Australian Journal of Crop Science
AJCS 9(10):976-986 (2015) ISSN:1835-2707

Normalized difference spectral indices and partial least squares regression to assess the yield and yield components of peanut

Salah Elsayed[1*], Mohamed Elhoweity[1], Urs Schmidhalter[2]

[1]Evaluation of Natural Resources Department, Environmental Studies and Research Institute, Sadat City University, Egypt
[2]Chair of Plant Nutrition, Department of Plant Sciences, Technische Universität München, Emil-Ramann-Str. 2, D-85350 Freising, Germany

***Corresponding author: sala7emam@ yahoo.com**

Abstract

High-throughput hyperspectral passive reflectance sensing can acquire timely information to make more informed management decisions in real time compared with the more laborious destructive measurements. Early prediction of yield and yield components of peanut by spectral reflectance measurements prior to harvest could reduce the phenotyping time and expenses compared to destructive measurements. In this study, the performance of hyperspectral passive reflectance sensing was tested at three growth stages, the beginning of pod development and at 50% and 80% pod development, to assess their relationship to the pod yield, seed protein content, seed oil content, and straw yield of peanut cultivars Two peanut cultivars, Giza 5 and Giza 6, were grown under field conditions and subjected to three levels of nitrogen application. Simple linear regression and partial least squares regression (PLSR) models were compared to analyse the spectral data. The closest relationships were obtained for the spectral index (R_{610} - R_{424})/(R_{610} + R_{424}) with the pod yield (R^2 = 0.70, significant at $p \leq 0.001$), as well as the straw yield (R^2 = 0.53, significant at $p \leq 0.001$) and the protein content (R^2 = 0.69, significant at $p \leq 0.001$. For the relationships between PLSR with the pod yield, protein content and oil content and the straw yield of peanut cultivars, the coefficients of determination reached values up to R^2 = 0.82 (significant at $p \leq$ 0.001) through the individual measurements. Both, the PLSR and normalized difference spectral index analysis of spectral data performed better for assessing the pod yield and protein content than the oil content and straw yield of peanut cultivars. In conclusion, phenotyping yield and quality related parameters of peanut by PLSR analysis of non-invasive reflectance measurements represent a promising strategy for management action as well as for screening peanut cultivars.

Keywords: Phenomics; phenotyping; proximal sensing; spectral reflectance; protein content.
Abbreviations: BBCH 71, 85 and 89_ BBCH growth stages indicating beginning of pod development and 50 and 80% of pod development, respectively; fed_feddan: 1 feddan = 0.42 hectares = 1.038 acres. GNDVI_the green normalized difference vegetation index; NDVI_the normalized difference vegetation index; PLSR_partial least square regression; NWI-3_the normalised water index 3; HPS_hyperspectral passive sensor.

Introduction

Peanut or groundnut (*Arachis hypogaea* L.) is one of the most important leguminous crops which are an important source of protein and oil. The high-energy value, protein content, and minerals make peanut a rich source of nutrition at a comparatively low price. Peanut seeds contain high amounts of edible oil (43 - 50%), protein (25-30%), carbohydrate (20%) and fiber (5%) and ash which make a substantial contribution to human nutrition (Fageria et al., 1997; Aboukheira, 2009). Nitrogen is an essential component of many compounds of plants, such as chlorophyll, proteins, nucleotides, alkaloids, enzymes, hormones (Tiwari and Dhakar 1997). Nitrogen is the most important constituent of plant proteins and is required throughout crop growth from the vegetative stage to the subsequent harvesting. Application of nitrogen is known to mainly increase the protein content and protein fractions. Many researchers have found that late season top dressed nitrogen addition as dry fertilizer material was most effective in attaining higher grain protein concentration, yield and increased fertilizer recovery and efficiency (Weiser and Seilmeier, 1998; Michael et al., 2000; Pendashteh et al., 2011). Barik et al. (1998) found that increasing the level of nitrogen fertilizer increased pod and seed yield and 100 seed and pod weight. Plant phenotyping is currently one of the key limiting factors of the agricultural research and will play a vital role in ensuring yield stability and continuous increase (Furbank and Tester 2011). Recent progress in DNA marker assays and sequencing technologies enable high throughput genotyping of many individual plants at relatively low cost (Peleman and van der Voort, 2003). In contrast, phenotyping of a large number of genotypes and mapping populations in field trials is still laborious and expensive. Spectral approaches have been proposed to be efficient, high-throughput and cost-effective and to allow phenotyping several traits in field trials (Winterhalter et al., 2012; Walter et al., 2012), and could bring about tremendous progress in plant genetic screening even to unravel the genetic basis of dynamic traits. High-throughput precision phenotyping, using spectral reflectance measurements, has further the potential to provide more information for making better-informed management decisions at the canopy scale in real time (Mistele and Schmidhalter, 2010; Elsayed et al., 2011; Rischbeck et al., 2014). Similar to that, for detecting grain yield, protein and oil content of plants in the field, numerous observations are required to characterize a field. Proximal remote sensing systems for phenotyping on the field scale can be based on passive reflectance sensing. Passive sensor systems depend on sunlight as a source of light which allows hyperspectral information to be obtained in the visible and near-infrared range (Erdle et al., 2011; Elsayed et al., 2015). Passive reflectance sensors have widely been used to measure several canopy variables such as plant water status, biomass, leaf area index, nitrogen status or grain yield in cereals (Mistele and Schmidhalter, 2008; Erdle et al., 2011; Winterhalter et al., 2013). In this study, the use of passive reflectance sensing for identifying promising peanut cultivars was assessed to test whether pod yield, straw, seed protein and oil content can be predicted before harvest. Early prediction of grain yield by spectral reflectance measurements prior to harvest could reduce phenotyping time and expenses compared to destructive measurements (Marti et al., 2007; Prasad et al., 2007). To the best of our knowledge, there is very little information available about the assessment of seed protein and oil content, straw and pod yield of peanut by using hyperspectral reflectance measurements under field conditions. The use of high throughput remote or proximal sensing for assessing seed yield and grain protein content is less studied and rarely reported in the literature. Basnet et al. (2003) found that Landsat TM data was related with barley and wheat, with maximum coefficients of determination of 0.64 and 0.70, respectively. The reflectance of Landsat TM short-wave infrared (band 5) derived from canopy spectra or image data was related with protein content of grain R^2 = 0.31 and 0.37, respectively (Zhao et al., 2005). Hansen et al. (2002) reported that no relationship was obtained between the reflectance measurements and protein content of barley or wheat. Some studies were done for assessment of grain yield in cereals. For example, the normalized difference vegetation index (NDVI) at the milk-grain stage was well correlated to the final wheat grain yield at two levels of nitrogen fertiliser application under rainfed and irrigated conditions. However, it was also observed that the NDVI $(R_{774} - R_{656})/(R_{774} + R_{656})$ was also reasonably correlated to the grain yield at the onset of stem elongation (Marti et al., 2007). Gutierrez et al. (2006) found that the relationship between the green normalized difference vegetation index (GNDVI) $(R_{780} - R_{550}/R_{780} + R_{550})$ and seed yield of bean was higher (R^2 = 0.77) than the relationship between NDVI $(R_{900} - R_{680}/R_{900} + R_{680})$. Erdle et al. (2013) found that the spectral index (R_{760}/R_{730}) was related to the grain yield of wheat cultivars under different levels of nitrogen fertilizer. Lobos et al. (2014) found that the normalised water index 3 ((NWI-3; $(R_{970} - R_{920})/(R_{970} + R_{920})$) and the normalised difference vegetation index NDVI $(R_{830} - R_{660})/(R_{830} + R_{660})$ were most closely related to the grain yield of wheat genotypes. Robson et al. (2004) found that, the hyperspectral analysis of growing peanut leaves delivered significant correlations (up to r= 0.73**: p = 0.006) in predicting the pod maturity with canopy reflectance using the normalized difference vegetation index (NDVI) derived from high resolution multispectral satellite imagery. For reflectance measurements of growing leaves with field spectroscopy, the explanation of variance was greater than 90% using PLSR analysis with the wavelengths 640, 640, 747, 964 and 1124 nm. An alternative approach is to use partial least square regression (PLSR) of hyperspectral reflectance. Partial Least Square Regression (PLSR) creates orthogonal latent variables across the input variables (single wavebands) and relates them to the target variables (Elsayed et al., 2015). Partial Least Square Regression (PLSR) analysis is a chemometric technique that generalizes and combines the methods of Principal Component Analysis (PCA) and multiple regressions; it is used to predict a set of dependent variables from a large set of independent ones (i.e., predictors) that may be correlated. In PLSR orthogonal components, unaffected by collinearity, are derived from all variables. Partial least square models of hyperspectral reflectance were used by Weber et al. (2011). Elsayed et al. (2015) reported that partial least square regression could improve the assessment of the grain yield and the normalized relative canopy temperature of barley cultivars. PLSR analysis enabled the prediction of the N status in wheat and corn from ground-based spectral data (Alchanatis et al., 2005; Bonfil et al., 2005) and in forests from hyperspectral images (Smith et al. 2002; Coops et al., 2003). The purpose of this work was to evaluate the performance of passive sensing to: i.e. (i) assess whether spectral indices obtained at three BBCH growth stages (Lancashire et al., 1991), the beginning of pod development and 50% and 80% of pod development (BBCH growth stages 71, 85 and 89) can reflect changes in the pod and straw yield, seed protein and oil content of peanut cultivars under three levels of nitrogen fertilizer application (ii) and to compare the performance of spectral reflectance indices and partial least square regression to assess the pod and straw yield, and seed protein and oil contents of two peanut cultivars.

