amount of fresh leaves per m^2 (g leaves m^{-2}) and $N_{density}$ was calculated on the basis on N_{conc} and the amount of dry leaves per m^2.

2.4. Data pretreatment

Four individual dual scans in each plot were recorded. The four replicates were evaluated for outlier tendencies by the method described in Hansen, Jørgensen, and Thomsen, 2003. One of the replicates showed outlier tendency in 98 cases, corresponding to 6.8% of all scans performed, and was excluded from further calculations. The remaining replicates were averaged and used in the analysis.

Table 1
Selected properties of the investigated crop variables

Crop variables	Unit	Mean	S.D.†	Min	Max	Range‡
GBM	g/m^2	1165	751	118	4124	4006
LAI	cm^2/cm^2	1.69	0.79	0.005	4.46	4.46
Chl_{conc}	mg/g	2.05	0.50	0.94	5.97	5.02
$Chl_{density}$	mg/m^2	1058	548	223	3272	3049
N_{conc}[a]	%	3.45	0.70	1.88	5.48	3.60
$N_{density}$	g/m^2	3.48	1.44	0.54	8.49	7.94

[a] % of dry matter.
† Standard deviation ($n = 360$).
‡ Difference between maximum (max) and minimum (min).

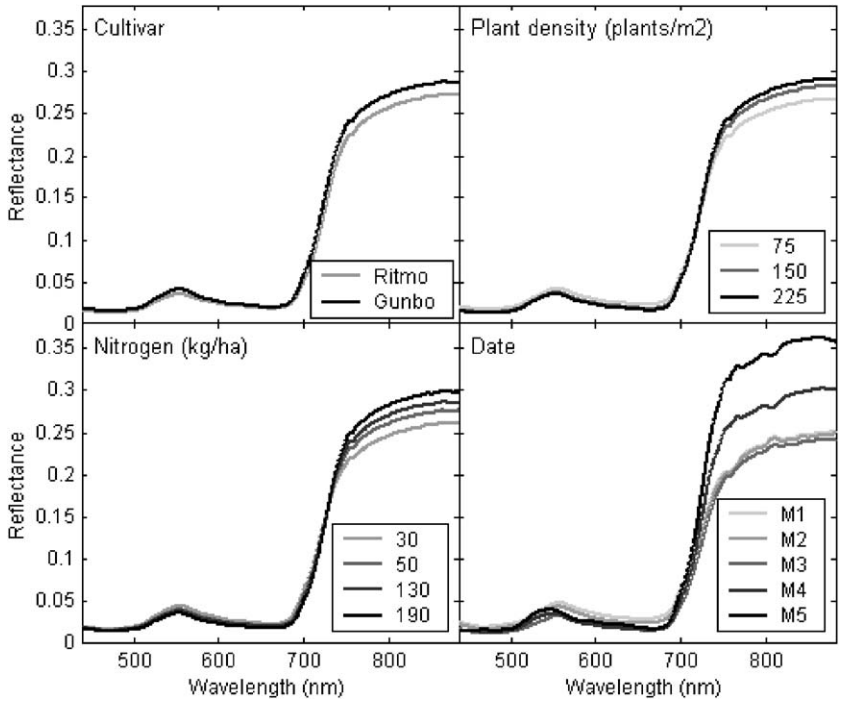

Fig. 1. Average reflectance spectrum of the different experimental treatments; cultivar ($n = 180$), plant density ($n = 120$), nitrogen application ($n = 90$), and date of measurement ($n = 72$).

2.5. Crop variable specific waveband selection

The basis of crop variable specific waveband selection was the two-band normalized difference vegetation index (Eq. (1)):

$$\mathrm{NDVI} = \frac{\lambda 2 - \lambda 1}{\lambda 2 + \lambda 1} \quad (1)$$

All possible two-pair combinations $\lambda 2 > \lambda 1$ of 441 wavelengths (97,020 combinations) were used in Eq. (1) and a linear regression was performed in order to determine the correlation coefficient (R^2). Linear regression was preferred in the initial analysis in order to operate with linear relationships. The procedure could not run with exponential fitting due to the requirement of initial parameter values. All the R^2 values were plotted in a matrix plot and the plot revealed a characteristic pattern with a number of "hot spots" with relatively high correlation coefficients. These spots were selected by choosing the wavelength combinations that showed an R^2 between 95% and 100% of the overall best index. The center wavelength and bandwidth

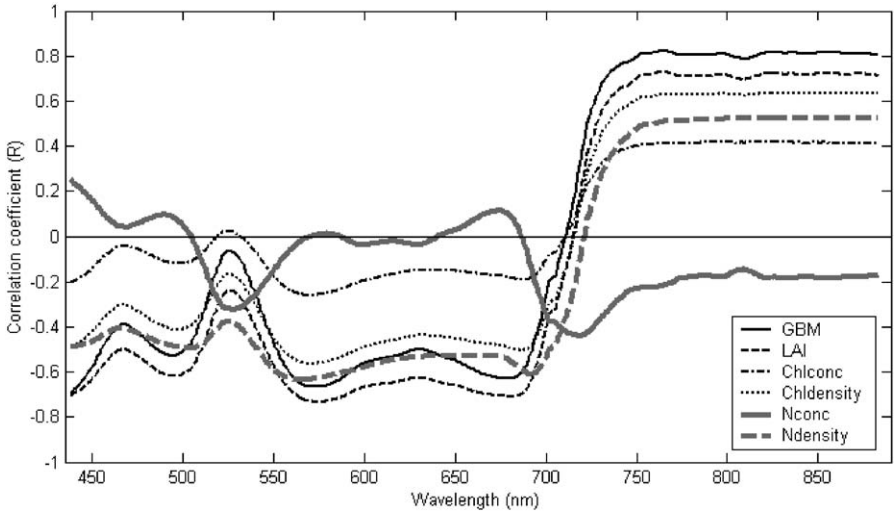

Fig. 2. Wavelength dependence of the coefficient of determination (R^2) derived from a linear regression analysis of canopy reflectance against GBM, LAI, leaf chlorophyll concentration ($\mathrm{Chl_{conc}}$), leaf chlorophyll density ($\mathrm{Chl_{density}}$), leaf nitrogen concentration ($\mathrm{N_{conc}}$), and leaf nitrogen density ($\mathrm{N_{density}}$), respectively.

for each of the selected spots were determined by fitting a rectangle that could hold the spot of interest inside its limits. Both linear ($y = ax + b$) and exponential ($y = ae^{bx}$) fitting procedures were tested using these new indices and their performance compared.

