European Journal of Agronomy 140 (2022) 126609

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Evaluating and defining agronomically relevant detection limits for spectral reflectance-based assessment of N uptake in wheat

Paul Heinemann [a], Stephan Haug [b], Urs Schmidhalter [a,*]

[a] Technical University of Munich, Chair of Plant Nutrition, Emil-Ramann-Str. 2, 85354 Freising-Weihenstephan, Germany
[b] Technical University of Munich, Chair of Mathematical Statistics, Parkring 11–13, 85748 Garching b, München, Germany

ARTICLE INFO

Keywords:
Digital agriculture
Drone
Fertilizer experiment
Nitrogen management
Nitrogen
Fertilization
Precision agriculture
Remote sensing
Spectral reflectance

ABSTRACT

Spectral detection of the N uptake of wheat is a widely used method, as it is non-destructive, rapid and cost-efficient. However, to date, agronomically supported spectral detection limits have not been sufficiently described. Concurrent statistical measures such as R^2, RMSE, or MAE do not fully satisfactorily address the agronomical relevance, bearing in mind that sensing is frequently carried out in either very low or excessively high nitrogen fertilization applications, which may not be indicative of current farming practice. This study, therefore, evaluates regression models of spectral indices in capturing N uptake using hyperspectral ground-based and multispectral unmanned aerial vehicles (UAV) based on data sets covering several years, sites, varieties, and developmental stages of wheat (*Triticum aestivum* L.). The results suggest that solely adopting commonly used statistical measures is not sufficient for an agronomic evaluation. Whereas the common statistical measure R^2 is essentially influenced by differentiating N uptake, which primarily occurs at later stages of development, the use of further statistics such as RMSE and MAE averages the error and should be extended by self-set confidence intervals based on agronomical decisions. For these studies, we therefore defined an appropriate error interval of ± 15 kg N uptake ha^{-1} up to BBCH 50, with a probability of at least 80%. Interval limits may be narrowed for earlier developmental stages and wider for later ones. Extreme N levels in field trials can bias models and should be limited to N-fertilization ranges that are indicative of the current practice in a given region, so as not to overemphasize the extremes. Differentiation of biomass was revealed to be more crucial than that of N content in detecting N uptake. Essentially, both terrestrial- and UAV-based sensing were equally well suited, with combinations of the REDEDGE and NIR bands being particularly effective for detecting the N uptake of wheat. Agronomically based detection limits should be considered besides common statistical measures in the spectral assessment of wheat N uptake.

1. Introduction

Wheat is one of the world's key cereal crops, along with rice and corn. Since wheat production is often limited by the availability of nitrogen (N), it frequently requires additional N-fertilization (Ladha et al., 2005). Although N plays a major role in the yield and quality of crops (Robertson and Vitousek, 2009; Barmeier et al., 2017; Prey et al., 2019a. b), some of the nitrogen used is lost to the air, water, and land, where it can cause both environmental and human health problems (Galloway et al., 2008). It is therefore important to minimize nitrogen losses by optimizing the process of N-fertilization.

N-fertilizer requirements are often determined on the basis of expected grain yields. Although Delogu et al. (1998) showed that both pre- and post-heading N uptake correlate closely with yield, the N uptake prior to anthesis benefits the grain N uptake and yield (Prey et al., 2019a,b; Ding et al., 2021). Recording N uptake during vegetation is therefore a key aspect of decision making concerning N-fertilization. N-fertilization can differ in terms of the absolute rate, splitting and application time (Darwinkel, 1983). In Western Europe, the overall quantity of N is typically divided into three splitting applications, which are applied approximately at tillering (BBCH 20–29) and at the beginning and the end of stem elongation (BBCH 30–39) (Swarbreck et al., 2019; BBCH is the abbreviation of „**B**iologische **B**undesanstalt, Bundessortenamt und **CH**emische Industrie"; Meier, 2018).