Results

Variation in the pod and straw yield, protein and oil content of peanut cultivars under three levels of nitrogen fertilizer

The cultivar Giza 5 presented the highest pod yield, straw, and seed protein content of 1.272 ton/fed, 1.95 ton/fed and 28.90 (%), respectively under T3 (N 40 kg/fed). While the cultivar Giza 6 presented the highest oil content of 43.79 (%) under T2 (N 20 kg/fed). The values of the mean pod yield, straw yield and seed protein content of peanut cultivars varied positively with the increasing nitrogen level, whereas the values of the mean oil content varied negatively with the increasing nitrogen level (Table 2). Significant differences ($P \leq 0.05$) were found for the pod yield, straw yield and seed protein and oil content of peanut cultivars among the treatments of nitrogen fertilizer application.

Contour map analysis of the hyperspectral passive data

A contour map analysis produced the mean coefficients of determination (R^2) of the three measurement dates for all dual wavelength combinations as a normalized difference spectral index. Contours of the matrices of the hyperspectral passive sensor presented generally more distinct relationships with the pod and straw yield, as well as the seed protein and oil content of peanut cultivars in the visible area than the combination of visible and near infrared wavelengths at the three measurements dates. The contour map analysis of the relationship between the normalized difference spectral indices with the pod yield, seed protein content and straw generally showed higher coefficients of determination than a contour map analysis of the oil content (Fig.1 and 2).

The relationship between spectral reflectance indices with the pod and straw yield, protein and oil content of peanut cultivars under three levels of nitrogen fertilizer application

Across the three measuring dates, spectral indices were more closely correlated with the pod yield, seed protein content

Table 1. Formula, index abbreviation and references of different previously developed and new spectral indices used in this study.

Formula	Index abbreviation	References
$(R_{780} - R_{550})/(R_{780} + R_{550})$	HPS[1] 780_550	Gutiérrez et al., 2010
$(R_{780} - R_{510})/(R_{780} + R_{510})$	HPS 780_510	Mistele et al., 2012
$(R_{760} - R_{730})/(R_{760} + R_{730})$	HPS 760_730	Barnes et al., 2000
$(R_{698} - R_{420})/(R_{698} + R_{420})$	HPS 698_420	this work
$(R_{620} - R_{470})/(R_{620} + R_{470})$	HPS 620_470	this work
$(R_{610} - R_{470})/(R_{610} + R_{470})$	HPS 610_470	this work
$(R_{610} - R_{450})/(R_{610} + R_{450})$	HPS 610_450	this work
$(R_{610} - R_{430})/(R_{610} + R_{430})$	HPS 610_430	this work
$(R_{610} - R_{424})/(R_{610} + R_{424})$	HPS 610_424	this work
$(R_{502} - R_{458})/(R_{502} + R_{458})$	HPS 505_458	this work
$(R_{500} - R_{460})/(R_{500} + R_{460})$	HPS 500_460	this work

[1]HPS indicates hyperspectral passive sensor.