2.6. Partial least squares regression

PLS is a bilinear calibration method using data compression by reducing the large number of measured collinear spectral variables to a few non-correlated principal components (PCs). The PCs represents the relevant structural information, which is present in the reflectance measurements to predict the dependent variable. The spectral data were mean-centered before analysis and the optimal number of principal components was determined by the guidelines described in Esbensen (2000). The basic PLS algorithm will not be described in this paper, but further information on the PLS model can be obtained in Ehsani, Upadhyaya, Slaughter, Shafii, and Pelletier (1999). However, in principle, PLS regression uses component projection successively to find latent structures. Visual inspection of score-plots and validation residual variance plots was used to find the optimal number of PCs, so that over-fitting was prevented. In most cases, this procedure can reduce the number of spectral variables to a few

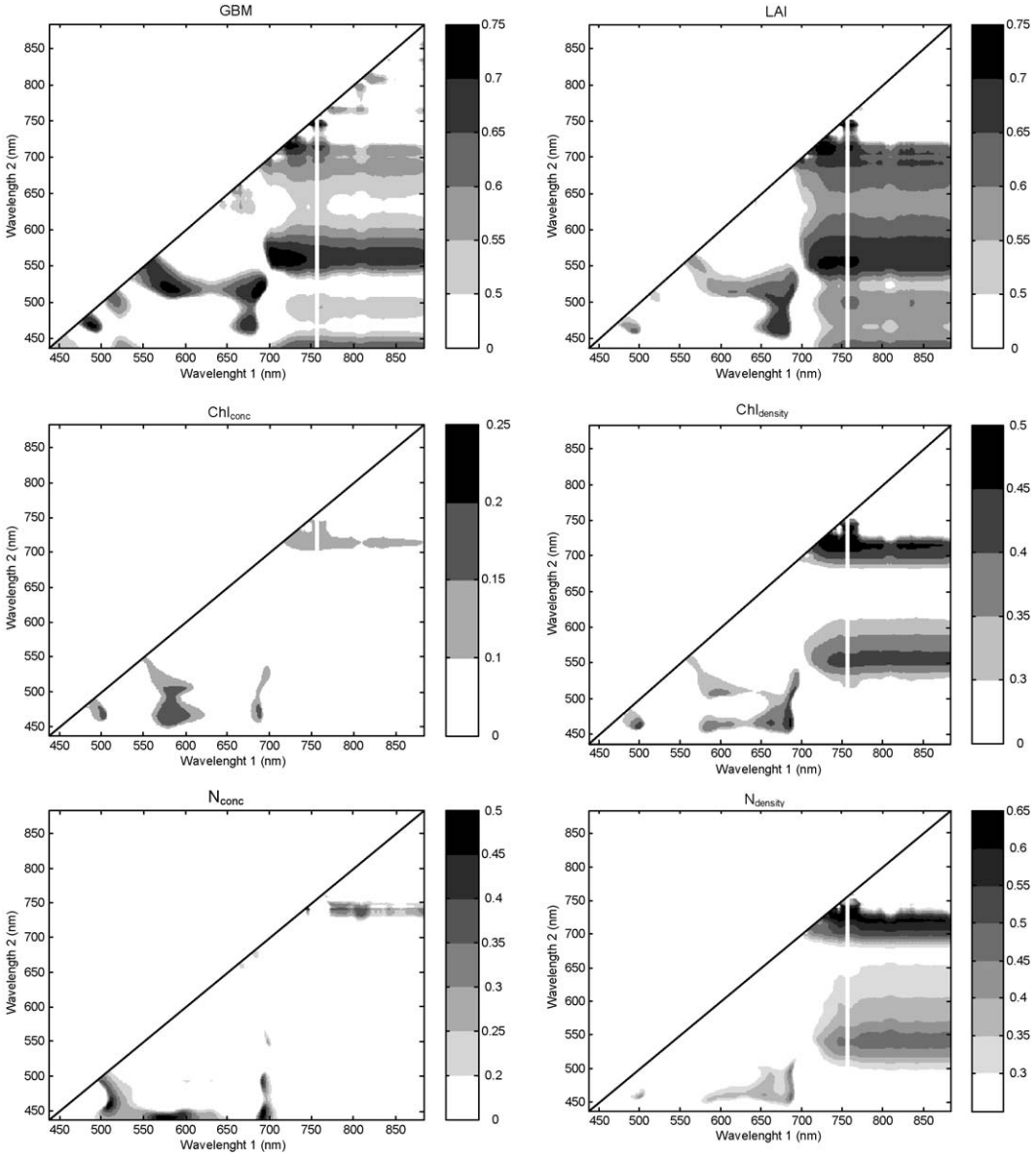

Fig. 3. Coefficient of determination (R^2) for the relation of all combinations of wavelengths used for a linear regression analysis of the normalized difference vegetation index=$(\lambda 1 - \lambda 2)/(\lambda 1 + \lambda 2)$ against GBM, LAI, leaf chlorophyll concentration (Chl_{conc}), leaf chlorophyll density ($Chl_{density}$), leaf nitrogen concentration (N_{conc}), and leaf nitrogen density ($N_{density}$), respectively. A total number of 97,020 combinations were investigated. Only the results from the matrix below the diagonal are shown, as the R^2 values were symmetrical.

independent PCs. The final model predicting $\hat{y}_i$ had the following form (Eq. (2)):

$$\hat{y}_i = b_0 + b_1 t_{1i} + b_2 t_{2i} + \cdots + b_n t_{ni} \quad (2)$$

where t_{1i} to t_{ni} were the scores from principal component (PC) 1 to n for variable i. The scores were calculated on the basis of mean-centered data. By linear regression of t versus y in the calibration iteration process, the regression coefficient b_n was obtained. Due to the initial centering of y, the centered mean b_0 was added in order to obtain $\hat{y}_i$.