There are several methods of assessing a plant's N requirements. At the beginning of vegetation, the additional N-fertilization demand is

* Corresponding author.
E-mail address: schmidhalter@wzw.tum.de (U. Schmidhalter).

https://doi.org/10.1016/j.eja.2022.126609
Received 12 February 2022; Received in revised form 10 July 2022; Accepted 8 August 2022
Available online 17 August 2022

calculated on the basis of the mineral nitrogen (nitrate-N + ammonium-N, "Nmin") already present in the soil (Heinemann and Schmidhalter, 2021). During vegetation, the plant itself is often used as an indicator, a common major method being the optical, non-destructive detection of the plant canopy's reflectance signature (Olfs et al., 2005; Li et al., 2014; Ali et al., 2017). The main measurement range of the sensors is between wavelengths of 400 nm and 1000 nm. Within this range, the vegetation displays a typical reflection signature, controlled in the visible wavelength range (400–700 nm) by the absorption of pigments (mainly chlorophyll *a* at 430/660 nm and chlorophyll *b* at 450/640 nm, plus other pigments such as carotenoids and xanthophylls at 450 nm) and in the near-infrared range (700–1100 nm) by reflection processes in the foliar layers (Lilienthal, 2014). Since N-fertilization affects the plant's aforementioned physiological parameters, the N supply, for example, for wheat, lends itself well to spectral recording (Fitzgerald et al., 2010; Mistele and Schmidhalter, 2010; Prey and Schmidhalter, 2019).

Many spectral indices are available for the detection of nitrogen traits in wheat, each differing in the spectral range (visible, red edge, or near-infrared) as well as in the wavelengths used (normalized difference, simple ratio, or red edge inflection point) (Prey et al., 2020). Both ground-based vehicles (Mistele and Schmidhalter, 2010; Erdle et al., 2011), and, since recently, unmanned aerial vehicles (UAV) (Gnyp et al., 2016; Chen et al., 2019; De Souza et al., 2021) have been shown to be suitable carriers of spectral sensors. The advantages of UAV's over ground-based systems is that they enable measuring without disrupting the surface, they can capture spatial information simultaneously, and they are able to generate high-resolution images (Aasen and Bolten, 2018).

The amount and timing of individual split applications of nitrogen are often decided on the basis of the farmer's experience. For example, high yield expectations in the 7–10 t ha^{-1} range necessitate an absolute nitrogen demand of 215–250 kg N ha^{-1} (Lfl, 2018). This absolute amount can be divided into split applications that are approximately within the 40–90 kg N ha^{-1} range. Both planned absolute and split application N-fertilizer rates can be modified during the growing season on the basis of changes in N uptake. Under these high-yield conditions, N uptake has a wide measuring range across the growing season, extending from less than 50 kg N ha^{-1} to more than 250 kg N ha^{-1}. In-season spectral measurements can be performed from tillering throughout the season, providing useful information on which to base decisions (Erdle et al., 2011; Li et al., 2014; Elsayed et al., 2018). If spectral measurements are used to determine the N uptake, the limits of the method's detectability must be taken into account. This not only impacts on the evaluation of the N uptake, it also influences impending N-fertilization decisions.

To determine the measurement quality of spectral sensors used for N uptake, nitrogen increase experiments are commonly performed. These require a priori determined N levels. To achieve a clear effect of the dependent factor N uptake, large differences between N levels are often chosen (Feng et al., 2015; Guo et al., 2018; Zhang et al., 2021). These N levels exceed the N fertilization range that is customary in practice in some cases to a significant degree. To determine the measurement quality, regressions between the dependent (e.g. N uptake) and the independent (e.g. spectral indices) variables are calculated and evaluated with the coefficient of determination (R^2). This raises the question as to the context in which the a priori defined N levels influence R^2 and whether the assessment of R^2 is crucial for the evaluation of the measurement quality. The detection quality of spectral sensors is another important factor in the range crucial for agricultural applications.