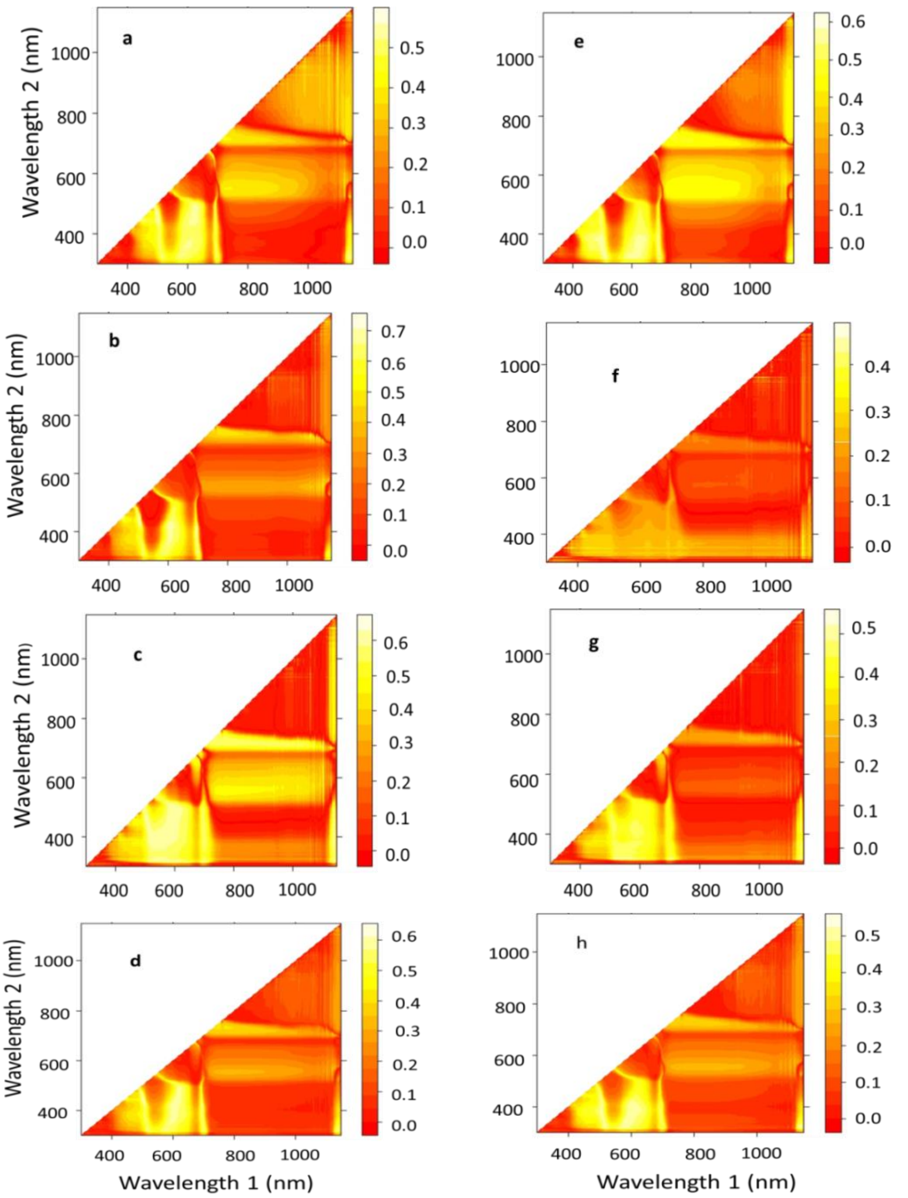

Fig 1. Correlation matrices (contour maps) showing the coefficients of determination (R^2) for all dual wavelength combinations in the range of 302–1148 nm (as a normalised difference index) of the hyperspectral passive reflectance sensor with the pod yield of two peanut cultivars: (a) at BBCH 71, (b) at BBCH 85, (c) at BBCH 88 and (d) mean coefficients of determination (R^2) of the spectral indices with the seed yield for three measurement dates and with the straw yield: (e) at BBCH 71, (f) at BBCH 85, (g) at BBCH 88 and (h) mean coefficients of determination (R^2) of the spectral indices with the straw yield for the three measurement dates.

Table 2. Average seed and straw yield, protein content and oil content of two Peanut cultivars at three levels of nitrogen fertilizer application in 2013. Values with the same letter are not significantly different ($P \geq 0.05$) among treatments according to Duncan's test. SD indicates standard deviation.

Cultivars	Treatments	Pod yield ton/fed	SD ton/ fed	Straw ton/fed	SD ton/ fed	Seed protein content (%)	SD (%)	Seed oil content (%)	SD (%)
Giza 5	N 10 kg/fed.	1.044 c	0.07	1.75 cd	0.08	22.08 d	0.64	42.61 b	0.82
	N 20 kg/fed	1.192 ab	0.15	1.81b c	0.08	27.19 b	1.5	40.81 c	0.91
	N 40 kg/fed	1.272 a	0.02	1.95 a	0.06	28.90 a	0.71	40.97 c	0.77
Giza 6	N 10 kg/fed.	0.800 e	0.07	1.58 e	0.07	21.10 d	0.45	42.30 b	0.94
	N 20 kg/fed	0.947 d	0.08	1.7 d	0.10	25.50 c	1.74	43.79 a	0.78
	N 40 kg/fed	1.112 bc	0.08	1.87 b	0.09	26.51bc	1.18	38.86 d	1.41

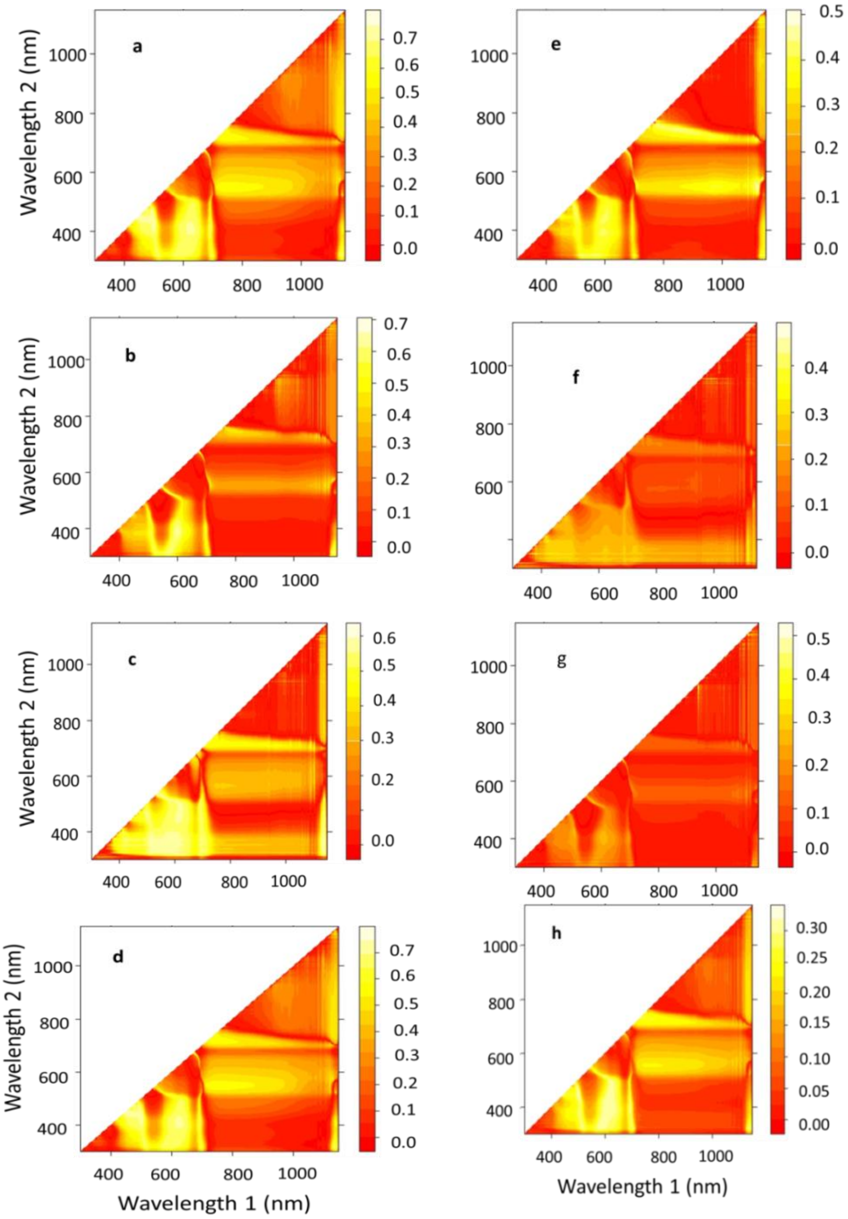

Fig 2. Correlation matrices (contour maps) showing the coefficients of determination (R^2) for all dual wavelength combinations in the range of 302–1148 nm (as a normalised difference index) of the hyperspectral passive reflectance sensor with the seed protein content of two peanut cultivars: (a) at BBCH 71, (b) at BBCH 85, (c) at BBCH 88 and (d) mean coefficients of determination (R^2) of the spectral indices with the seed protein content for three measurement dates and with the seed oil content: (e) at BBCH 71, (f) at BBCH 85, (g) at BBCH 88 and (h) mean coefficients of determination (R^2) of the spectral indices with the seed oil content for three measurement dates.