Validation of the models was performed by comparing differences in R^2 and root mean square error (RMSE). RMSE values were calculated according to Eq. (3):

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (3)$$

where $\hat{y}_i$ and y_i were the predicted and measured crop variables, respectively, and n the number of samples ($n = 360$). RMSE provides a direct estimate of the modeling error expressed in original measurement units (Kvalheim, 1987).

All data handling and regression calculations were made in Matlab 6.0 including the Optimization Toolbox (Mathworks, USA). The PLS analysis was performed using the PLS toolbox 2.0 (Eigenvector, USA).

3. Results

The experimental treatments, including different cultivars, plant densities, and nitrogen application strategies together with the temporal timing of plant sampling, caused a wide range of variation within the investigated crop variables, i.e. 35-fold variation in GBM, almost 900-fold variation in LAI, 6- to 15-fold variation in leaf chlorophyll and 3- to 16-fold variation in leaf nitrogen (Table 1).

The wide range in the investigated crop variables (Table 1) was planned in order to make the relationship between plant performance and reflectance measurements (Fig. 1) as realistic and universal as possible. The reflectance was at all wavelengths on average higher in Gunbo compared to Ritmo (Fig. 1a). Gunbo covered even from early growth stages a higher proportion of the soil surface compared to Ritmo despite a slightly lower GBM (data not shown). Increasing seeding rate and nitrogen supply caused on average lower reflectance in the visible spectral range (438–690 nm), while the reflectance was higher in the near infrared spectral range [750–883 nm, Fig. 1b and c]. The canopy reflection in the near infrared spectral range in-

Table 2
The best NDVI indices from the top 5% of all combinations of band centers $\lambda 1$ and $\lambda 2$ sorted in decreasing order with respect to the single highest coefficient of determination (R^2) for a given waveband combination observed in each spot

Crop variable	Band center and width (nm)	Band centers ($\lambda 1$ and $\lambda 2$) and band widths ($\Delta\lambda 1$ and $\Delta\lambda 2$) for two-band vegetation indices					
		Index 1 (nm)	Index 2 (nm)	Index 3 (nm)	Index 4 (nm)	Index 5 (nm)	Index 6 (nm)
GBM (g/m^2)	$\lambda 1$	521	565	718	520	469	748
	$\Delta\lambda 1$	23	23	10	8	3	2
	$\lambda 2$	689	708	720	581	488	750
	$\Delta\lambda 2$	14	36	14	12	4	3
LAI (cm^2/cm^2)	$\lambda 1$	711	746	556	556	–	–
	$\Delta\lambda 1$	20	6	14	11	–	–
	$\lambda 2$	720	750	730	760	–	–
	$\Delta\lambda 2$	25	12	23	15	–	–
Chl_{conc} (mg/g)	$\lambda 1$	471	501	–	–	–	–
	$\Delta\lambda 1$	24	4	–	–	–	–
	$\lambda 2$	584	586	–	–	–	–
	$\Delta\lambda 2$	20	7	–	–	–	–
$Chl_{density}$ (mg/m^2)	$\lambda 1$	717	–	–	–	–	–
	$\Delta\lambda 1$	19	–	–	–	–	–
	$\lambda 2$	732	–	–	–	–	–
	$\Delta\lambda 2$	22	–	–	–	–	–
N_{conc} (%)	$\lambda 1$	440	447	459	–	–	–
	$\Delta\lambda 1$	5	9	5	–	–	–
	$\lambda 2$	573	692	509	–	–	–
	$\Delta\lambda 2$	14	3	4	–	–	–
$N_{density}$ (g/m^2)	$\lambda 1$	734	730	717	720	–	–
	$\Delta\lambda 1$	24	20	8	8	–	–
	$\lambda 2$	750	759	770	839	–	–
	$\Delta\lambda 2$	22	11	39	46	–	–

The symbols $\Delta\lambda 1$ and $\Delta\lambda 2$ denote the bandwidths to the matching center wavelength.

creased with growth stage. The lack of canopy development between measurement occasions M2 and M3 (Fig. 1d) was due to a short period of mild drought.

A correlogram was constructed by sequential regression of the reflectance at each of the 441 individual narrow-bands against one crop variable at a time and subsequent plotting of the correlation coefficient (R) against wavelength (Fig. 2). The correlograms for the area related variables (GBM, LAI, $Chl_{density}$, and $N_{density}$) showed three broad regions of relatively high correlation with negative coefficients in the blue and red region and positive in the near infrared region. Chl_{conc} and N_{conc} behaved oppositely to the area-related crop variables, with low correlation coefficients in the three broad regions blue, red, and near infrared. Two narrow regions positioned at 530 nm (green peak) and 710 nm (red edge) showed a low correlation coefficient for all the investigated variables.

A number of "hot spots" (Fig. 3) with high correlation coefficients were revealed in a linear regression analysis of the individual crop variables against NDVIs calculated according to Eq. (1) for all possible two-pair combinations of the reflectance measured at the 441 wavelengths. Analysis of the center wavelength and bandwidth in both directions revealed from one to six indices depending on crop variable (Table 2).