N uptake is the product of the N content and the dry weight of the aboveground biomass (DW). For many spectral indices, biomass is a crucial parameter (Mistele et al., 2004). However, many studies indicate that the detection of N content by spectral means is much less effective (Erdle et al., 2011; Lilienthal, 2014; Elsayed et al., 2018). Better understanding the potential for the spectral detection of N uptake using commonly applied indices requires information on the spectral detectability of both N content and aboveground biomass (DW), as well as the interaction between them, particularly in terms of the plant growth for the specific year.

The nitrogen nutrition index (NNI) was created to better understand the respective plant nitrogen supply. The NNI is the quotient of the current ($N_{act.}$) and critical ($N_{crit.}$) N content (% of dry matter). $N_{act.}$ is the measured N content of the plant samples under investigation. $N_{crit.}$ is the threshold value of the N content derived by Justes et al. (1997) and depends on the aboveground biomass (in the range 1.55–12 t ha^{-1}), below which N deficiency leads to reduced biomass production. NNI values ≥ 1 indicate a current N content that does not impede biomass growth, whereas values < 1 indicate limited biomass production due to currently insufficient N content (Justes et al., 1997; Mistele and Schmidhalter, 2008).

This study uses data taken from a newly established field trial (Experiment 1) in conjunction with published data (Experiments 2 and 3) to assess the influence of location, weather conditions, varieties, and spectral sensor types.

Using three different data sets (Experiments 1–3), the study aims, at the canopy level for wheat, (i) to compare different statistical measures of goodness by which to evaluate the model performance, (ii) to examine the limitations of the spectral detection of N uptake, and (iii) to assess the influence of the a priori determined N levels on the coefficient of determination (R^2). In addition, Experiment 1 was used (iv) to compare commonly used indices of passive multispectral sensors (UAV) with ground-based hyperspectral sensors of a mobile ground-based platform, (v) to determine the sensitivity of indices frequently used for detecting N uptake taking into account the developmental stage, and (vi) to assess the importance of N content and DW for the detection of N uptake using frequently applied indices.

2. Materials and methods

2.1. Field trials

Field experiment 1 was conducted at the Dürnast research station of the Technical University of Munich in Germany (48°23′60″ N, 11°41′60″ E). The soil of the experimental field is described in Table 1. The different soil classifications within the field are reflected in the wide range of the apparent electrical conductivity values, as measured by an EM38 device (Heil and Schmidhalter, 2003). The average annual temperature is 8 °C, and average precipitation is approximately 800 mm..

Winter wheat (*Triticum aestivum* L., variety *Apostel*) was sown in mid-October 2019 at a rate of 350 kernels per square meter. The preceding crop was spring barley. The field was managed conventionally and plant protection activities were in line with local standards. Potassium and phosphorus were supplied by fertilization as part of the crop rotation. The nitrate-N content of the 0–60 cm soil layer was measured at the beginning of vegetation in the experimental year and amounted to 22 kg ha^{-1}.

The experiment had a randomized complete block design with four replicates (Fig. 1). Each block contained all N levels (n = 10), and each N

Table 1
Description of field site, showing elevation, apparent soil electrical conductivity (EC_a), and soil classification, with a predominant soil texture according to FAO et al. (1998).

Field	Elevation [m]	EC_a [mS m^{-1}]	Soil classification
D4	457–470	20–63	Cambisol: Silty clay loam Silty loam Loam Sandy loam Skeletic Cambisol Cumulic Cambisol Cumulic Anthrosol