and straw yield than the seed oil content of the two peanut cultivars. The obtained coefficients of determination (R^2) are shown in Tables 3 and 4. Linear relationships were chosen to assess the relationship between the normalized difference spectral indices with the pod and straw yield, the seed protein and oil content. The closest significant relationships for the hyperspectral passive sensor were found for the pod yield with R^2 values ranging from 0.16*** to 0.70***, straw yield with R^2 values ranging from 0.10** to 0.53*** as shown in Table 3, as well as the seed protein content with R^2 values ranging from 0.28*** to 0.69***, and the seed oil content with R^2 values ranging from 0.05* to 0.43*** as indicated in Table 4. The normalized spectral index of HPS 610_424 showed the highest coefficients of determination (0.70***, 0.69*** and 0.53***) with pod yield, seed protein content and straw yield, respectively, at the growth stage BBCH 85 as shown in Tables 3 and 4. For the normalized difference spectral index HPS 500_470 a coefficient of determination of 0.43*** was obtained for the oil content at the growth stage BBCH 71 (Table 3).

Partial least squares regression analysis to predict pod and straw yield, protein and oil content of peanut cultivars

In Table 5 the quality of the PLSR models is presented through adjusted coefficients of determination of calibration (R^2 cal) and validation (R^2 val), root mean square errors (RMSE cal and val) and the slope of the linear regressions for calibration and validation models at the three measurement dates. Across all calibration data sets, the closest relationships for pod yield ($R^2 = 0.80$***), for straw yield ($R^2 = 0.71$***) were recorded for seed protein content ($R^2 = 0.82$***) at growth stage BBCH 71 and for seed oil content ($R^2 = 0.68$***) at growth stage BBCH 88. Across all validation data set formations, the highest coefficients of determination, with $R^2 = 0.7$*** for pod yield at growth stage BBCH 71, with $R^2 = 0.45$*** for straw yield at growth stage BBCH 88, and $R^2 = 0.69$*** for seed protein content at growth stage BBCH 71 and with $R^2 = 0.49$*** for oil content were recorded. Across all calibration and validation data set formations, the RMSEC varied from 0.08 to 0.12 ton per feddan for pod yield, from 0.08 to 0.11 ton per feddan for straw yield, from 1.25 % to 1.95 % for seed protein content and from 1.02 % to 1.52 % for seed oil content. The highest slope values for calibration and validation data sets (0.82 and 0.72) were recorded at growth stage BBCH 71.

Discussion

A hyperspectral passive reflectance sensor was used in this study to assess the pod yield, straw yield, seed protein and oil content of two peanut cultivars under three levels of nitrogen fertilizer application. The difference in the doses of nitrogen fertilizer which were added to the soil affected the pod and straw yield, seed protein and oil content of two peanut cultivars (Table 2). Weiser and Seilmeier (1998) reported that the protein content was strongly influenced by nitrogen. The content of protein increased significantly with increasing dose of nitrogen application. El-Habbasha et al. (2013) found that there were positive relationships between increased nitrogen levels from 30 and 40 N kg/feddan and pod yield, seed protein content and straw yield. In contrast there was a negative relationship between increased nitrogen fertilizer application and seed oil content. These results agree with our results obtained for two peanut cultivars (Table 2), where the increased nitrogen fertilizer dose and the pod yield, seed protein content and straw yield increased whereas the seed oil content decreased. Between the peanut cultivars, significant differences in mean pod yield, straw, seed protein and oil content were found at three levels of nitrogen fertilizer application (Table 3). These results agree with the findings of Gohari et al. (2010) and Moraditochaee (2012), who reported that there were variations in the pod yield and straw yield of peanut cultivars under different levels of nitrogen fertilizer application. In this study, high-throughput passive sensing was found to present a major advantage. Spectral measurements could be performed simultaneously and in a short time by using a tractor as mobile carrier platform. Fast measurements can reduce disturbances caused by shifting illumination. The passive optical sensor system was mounted on a frame in front of a tractor-based measuring platform. The sensor was driven to measure the spectral reflectance of all plots. In several other studies (Ferrio et al., 2005; Inman et al., 2007; Prasad et al., 2007; Gutierrez et al., 2010) handheld sensors were used for spectral measurements of plants. This method is more time consuming and may, therefore, be more affected by external factors, such as ambient climatic conditions. To further minimize disturbing effects by external factors, bi-directional measurements of the incident and reflected radiation were used to calculate optimised vegetation indices by creating matrix contour maps for the three measurements dates (Figs 1 and 2) and the more stable and strong spectral indices were chosen (Table 1).

Averages of the correlation matrices resulting from the three measurements dates as indicated by the coefficients of determination (R^2) for all dual wavelengths combinations in the range of 400–750 nm (as a normalised difference index for pod and straw yield, protein and oil content of two peanut cultivars (Fig. 1 d and h and Fig. 2 d and h) presented higher R^2-values compared to the other combinations of two wavelengths. Maybe this is due to the range from 400 to 750 nm is more affected by the chlorophyll *a* and *b* as well as by the leaf area index. In this study to create a variation in the straw yield, as well as the protein and oil content, the peanut cultivars were exposed to three doses of nitrogen fertilizer. These results agree with Gates et al. (1965) and Townsend et al. (2003), who reported that the range from 400 to 530 nm rapidly changes from relatively low reflectance in the blue region to higher reflectance in the green region and in the range from 680-730 nm straddling the red edge. The spectra in the 450–530 nm range are strongly influenced by the presence and abundance of chlorophyll *a* and *b*. In contrast, the spectra in the 680-730 nm range may be correlated with the leaf area index. Our assessment of reflectance indices as a method to predict the pod and straw yield, and the seed protein content of peanut cultivars demonstrated that the selected four indices such as HPS 610_424, HPS 610_430, HPS 698_420, HPS 620_470 which were derived from the visible region as well as HPS 760_730 and HPS 780_550, which were derived from the near infrared range or combinations between visible and near infrared regions are apparently useful for describing these parameters (Tables 3 and 4). The use of the full spectral range from 302 - 1148 nm for the PLSR analysis increased the accuracy of the estimates of the pod and straw yield, seed protein and oil content compared with the use of the normalised difference spectral indices (Tables 3, 4 and 5). The calibration model of the PLSR was strongly related to the pod and straw yield, as well as the seed protein and oil content compared with the normalised difference spectral indices (Tables 3, 4 and 5). Comparably, in peanut cultivars, the assessment of the pod yield, straw and seed protein and oil content was stronger and more robust when using PLSR models than with the previously assayed normalised difference spectral indices.