Bands with center wavelengths in the red edge spectral region from 680 to 750 nm were represented in 50% of all selected bands for all crop variables in total. The bands ($\lambda 1$ and $\lambda 2$) were often paired so that both bands were closely spaced in the steep linear shift between R_{red} and R_{nir} reflection (Fig. 1 and Table 2). This was true for narrow-band combinations referring to GBM, LAI, $Chl_{density}$, and $N_{density}$, which all were expressed on the basis of ground surface area or canopy surface area. Another effective band combination estimating GBM and LAI was provided by $\lambda 2$ at 690–760 nm combined with $\lambda 1$ in the green area 521–565 nm ($\lambda 1$). For $N_{density}$, a relatively broadband ($\lambda 2$) in the NIR region at 770 and 839 nm together with a narrow band ($\lambda 1$) at the red shift (717 and 720 nm) was efficient (indices 3 and 4 in Table 2). The red edge region was almost absent for bands connected to Chl_{conc} and N_{conc}. The bands for these two variables were exclusively found in the VIS spectral range, mainly in the blue region ($\lambda 1$) paired with a green or a red band ($\lambda 2$).

Table 3
Coefficients of determination (R^2) between various NDVIs and specific crop variables

Band type	Index	MSS number/center wavelength		GBM	Coefficient of determination (R^2)				
		$\lambda 1$	$\lambda 2$		LAI	Chl_{conc}	$Chl_{density}$	N_{conc}	$N_{density}$
Best NDVI from Table 2									
Linear fit	Index 1			0.72	0.72	**0.24**	**0.53**	**0.55**	**0.68**
	Index 2			0.74	**0.73**	0.23	–	**0.55**	0.67
	Index 3			**0.75**	0.71	–	–	0.54	0.64
	Index 4			0.72	0.71	–	–	–	0.66
	Index 5			0.72	–	–	–	–	–
	Index 6			0.72	–	–	–	–	–
Best NDVI from Table 2									
Exponential fit	Index 1			0.78	**0.75**	**0.24**	**0.57**	**0.56**	0.68
	Index 2			**0.84**	**0.75**	**0.24**	–	0.55	**0.69**
	Index 3			0.77	**0.75**	–	–	0.55	**0.69**
	Index 4			0.76	**0.75**	–	–	–	**0.69**
	Index 5			0.75	–	–	–	–	–
	Index 6			0.75	–	–	–	–	–
Broadbands	RVI	4	3	0.64	0.71	0.13	0.49	0.00	0.52
	$VI1_{broad}$	4	3	0.53	0.61	0.07	0.33	0.01	0.36
	$VI2_{broad}$	4	2	**0.61**	**0.67**	0.11	**0.43**	0.00	**0.48**
	$VI3_{broad}$	4	1	0.51	0.58	0.06	0.32	0.01	0.36
	$VI4_{broad}$	2	3	0.37	0.48	0.02	0.20	0.04	0.19
	$VI5_{broad}$	1	2	0.01	0.00	0.09	0.05	**0.13**	0.05
	$VI6_{broad}$	1	3	0.57	0.66	**0.13**	**0.43**	0.00	0.42
Short-bands	RVI	790	680	0.69	0.73	0.14	0.50	0.00	0.51
	$VI1_{short}$	790	680	0.56	0.62	0.07	0.34	0.01	0.36
	$VI2_{short}$	790	550	**0.64**	**0.68**	**0.13**	**0.46**	0.00	**0.51**
	$VI3_{short}$	790	450	0.57	0.59	0.07	0.32	0.04	0.32
	$VI4_{short}$	550	680	0.46	0.53	0.04	0.23	0.06	0.21
	$VI5_{short}$	450	550	0.23	0.18	0.00	0.03	**0.32**	0.00
	$VI6_{short}$	450	680	0.31	0.49	0.08	0.33	0.04	0.44

The upper half of the table shows R^2 values obtained by linear or exponential fitting to the best indices in Table 2. The lower half of the table shows R^2 values obtained by exponential fitting of NDVIs to different combinations of blue, green, red and near infrared broadbands (TM MSS; $VI1_{broad}$–$VI6_{broad}$) and short-bands ($VI1_{short}$–$VI6_{short}$). The simple ratio red and near infrared (RVI) was included for comparison. The "bold" font indicates the best index in each section of the table.

Table 3 shows the coefficients of determination based on a linear regression analysis of crop variable against each of the NDVI indices calculated on the basis of the band centers in Table 2. The differences in R^2 values inside each of the six crop variables were very small (4–6%), while among crop variables there was a considerable variation. The two R^2 extremes were Chl_{conc} and GBM with 0.24 and 0.75, respectively. However, most of the observed relationships were not linear, but rather followed an exponential interdependence (Fig. 4). A total test of all two-band combinations and their exponential fit were not possible due to limitations in computer power. However, if the selected wavelengths and wavebands in the linear fit were regarded as optimal also in the exponential fit, some of the crop variables could be more precisely described using an exponential relationship (Table 3). This was, e.g. the case for GBM, LAI, $Chl_{density}$, and $N_{density}$, where R^2 increased 12%, 5%, 7%, and 7%, respectively, compared to the linear regression. R^2 values for Chl_{conc} or N_{conc} did not improve by exponential fitting.

The relationship between broadband and short-band NDVIs for the crop physiological variables was investigated and the performance compared to the selected narrow-band NDVI models (Table 3). The best coefficient of determination (R^2) using linear regression was obtained for estimates of GBM and LAI. The band combination used in both broadband and short-band NDVI was near infrared and green (GNDVI). Furthermore, it was in all cases observed that short-band combinations performed equally well or better than broadband NDVIs with 0–60% improvement. The R^2 values obtained using narrow-band indices with linear fit were 7–46% higher than those obtained using the best short-band indices.