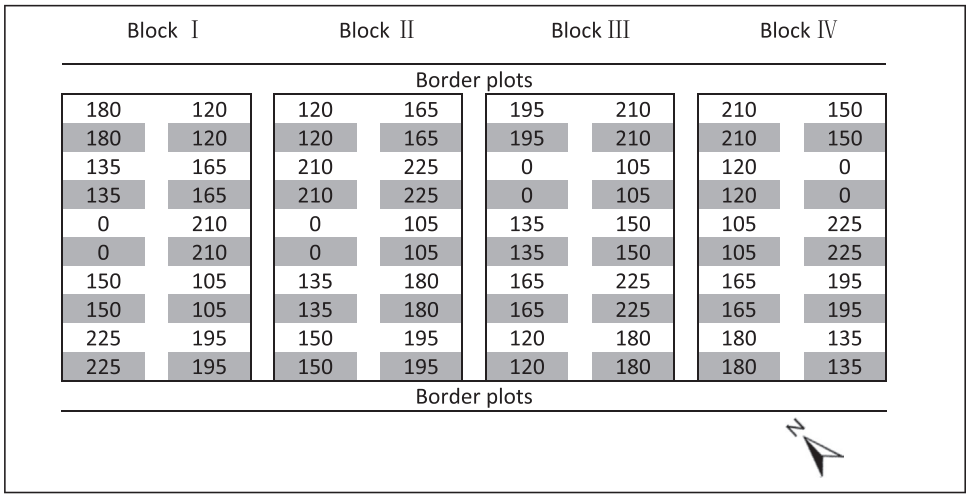

Fig. 1. Design of the field trial. The numbers indicate the fertilized quantity of nitrogen in kg N ha^{-1}, whereby each N-fertilization level per block (n = 4) always consists of two adjacent plots. White cells indicate the plots for biomass sampling while grey cells indicate those for combine harvesting and spectral measurements.

level treatment consisted of two neighboring plots, one for combine harvest and spectral measurements and one for biomass sampling. The two types of plots were located directly alongside each other. Each plot measured 10.5 m in length and 1.50 m in width.

Table 2 shows the N-fertilizer treatments. The absolute N-fertilization quantity for each N level was divided into three equally split applications using calcium ammonium nitrate (27% N). The timing was chosen both to enhance tillering and, to increase the number of spikelets and the grain protein content. The range of the absolute fertilized N levels (105–225 kg N ha^{-1}) roughly reflects that of N-fertilization in agricultural practice for yield expectations typical in large parts of Europe. The variant without N-fertilization (0 kg N ha^{-1}) provides additional information on the site and year-specific nitrogen supply of the experimental field. The fertilizer quantity could be precisely dosed and distributed on the plot in fine gradations thanks to the exact machine adjustments that are possible with a box spreader (Fiona® G-85, Denmark).

Experiment 2 was performed at the same site as Experiment 1. Table 3 indicates the data sets used. Extreme N levels were present in these experiments. To evaluate the influence of high N-fertilization, we divided the data sets in this study for further analysis. One set included all the N levels (0–60–120–180–240–300–360–420 kg N ha^{-1}), while the other only included the site-specific, agronomically relevant N levels (0–60–120–180–240 kg N ha^{-1}). A detailed description of the Experiment 2 can be found in Prey and Schmidhalter (2019) and Li et al. (2012).

Field Experiment 3 was conducted at the Roggenstein research station of the Technical University of Munich in Germany (48°10′44,9″ N, 11°19′14,5″ E). In 2017, data were collected for the RGT Reform variety, which was fertilized with calcium ammonium nitrate. The range of N levels was 0–245 kg N ha^{-1}. A detailed description of the Experiment 3 can be found in Westermeier and Maidl (2019).

Table 2
Quantity distribution and growing stage of N-fertilization. At the beginning of vegetation (VB) the wheat was at the tillering stage of development. The main developmental stage for BBCH 32 was stem elongation, while for BBCH 49 it was booting.

Fertilizer treatments	VB	BBCH 32	BBCH 49	Total
1	0	0	0	0
2	35	35	35	105
3	40	40	40	120
4	45	45	45	135
5	50	50	50	150
6	55	55	55	165
7	60	60	60	180
8	65	65	65	195
9	70	70	70	210
10	75	75	75	225

Table 3
Data sets used in Experiment 2.