Table 3. Coefficients of determination of linear regressions of pod and straw yield with spectral indices of the hyperspectral passive sensor (HPS) (calculated as normalised difference indices) for peanut cultivars subjected to three levels of nitrogen.

Spectral indices	Parameters	3 August 2013 BBCH 71	24 August2013 BBCH 85	2 September 2013 BBCH 88
HPS 780_550	Pod yield	0.40***	0.35***	0.47***
	Straw	0.45***	0.26***	0.20**
HPS 760_730	Pod yield	0.37***	0.43***	0.56***
	Straw	0.45***	0.27***	0.24***
HPS 760_510	Pod yield	0.25***	0.16*	0.23**
	Straw	0.38***	0.10*	0.03
HPS 698_420	Pod yield	0.46***	0.67***	0.57***
	Straw	0.42***	0.45***	0.45***
HPS 620_470	Pod yield	0.53***	0.61***	0.59***
	Straw	0.49***	0.38***	0.41***
HPS 610_450	Pod yield	0.52***	0.62***	0.60***
	Straw	0.46***	0.42***	0.42***
HPS 610_430	Pod yield	0.53***	0.68***	0.59***
	Straw	0.48***	0.49***	0.42***
HPS 610_424	Pod yield	0.51***	0.70***	0.60***
	Straw	0.49***	0.53***	0.42***
HPS 502_458	Pod yield	0.56***	0.68***	0.58***
	Straw	0.51***	0.48***	0.44***
HPS 500_470	Pod yield	0.56***	0.62***	0.58***
	Straw	0.38***	0.10*	0.45***
HPS 500_460	Pod yield	0.55***	0.67***	0.57***
	Straw	0.52***	0.47***	0.44***

*, **, *** Statistically significant at P ≤ 0.05; P ≤ 0.01 and P ≤ 0.001, respectively.

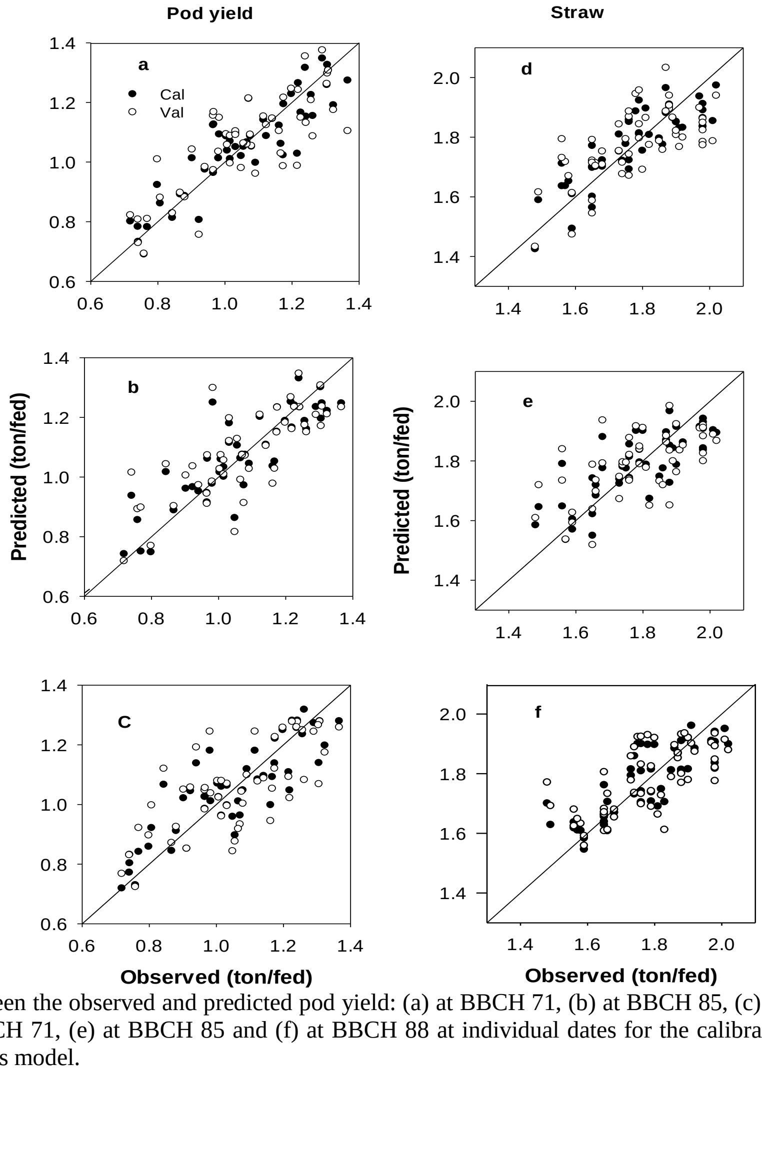

Fig 3. Relationships between the observed and predicted pod yield: (a) at BBCH 71, (b) at BBCH 85, (c) at BBCH 88, as well as for the straw yield (d) at BBCH 71, (e) at BBCH 85 and (f) at BBCH 88 at individual dates for the calibration and validation datasets using a partial least squares model.

Table 4. Coefficients of determination of linear regressions of seed protein and oil content with spectral indices of the hyperspectral passive sensor (HPS) (calculated as normalised difference indices) for peanut cultivars subjected to three levels of nitrogen.

Spectral indices	Parameters	3 August 2013 BBCH 71	24 August2013 BBCH 85	2 September 2013 BBCH 88
HPS 780_550	Protein content	0.53***	0.37***	0.47***
	Oil content	0.33***	0.13*	0.10*
HPS 760_730	Protein content	0.53***	0.43***	0.54***
	Oil content	0.33***	0.13*	0.16*
HPS 760_510	Protein content	0.32***	0.17*	0.28***
	Oil content	0.17***	0.05	0.03
HPS 698_420	Protein content	0.54***	0.61***	0.58***
	Oil content	0.33***	0.18**	0.23***
HPS 620_470	Protein content	0.66***	0.59***	0.59***
	Oil content	0.39***	0.20***	0.22***
HPS 610_450	Protein content	0.64***	0.59***	0.57***
	Oil content	0.41***	0.21**	0.22**
HPS 610_430	Protein content	0.54***	0.69***	0.57***
	Oil content	0.41***	0.24***	0.23***
HPS 610_424	Protein content	0.62***	0.69***	0.57***
	Oil content	0.40***	0.25***	0.23**
HPS 502_458	Protein content	0.56***	0.65***	0.57***
	Oil content	0.43***	0.24***	0.24***
HPS 500_470	Protein content	0.67***	0.57***	0.57**
	Oil content	0.43***	0.21**	0.24***
HPS 500_460	Protein content	0.66***	0.65***	0.53***
	Oil content	0.42***	0.24***	0.25***

*, **, *** Statistically significant at $P \leq 0.05$; $P \leq 0.01$ and $P \leq 0.001$, respectively.