The performance of the best of the selected indices was compared to that obtained in a multivariate calibration based on bilinear partial least squares regression (PLS). The number of components in the final model ranged from 2 in LAI to 6 in GBM and N_{conc} models. PLS increased all R^2 values with 0.02 to 0.16 compared to exponential fitting. The most significant improvement was observed for N_{conc} where R^2 increased from 0.56 to 0.71. Evaluated on the basis of the root mean square error (RMSE), which represents the average error, PLS improved the models related to GBM with 22%, $Chl_{density}$ 4%, N_{conc} 24%, and $N_{density}$ 3% as compared to exponential fitting used the selected narrow-bands (Table 4). No improvements were observed using PLS for LAI and Chl_{conc}.

The different components can be defined by their respective scores and loadings. The scores are related to the single samples, while the loadings quantify the contribution of different wavelengths to the model. The loading weights (LWs) are loadings orientated, allowing the optimal fit to be achieved for the specific crop variable of interest. The LWs related to the six investigated variables are shown in Fig. 5.

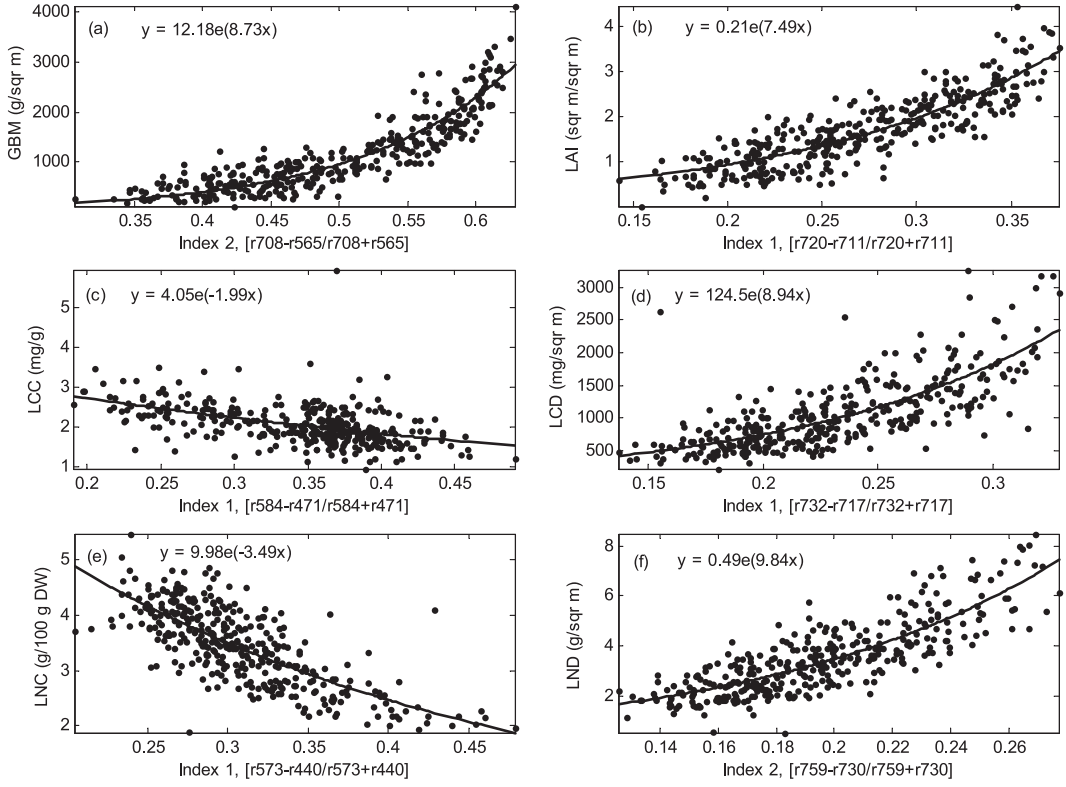

Fig. 4. The exponential relationship between the best narrow-band selected indices and the six investigated crop variables: (a) GBM, (b) LAI, (c) leaf chlorophyll concentration, (d) leaf chlorophyll density, (e) leaf nitrogen content, and (f) leaf nitrogen density. Index 1 and Index 2 refer to the indexes given in Table 3.

Table 4
Results of the partial least squares regression (PLS) modeling the investigated crop biophysical variables; GBM, LAI, leaf chlorophyll concentration, leaf chlorophyll density, leaf nitrogen concentration, and leaf nitrogen density

	PLS			The best narrow-band NDVI linear fit		The best narrow-band NDVI exponential fit	
	PCs	R^2	RMSE	RMSE	Percent difference to PLS	RMSE	Percent difference to PLS
GBM	6	0.89	250	387	+55	304	+22
LAI	2	0.75	0.40	0.42	+6	0.39	−2
Chl_{conc}	4	0.30	0.42	0.44	+5	0.44	0
$Chl_{density}$	3	0.60	345	373	+8	359	+4
N_{conc}	6	0.71	0.38	0.47	+24	0.47	+24
$N_{density}$	3	0.71	0.78	0.83	+6	0.80	+3

The results of R^2 and RMSE were compared to the best narrow-band NDVI with linear and exponential fit, respectively.

The first three LWs showed some common trends, reflecting the relationship between the spectral data and the canopy. Two zones at approximately 550 and 700–740 nm of major importance for the PLS models could be identified. These zones showed either a shift or a significant peak.

There was a high degree of coincidence between the selected narrow bands for the best NDVI index and the size of the numerical PLS regression coefficients (comparison of Table 2 and Fig. 5). This means that the same wavelengths were important in both methods. A clear example was $N_{density}$, where the best linear fit used $\lambda 1 = 734$ nm and $\lambda 2 = 750$ nm as center wavelengths (index 1 in Table 2). This was also exactly the area where the PLS regression coefficient shifted from a global low to a global high value. The rest of the selected indices (indices 2 to 4) were located in the NIR region, where most of the structural information also was located (Fig. 5). On the contrary, there was little coincidence when the difference in RMSE was high between analysis methods, as was the case for GBM and N_{conc}.

4. Discussion

The overall level of reflectance was directly influenced by the different treatments (Fig. 1). However, the pattern of

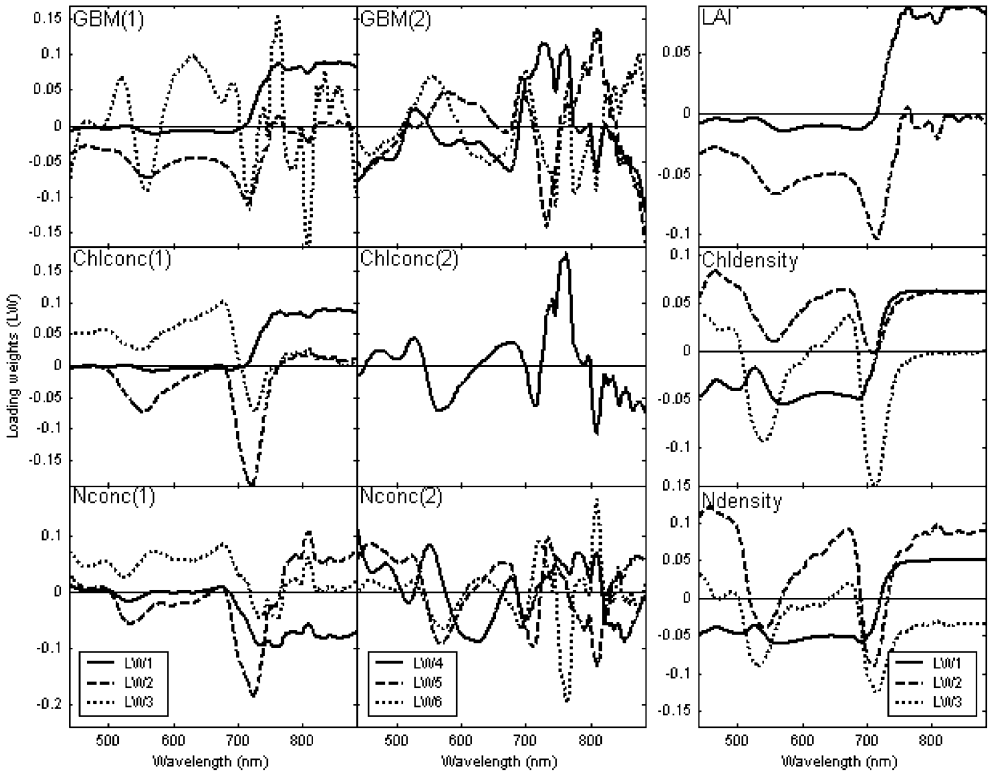

Fig. 5. Loading weights (LW) for the PLS regression models calibrated to estimate green biomass (GBM), leaf area index (LAI), leaf chlorophyll concentration (Chl_{conc}), leaf chlorophyll density ($Chl_{density}$), leaf nitrogen concentration (N_{conc}), and leaf nitrogen density ($N_{density}$). Subplot columns 1 and 2 show the relevant LWs from 1 to 3 and from 4 to 6, respectively, for GBM, Chl_{conc} and N_{conc}. High numerical values of regression coefficients indicate high importance of the reflected wavelength in the PLS analysis.

change was in general the same for all the treatments. Increased plant density, nitrogen supply and growth stage caused increased greenness, reflecting a combination of biomass and chlorophyll density. The effect on the canopy reflectance followed the general assumption that an increase in greenness create lower R_{vis}, because of increased light absorption due to a higher density of pigments per unit soil surface area (Hinzman, Bauer, & Daughtry, 1986). The opposite effect was observed for R_{nir}. The relative high reflectance observed is due to the canopy and leaf internal scattering (Colwell, 1974).

A large number of two-band reflectance combinations were tested and used in the NDVI equation (Eq. (1)). The traditional two-band combination is R_{red} and R_{nir}, which has been used in satellite, aerial, and ground-based applications and found to be sensitive to different crop variables (Broge & Mortensen, 2002; Carlson & Ripley, 1997; Casanova, Epema, & Goudriaan, 1998; Colwell, 1974). A number of new two-band combinations designed to be optimally related to a specific crop variable have also been tested (Blackburn, 1998b; Gitelson et al., 1996).

Broad- (>50 nm) and short-bands (5–10 nm) have been used in various vegetation indices for comparative studies (Broge & Leblanc, 2001; Thenkabail et al., 2001). In the present work, six different two-band combinations of blue, green, red and NIR were compared in order to evaluate the sensitivity of broad- and short bands used in NDVI (Table 3). The broadband data were analogous to the bandwidths of the Landsat Thematic Mapper using the TM-spectrometer, while the short-bands were set to be 10 nm wide with center wavelengths at 450, 550, 680, and 790 nm. A slightly improved performance of short-bands compared to broadbands was observed for all the investigated crop variables, which is in good agreement with comparative studies of measured and simulated data (Broge & Leblanc, 2001; Elvidge & Chen, 1995; Thenkabail et al., 2001).