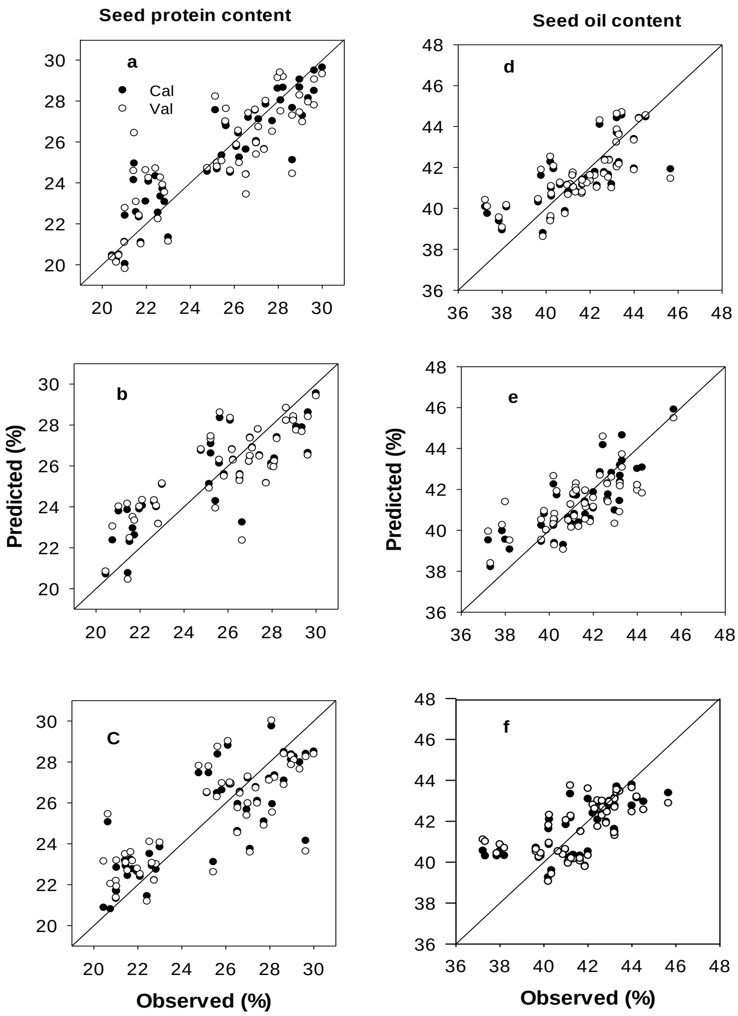

Fig 4. Relationships between the observed and predicted seed protein content: (a) at BBCH 71, (b) at BBCH 85, (c) at BBCH 88, as well as for the seed oil content (d) at BBCH 71, (e) at BBCH 85 and (f) at BBCH 88 at individual dates for the calibration and validation datasets using a partial least squares model.

Table 5. Calibration (R² cal, RMSEC and slope cal), and 7-fold cross-validation (R² val, RMSEV and slope val) of the statistics of partial least square regression models of the spectral reflectance from 302 to 1148 nm with the pod and straw yield, protein content and oil content of peanut cultivars.

Growth stages	Statistical parameters	[1]PCs	Pod yield (ton/fed)	PCs	Straw yield (ton/fed)	PCs	Seed protein content (%)	PCs	Seed oil content (%)
BBCH 71		5		5		6		4	
	Cal R^2		0.80***		0.71***		0.82***		0.59***
	Val R^2		0.65***		0.43***		0.69***		0.49***
	RMSEC		0.08		0.08		1.25		1.17
	RMSEV		0.11		0.11		1.67		1.31
	Slope cal		0.80		0.71		0.82		0.59
	Slope val		0.70		0.54		0.72		0.54
BBCH 85		5		4		3		6	
	Cal R^2		0.73***		0.60***		0.71***		0.68***
	Val R^2		0.60***		0.31***		0.60***		0.29***
	RMSEC		0.09		0.09		1.54		1.02
	RMSEV		0.11		0.11		1.85		1.52
	Slope cal		0.73		0.60		0.71		0.68
	Slope val		0.63		0.45		0.69		0.58
BBCH 88		4		4		3		3	
	Cal R^2		0.75***		0.67***		0.69***		0.50***
	Val R^2		0.57***		0.45***		0.59***		0.34***
	RMSEC		0.09		0.08		1.68		1.29
	RMSEV		0.12		0.11		1.94		1.48
	Slope cal		0.75		0.67		0.69		0.50
	Slope val		0.61		0.53		0.60		0.42

*** Statistically significant at $P \leq 0.001$, respectively [1]PCs, Number of latent variables. Cal, Calibration, Val, Validation,RMSEC, Root mean square error for calibration, RMSEV, Root mean square error for validation

This is shown by the improvement of the coefficients of determination in Tables 3 and 4 and the results of the cross validation in Table 5. These results agree with Sharabian et al. (2014), who found that strong relationships existed between the predicted and observed values for a validation dataset of grain yield (R^2 = 0.87, RMSE = 301) and protein content (R^2 = 0.80, RMSE = 6.8) and SPAD values (R^2 = 0.84, RMSE = 1.94). Li et al. (2014), who found that the PLSR is a potentially useful approach to derive the canopy nitrogen concentration of winter wheat across growth stages, compared with spectral indices, and the average value of the coefficient of determination for the PLSR model increased to 76.8% and 75.5% in the calibration and validation datasets, respectively. The advantage of PLSR models compared with spectral index models is that the PLSR in this study used information from every spectral band from 302 to 1148 nm and selected the number of factors to best represent the calibration data without overfitting. PLSR had no limitation in predicting the pod yield, straw and seed protein and oil content and the relationship between the observed and predicted values was linear. The calibration model of the PLSR and spectral indices (Tables 3, 4, 5) presented higher coefficients of determination with the pod yield and the seed protein content than with the straw yield and the seed oil content.

Materials and Methods

Field experiments and design

Field experiments were conducted at the research station of the Sadat City University in Egypt. The research station of the Sadat City University is characterized by a semi-arid climate with moderate cold winters and warm summers. The experiment was a two factorial set up with two peanut cultivars, Giza 5 and Giza 6, three fertilizer rates with T1 (N 10 kg/feddan), T2 (N 20 kg/feddan) and T3 (N 40 kg/feddan) and nine replicates for each treatment. The fertilizer was added in two equal doses, immediately at sowing and 21 days later as ammonium nitrate (NH_4NO_3 33.5%). All treatments received the recommended dose from superphosphate (15.5% P_2O_2) at a rate of 200 kg/feddan and potassium sulfate (48% K_2O) at a rate of 50 kg/feddan. Drip irrigation was used. The peanut cultivars were sown on 15 May 2013 in sandy loam soil that contains (72.8 % sand, 19. 4 % silt and 7.9 % clay). The soil was characterized by an electrical conductivity of 1.82 dS m^{-1}, 0.36% organic matter and 5% calcium carbonate. The plots consisted of 3 rows spaced 60 cm apart and had a length of 4 m. Herbicide and fungicide treatments were applied in all trials when necessary. At harvest time, a random sample of 10 plants was taken from each plot to determine the averages of the pod and straw yield. After pod drying, the pod yield (ton/fed) was determined. In addition, samples of seeds obtained from 100-pods collected randomly from each plot were used to determine: (1) total nitrogen by using the micro-Kjeldahl method, as described by A.O.A.C. (1980), the protein content in seeds was calculated by multiplying the total nitrogen percentage by a factor of 6.25. The seed oil percentage was determined by the NMR method as described by Granhund and Zimmerman (1975).