The band selection has been carried out in relation to known canopy characteristics (Chappelle, Kim, & Mcmurtrey, 1992). The approach in the present experiment was different, because all available combinations of wavelengths were tested in a NDVI. The knowledge of the spectral properties of the canopy was not used directly in this analysis. Instead, a matrix plot was created, which is a sort of two-dimensional correlogram (Fig. 3). The same kind of matrix plots were used to find the best wavebands using the simple ratio pigment index (SRPI), normalized difference pigments index (NDPI) and structure insensitive pigment index (SIPI) estimating Chl*a*, Chl*b* and the ratio *Cars*/Chl*a* in leaves of from a number of species, among them wheat, grown in greenhouse (Penuelas, Baret, & Filella, 1995). They were also used in ground-based hyperspectral reflectance measurements estimating biomass and LAI in cotton, potato, soybeans, corn, and sunflower in Syria (Thenkabail et al., 2001). The advantages of the matrix plots are that they give a quick overview of thousands of wavelength combinations and make it possible to detect wavelengths of interest for further analysis.

The selected bands with center wavelengths in the red edge spectral region from 680 to 750 nm were represented in 87% of the selected indices related to GBM, LAI, $Chl_{density}$, and $N_{density}$. The band combinations were often paired so that both bands were closely spaced in the steep linear shift between R_{red} and R_{nir}. This index proved in fact to give the best description of the four crop variables GBM, LAI, $Chl_{density}$, and $N_{density}$, all expressing quantity per unit soil surface or canopy area. Use of a such normalized difference type of index implies that closely spaced narrow bands indicate the slope of the reflectance, which often is referred to as a derivative vegetation index (Elvidge & Chen, 1995). Red edge has proved to respond more linearly to LAI index and chlorophyll compared to classical NDVI, which often suffer from saturation problems even at relatively low LAI values (<3.0; Danson & Plummer, 1995).

The red edge region was almost absent for bands related to Chl_{conc} and N_{conc}. The bands for these variables were exclusively found in the VIS spectral range, mainly in the blue region characterized by a strong light absorption due to chlorophylls *a* and *b*. The blue band from 440 to 501 nm ($\lambda 2$) was paired with bands in the green [573–586 nm, ($\lambda 1$)] or a band in the red region [692 nm, ($\lambda 1$)]. It is not obvious how the selected band combinations used in indices related to Chl_{conc} and N_{conc} were connected to the results from the correlogram. However, the chlorophyll concentration, and to some extent other pigments in the leaves, plays a major role in coloring the canopy during vegetative growth stages. This agrees with the fact that visible wavelengths seem to be important for indices related to crop variables expressing a concentration of either chlorophyll or leaf nitrogen.

Problems due to saturation have previously been reported for NDVI and soil adjusted vegetation index (SAVI), using red and NIR, while it was less pronounced for green NDVI (Daughtry, Walthall, Kim, de Colstoun, & Mcmurtrey, 2000; Gitelson et al., 1996). In the present work, saturation was not critical because measurements were made relatively early in the growing season (last measurement occasion was just before heading), i.e. over a period with an average LAI of 1.7 and a maximum LAI of 4.5 (Table 1). Problems with saturation are observed for LAI above 2.5. The coefficient of determination (R^2) between selected indices and concentration of chlorophyll (Chl_{conc}) and nitrogen (N_{conc}) was not improved by exponential fitting. These crop variables are supposed to be related to the color of the canopy. The amount of biomass is therefore of less importance at least when LAI exceeds 2.5. The consequence is that saturation problems related to the amount of biomass have a minor effect, resulting in a more linear relation.

Multivariate calibration methods have until now only been used in very few cases within remote sensing in agriculture. The present work represents an initial step in evaluating bilinear partial least squares regression (PLS) compared to vegetation indices. The PLS models performed

equally or better (up to 23%) compared to the best of the selected indices based on exponential curve fitting, showing that PLS indeed is a potentially useful method. It is very critical to select the correct number of components in order to avoid over-fitting caused by too many components. In the present work, the highest number of components was 6 for GBM and N_{conc}. These two variables were at the same time those for which the highest improvement was observed. Thus, selection of two narrow bands used for a NDVI leaves out valuable information which could contribute to a better explanation of the relation between spectral reflectance data and specific crop variables, while this information is still included in the PLS models. Consequently, PLS provides a useful explorative tool for interpretation of the relationship between canopy spectral reflectance and crop variables. Furthermore, PLS seems to be a good alternative to univariate statistical methods.

5. Conclusions

NDVI is a potentially useful vegetation index. However, selection of the correct wavelengths and bandwidths are important. Short bands (10 nm) perform better than broadbands (TM-bands) using standard Red/NIR and Green/NIR NDVIs. Selection of the optimal waveband combination in NDVI improves the relation to the investigated crop variables as compared to the performance on standard short band NDVI using a linear regression method. More variation is explained if the selected indices are used in an exponential relationship to the crop variables. PLS can further improve this relation especially for GBM and N_{conc}.

Acknowledgements

The work was supported by the Danish Ministry of Food, Agriculture and Fisheries and Risoe National Laboratory, Plant Research Department. We thank Rasmus N. Jørgensen, Risoe National Laboratory for his inspiring comments and suggestions during the preparation of this paper. The technical assistance of Mrs. L. Brandt and Mr. F. Hasselbalch is gratefully acknowledged.

References

Ahlrichs, J. S., & Bauer, M. E. (1983). Relation of agronomic and multispectral reflectance characteristics of spring wheat canopies. *Agronomy Journal*, *75*, 987–993.

Aparicio, N., Villegas, D., Casadesus, J., Araus, J. L., & Royo, C. (2000). Spectral vegetation indices as non-destructive tools for determining durum wheat yield. *Agronomy Journal*, *92*, 83–91.

Asseng, S., van Keulen, H., & Stol, W. (2000). Performance and application of the APSIM Nwheat model in the Netherlands. *European Journal of Agronomy*, *12*, 37–54.

Best, R. G., & Harlan, J. C. (1985). Spectral estimation of green leaf area index of oats. *Remote Sensing of Environment*, *17*, 27–36.

Blackburn, G. A. (1998a). Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches. *Remote Sensing of Environment*, *66*, 273–285.

Blackburn, G. A. (1998b). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing*, *19*, 657–675.

Broge, N. H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, *76*, 156–172.

Broge, N. H., & Mortensen, J. V. (2002). Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. *Remote Sensing of Environment*, *81*, 45–57.

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, *62*, 241–252.

Casanova, D., Epema, G. F., & Goudriaan, J. (1998). Monitoring rice reflectance at field level for estimating biomass and LAI. *Field Crops Research*, *55*, 83–92.

Chappelle, E. W., Kim, M. S., & Mcmurtrey, J. E. (1992). Ratio analysis of reflectance spectra (RARS). An algorithm for the remote estimation of the concentrations of chlorophyll *A*, chlorophyll *B*, and carotenoids in Soybean Leaves. *Remote Sensing of Environment*, *39*, 239–247.

Christensen, S., & Goudriaan, J. (1993). Deriving light interception and biomass from spectral reflectance ratio. *Remote Sensing of Environment*, *43*, 87–95.

Colwell, J. E. (1974). Vegetation canopy reflectance. *Remote Sensing of Environment*, *3*, 175–183.

Danson, F. M., & Plummer, S. E. (1995). Red-edge response to forest leaf area index. *International Journal of Remote Sensing*, *16*, 183–188.

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

Ehsani, M. R., Upadhyaya, S. K., Slaughter, D., Shafii, S., & Pelletier, M. (1999). A NIR technique for rapid determination of soil mineral nitrogen. *Precision Agriculture*, *1*, 217–234.

Elvidge, C. D., & Chen, Z. K. (1995). Comparison of broad-band and narrow-band red and near-infrared vegetation indexes. *Remote Sensing of Environment*, *54*, 38–48.

Esbensen, K. H. (2000). *Multivariate data analysis—in practice*. Corvallis: CAMO, 598 pp.

Fernández, S., Vidal, D., Simón, E., & Solé-Sugrañes, L. (1994). Radiometric characteristics of *Triticum aestivum* cv. Astral under water and nitrogen stress. *International Journal of Remote Sensing*, *15*, 1867–1884.

Filella, I., Serrano, L., Serra, J., & Penuelas, J. (1995). Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Science*, *35*, 1400–1405.

Gamon, J. A., Penuelas, J., & Field, C. B. (1992). A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment*, *41*, 35–44.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*, 289–298.

Gitelson, A. A., & Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, *18*, 2691–2697.

Gitelson, A. A., Merzlyak, M. N., & Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology*, *148*, 501–508.

Hansen, P. M., Jørgensen, J. R., & Thomsen, A. (2003). Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. *Journal of Agricultural Science, Cambridge*, *139*, 307–318.

Hinzman, L. D., Bauer, M. E., & Daughtry, C. S. T. (1986). Effects of

nitrogen fertilization on growth and reflectance characteristics of winter wheat. *Remote Sensing of Environment*, *19*, 47–61.

Jamieson, P. D., Porter, J. R., Goudrian, J., Ritchie, J. T., van Keulen, H., & Stol, W. (1998). A comparison of the models AFRCWHEAT2, CERES-wheat, Sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. *Field Crops Research*, *55*, 23–44.

Kvalheim, O. M. (1987). Latent-structure decompositions (projections) of multivariate data. *Chemometrics and Intelligent Laboratory Systems*, *2*, 283–290.

Lancashire, P. D., Bleiholder, H., Langelüddecke, P., Stauss, R., Van den Boom, T., Weber, E., & Witzen-Berger, A. (1991). An uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, *119*, 561–601.

Lichtenthaler, H. K. (1987). Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes. *Methods in Enzymology*, *148*, 350–382.

Lukina, E. V., Stone, M. L., & Raun, W. R. (1999). Estimating vegetation coverage in wheat using digital images. *Journal of Plant Nutrition*, *22*, 341–350.

Maccioni, A., Agat, G., & Mazzinghi, P. (2001). New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *Journal of Photochemistry and Photobiology. B, Biology*, *61*, 52–61.

Penuelas, J., Baret, F., & Filella, I. (1995). Semi-empirical indices to assess carotenoids/chlorophyll *a* ratio from leaf spectral reflectance. *Photosynthetica*, *31*, 221–230.

Penuelas, J., Camon, J. A., Griffin, K. L., & Field, C. B. (1993). Assessing community type, plant biomass, pigment composition, and photosynthetic efficiency of aquatic vegetation from spectral reflectance. *Remote Sensing of Environment*, *46*, 110–118.

Penuelas, J., Filella, I., & Gamon, J. A. (1995). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist*, *131*, 291–296.

Penuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., & Field, C. B. (1994). Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment*, *48*, 135–146.

Serrano, L., Filella, I., & Penuelas, J. (2000). Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Science*, *40*, 723–731.

Thenkabail, P. S., Smith, R. B., & de Pauw, E. (2001). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, *71*, 158–182.

White, J. D., Trotter, C. M., Brown, L. J., & Scott, N. (2000). Nitrogen concentration in New Zealand vegetation foliage derived from laboratory and field spectrometry. *International Journal of Remote Sensing*, *21*, 2525–2531.

Wiegand, C. L., Maas, S. J., Aase, J. K., Hatfield, J. L., Pinter, P. J., Jackson, R. D., Kanemasu, E. T., & Lapitan, R. L. (1992). Multisite analyses of spectral–biophysical data for wheat. *Remote Sensing of Environment*, *42*, 1–21.

Yoder, B. J., & Pettigrew-Crosby, R. E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, *53*, 199–211.

Yoder, B. J., & Waring, R. H. (1994). The normalized difference vegetation index of small douglas-fir canopies with varying chlorophyll concentrations. *Remote Sensing of Environment*, *49*, 81–91.