Spectral reflectance measurements

For spectral reflectance measurements, a passive bi-directional reflectance sensor (tec5, Oberursel, Germany) measuring at wavelengths between 302 - 1148 nm with a bandwidth of 2 nm and connected to a portable computer and geographical positioning system (GPS) was used. The passive reflectance sensor consists of two units; one unit was linked with a diffuser and measured the light radiation as a reference signal. The second unit simultaneously measured the canopy reflectance with a fiber optic (Mistele and

Schmidhalter, 2008; Elsayed et al., 2011).The aperture of the optics was 12° and the field of view was 0.2 m^2 from one meter height. Collecting information in the field, the sensor outputs were co-recorded along with the GPS coordinates. For each position, the actual sensor output was co-referenced and recorded. Afterwards, readings within one plot were averaged to single values per plot. With the readings from the spectrometer unit the canopy reflectance was calculated and corrected with a calibration factor obtained from a reference grey standard. Spectral measurements were were taken mostly on sunny days at the nadir direction about 0.75 m above the canopy. The sensor was mounted on a frame in front of a tractor and was driven to measure the spectral reflectance of all plots at three measurement dates on 3 August 2013 at BBCH 71, 24 August 2013 at BBCH 85 and 2 September 2013 at BBCH 88.

Statistical analysis

Selection of spectral reflectance indices and single reflectance bands

In Table 1 eleven spectral indices from different sources are listed with reference. We calculated and tested both known and novel indices. A contour map analysis for all wavelengths of the hyperspectral passive sensor (from 302 to1048 nm) was used to select some normalized difference indices, which generally presented more stable and strong relationships with pod yield, straw and seed protein and oil content of peanut cultivars under three levels of nitrogen fertilizer (Figs.1 & 2). All possible dual wavelengths combinations were evaluated depending on a contour map analysis of the reflectance measurements. Contour maps are matrices of the coefficients of determination of pod and straw yield, seed protein and oil content of peanut cultivars with all possible combinations of binary, normalized spectral indices (Fig. 1 and 2). The R package "lattice" from the software R statistics version 3.0.2 (R foundation for statistical computing 2013) was used to produce the contour maps from the hyperspectral reflectance readings, seventeen wavelengths (780, 760, 730, 698, 620, 610, 550, 510, 502, 500, 470, 460, 458, 450, 430, 424, and 420nm) were subsequently selected for the evaluation of optimized reflectance indices.

Modelling of the pod and straw yield, seed protein and oil content of peanut cultivars

Sigmaplot for Windows v.12 (Systat software Inc., Chicago), and SPSS 16 (SPSS Inc., Chicago, IL, USA were used for the statistical analysis. Simple linear regressions were calculated to analyse the relationship between spectral indices listed in Table 1 with the pod and straw yield, seed protein and oil content (Tables 3 and 4). Coefficients of determination and significance levels were determined; nominal alpha values of 0.05, 0.01 and 0.001 were used (Tables 3 and 4). The Unscrambler X multivariate data analysis software version 10.2 (CAMO Software AS, Oslo) was used to calibrate and validate partial least square models. Single wavebands derived from the same spectra usually contain redundant information (Sharabian et al., 2014). Partial Least Square Regression (PLSR) creates orthogonal latent variables across the input variables (single wavebands) and relates them to the target variables (seed and straw yield, seed protein and oil content). This is a way to cope with redundancy in the input variables. The PLSR searches the sensitive information from spectral reflectance for all wavelengths. For the hyperspectral passive sensor, all wavelengths from 302 to 1148 were used as input variables in the PLSR models shown in Table 5. For the model the datasets from the three measurements dates at 3 August 2014, 24 August 2013 and 2 September 2013 were used. For determining the model quality one approach of validation were used. In Table 5 a (7 fold) cross validation approach was applied for the PLSR models. Calibration and validation quality of models is presented through adjusted coefficients of determination of calibration (R^2cal) and validation (R^2val), root mean square errors for calibration (RMSEC) and for validation (RMSEV) and the slope of the linear regressions between observed and predicted values of the pod yield, straw and seed protein and oil content. Scatter plots of seed yield, straw and seed protein and oil content predicted from calibration and validation models with observed data for three measurement dates are shown in Figures 3& 4.

Conclusions

The results show that the models developed from the normalised difference spectral indices analysis and PLSR analysis reliably assessed the pod and straw yield, as well as the seed protein and oil content of two peanut cultivars. Partial least square regression models of reflectance measurements potentially improve non-invasive measurements of pod and straw yields, as well as seed protein and oil contents of peanut cultivars compared with normalised difference spectral indices analysis.

Acknowledgements

This research was partly supported by the International Bureau of the German Federal Ministry of Education and Research (Project No: 01DH12046)

References

AbouKheira AA (2009) Macro management of deficit-irrigated peanut with sprinkler irrigation. Agric Water Manage. 96:1409-1420.

Alchanatis V, Schmilovitch Z, Meron M (2005) In-field assessment of single leaf nitrogen status by spectral reflectance measurements. Precis Agric. 6:25-39.

A.O.A.C (1980) Official methods of analysis. 13th Ed. Association of official analytical chemists Washington, D.C., USA.

Barik AK, Mukherjee AK, Mendal BK (1998) Growth and yield of Sorghum (*Sorghum bicolor*) and groundnut (*Arachis hypogaea* L.) grown as sole and intercrops under different regimes. Indian J Agron. 43:27-32.

Barnes EM, Clarke TR, Richards SE, Colaizzi PD, Haberland J, Kostrzewski M (2000) Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, MN, USA.

Basnet BB, Apan AA, Kelly RM, Jensen TA, Strong WM, Butler DG (2003) Relating satellite imagery with grain protein content. Proceedings of the First Spatial Sciences Coalition Conference, 22-26 September 2003, Canberra, pp. 11.

Bonfil DJ, Karnieli A, Raz M, Mufradi I, Asido S, Egozi H, Hoffman A, Schmilovitch Z (2005) Rapid assessing of water and nitrogen status in wheat flag leaves. J Food Agric Environ. 3:148-153.

Coops NC, Smith ML, Martin ME, Ollinger SV (2003) Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. IEEE T Geosci Remote. 41:1338-1346.

El-Habbasha SF, Taha MH, Jafar NA (2013) Effect of nitrogen fertilizer levels and zinc foliar application on yield, yield attributes and some chemical traits of groundnut. Res J Agr Biol Sci. 9:1-7.

Elsayed S, Mistele B, Schmidhalter U (2011) Can changes in leaf water potential be assessed spectrally? Funct Plant Biol. 38:523-533.

Elsayed S, Rischbeck P, Schmidhalter U (2015) Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars. Field Crop Res. 177:148-160.

Erdle K, Mistele B, Schmidhalter U (2011) Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. Field Crop Res. 124:74-84.

Erdle K, Mistele B, Schmidhalter U (2013) Spectral assessments of phenotypic differences in spike development during grain filling affected by varying N supply in wheat. J Plant Nutr Soil Sci. 6:952-963.

Fageria NK, Baligar VC, Jones C (1997) Growth and mineral nutrition of field crops 2nd Ed. Marcel Dekker, Inc, New York 1001 k, pp. 494.

Ferrio JP, Villegas D, Zarco J, Aparicio N, Araus JL, Royo C (2005) Assessment of durum wheat yield using visible and near-infrared reflectance spectra of canopies. Field Crop Res. 94:126-148.

Furbank RT, Tester M (2011) Phenomics – technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 16:635-643.

Gates DM, Keegan HJ, Schleter JC, Weidner VR (1965) Spectral properties of plants. Appl Optics. 4:11-20.

Gohari A, Niyaki S (2010) Effects of iron and nitrogen fertilizers on yield and yield components of peanut (*Arachis hypogaea* L.) in Astaneh Ashrafiyeh, Iran. Am-Euras J Agric Environ Sci. 9:256-262.

Granhund M, Zimmeren DC (1975) Effect of drying conditions on soil content of sunflower (*Helianthus annus* L.) seeds as determined by wide–line nuclear magnetic resonance (NMR). North Dacota Academy of Science, Grand Forks ND.

Gutiérrez M, Estrada J, Gonzales M, Reynolds, MP (2006) Canopy reflectance indices and its relationship with yield in common bean plants (*Phaseolus vulgaris* L.) with phosphorous supply. Int J Agr Biol. 8:203-207.

Gutiérrez M, Reynolds MP, Raun WR, Stone, ML, Klatt AR (2010) Spectral water indices for assessing yield in elite bread wheat genotypes in well irrigated, water stressed, and high temperature conditions. Crop Sci. 50:197-214.

Hansen PM, Jorgensen JR, Thomas A (2002) Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression. J Agr Sci. 139:307-318.

Inman D, Khosla R, Reich RM, Westfall DG (2007) Active remote sensing and grain yield in irrigated maize. Precis Agric. 8:241-252.

Lancashire PD, Bleiholder H, Boom TVD, Langelüddeke P, Stauss R, Weber E, Witzenberger A (1991) A uniform decimal code for growth stages of crops and weeds. Ann Appl Biol. 119:561-601.

Li F, Mistele B, Hu Y, Chen X, Schmidhalter U (2014) Reflectance estimation of canopy nitrogen content in winter wheat using optimized hyperspectral spectral indices and partial least squares regression. Eur J Agron. 52:198-209.

Lobos GA, Matus I, Rodriguez A, Romero-Bravo S, Araus JL, Pozo AD (2014) Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. J Integr Plant Biol. 56:470-479.

Marti J, Bort J, Slafer GA, Araus JL (2007) Can wheat yield be assessed by early measurements of normalized difference vegetation index? Ann Appl Biol. 150:253-257.

Michael J, Ottman T, Doerge A, Edward C, Martin E (2000) Durum grain quality as affected by N–fertilizer near anthesis and irrigation during grain fills. Agron J. 92:1035-1041.

Mistele B, Elsayed S, Schmidhalter U (2012) Assessing water status in wheat under field conditions using laser induced chlorophyll fluorescence and hyperspectral measurements. 11th International Conference on Precision Agriculture. Indianapolis, Indiana USA.

Mistele B, Schmidhalter U (2008) Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. Eur J Agron. 29:184-190.

Mistele B, Schmidhalter U (2010) Tractor-based quadrilateral spectral reflectance measurements to detect biomass and total aerial nitrogen in winter wheat. Agron J. 102:499-506.

Moraditochaee M (2012) Effects of humic acid foliar spraying and nitrogen fertilizer management on yield of peanut (*Arachis hypogaea* L.) in Iran. ARPN J Agr Biol Sci. 7:289-293.

Peleman JD, van der Voort JR (2003) Breeding by design. Trends Plant Sci. 8:330-334.

Pendashteh M, Tarighi F, Doustan HZ, Keshavarz A, Azarpour E, Moradi M, Bozorgi HR (2011) Effect of foliar zinc spraying and nitrogen fertilization on seed yield and several attributes of groundnut (*Arachis hypogaea* L.). World Appl Sci J. 13:1209-1217.

Prasad B, Carver BF, Stone ML, Babar MA, Raun WR, Klatt AR (2007) Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under Great Plains conditions. Crop Sci. 47:1426-1440.

Rischbeck P, Baresel JP, Elsayed S, Mistele B, Schmidhalter U (2014) Development of a diurnal dehydration index for spring barley phenotyping. Funct Plant Biol. 41:1249-1260.

Robson A, Wright G, Phinn S (2004) Assessment of peanut crop maturity and yield forecasting with quick bird multispectral satellite imagery. Spat Sci Queensland. 4:35-37.

Sharabian RV, Noguchi N, Ishi K (2014) Significant wavelengths for prediction of winter wheat growth status and grain yield using multivariate analysis. Eng Agr Environ Food. 7:14-21.

Smith ML, Ollinger SV, Martin M E, Aber JD, Hallett RA, Goodale CL (2002) Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. Ecol Appl. 12:1286-1302.

Tiwari RB, Dhakar LL (1997) Productivity and economics of summer groundnut (*Arachis hypogaea*) as affected by irrigation, fertilizes and weed control. Indian J Agron. 42:490-494.

Townsend PA, Foster JR, Chastain RA, Currie WS (2003) Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. IEEE T Geosci Remote. 41:1347-1354.

Walter A, Studer B, Kölliker R (2012) Advanced phenotyping offers opportunities for improved breeding of forage and turf species. Ann Bot. 110: 1271-1279.

Weber VS, Araus JL, Cairns JE, Sanchez C., Melchinger AE, Orsini E (2012) Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. Field Crop Res. 128:82-90.

Wieser H, Seilmeier W (1998) The influence of nitrogen fertilization on quantities and proportions of different protein types in wheat flour. J Sci Food Agr. 76:127-131.

Winterhalter L, Mistele B, Schmidhalter U (2012) Assessing the vertical footprint of reflectance measurements to characterize nitrogen uptake and biomass distribution in maize canopies. Field Crop Res. 129:14-20.

Winterhalter L, Mistele B, Schmidhalter U (2013) Evaluation of active and passive sensor systems in the field to phenotype maize hybrids with high-throughput. Field Crop Res. 154:236-245.

Zhao C, Liu L, Wang J, Huang W, Song X, Li C (2005) Predicting grain protein content of winter wheat using remote sensing data based on nitrogen status and water stress. Int J Appl Earth Obs. 7:1-9.