Year	Variety	N fertilizer
2009	Nongda318	Urea
2016	Diskus, Rumor	Ammonium nitrate
2018	Diskus, Rumor	Ammonium nitrate

2.2. Spectral measurements

In Experiment 1, spectral aerial- (UAV) and ground-based measurements were conducted under cloud-free conditions on the day of the biomass harvest. Care was taken to ensure that the time offset between the use of the measurement platforms was as short as possible.

A Phenotrac IV vehicle sensor platform was used to perform ground-based measurements, with a hyperspectral bidirectional passive sensor spectrometer (tec5, Oberursel, Germany) mounted at the front center of the Phenotrac IV. The sensor had a nominal resolution of approximately 3.3 nm, a measuring range from 300 to 1000 nm, and a field of view (FOV) of 24°. The measuring distance to the canopy surface was approximately 0.8 m. The driving speed during measurement was approximately 5 km h^{-1}. A more detailed description of the Phenotrac IV can be found in Barmeier and Schmidhalter (2017).

Aerial multispectral sensing was performed with eBee X (BBCH 37) and ebee Plus (BBCH 61) fixed-wing aircraft (SenseFly®, Lausanne, Switzerland). Both UAVs were equipped with the same multispectral camera (Sequoia^{+} camera, Parrot, Paris, France), which recorded four spectral bands of the electromagnetic spectrum: green (550 nm, ~40 nm bandwidth), red (660 nm, ~40 nm bandwidth), red edge (735 nm, ~10 nm bandwidth), and NIR (790 nm, ~40 nm bandwidth). A white balance card was used to calibrate the reflectance. The flights were conducted at a height of 80 m above ground, resulting in ground resolutions of approximately 8 cm/pixel. This was the minimum possible flight altitude, as it was limited by the presence of power lines. The individual images were merged into a whole using Pix4D software (Pix4D S.A., Prilly, Switzerland). A detailed description of the UAV equipment can be found in Hu et al. (2020). A polygon for each plot was subsequently created for the complete image of each individual band using the ArcGIS program (ESRI®, Germany, Version 10.5.0.6491),

using the same measurement area and position as the terrestrial Phenotrac IV measurements for each plot. The mean value per polygon was then calculated for each individual band.

Indices were calculated from the reflectance data (Table 3). The indices for nitrogen traits in wheat were evaluated in previous studies at this site (Mistele and Schmidhalter, 2008, 2010; Erdle et al., 2011; Prey and Schmidhalter, 2019; De Souza et al., 2021) (Table 4).

Experiment 2, ground-based spectral measurements were performed with a Handy-Spec® field spectrometer in 2009 (Li et al., 2012) at BBCH 37, as well as with the Phenotrac IV mobile sensor platform in 2016 at BBCH 37, and in 2018 at BBCH 45 (Prey and Schmidhalter, 2019). In Experiment 3, a two-channel spectrometer was used at BBCH 32, 39, and 65 (Westermeier and Maidl, 2019). For a more detailed description of the sensors used, the reader is referred to the publications indicated above. Of the reflectance data obtained in Experiments 2 and 3, only the index REIP was used in this study. This index displayed a very good performance in previous publications compared with all the evaluated indices.

2.3. Destructive data collection and laboratory analyses

In Experiment 1, plant samples were taken from each biomass plot at indicative stages of development (BBCH 37 and 61) to assess the winter wheat nitrogen status. Four adjacent rows of plants at a mutual distance of 12.5 cm and with a length of 1.5 m (0.75 m^2) were manually sampled. The fresh biomass samples were oven-dried at 60 °C until no further water loss occurred and weighed to determine the DW, expressed as t ha^{-1}. The dried samples were then milled and sieved to 0.5 mm (Brabender®, Duisburg, Germany) for subsequent analysis in the laboratory to determine the N content (% of dry matter). This was done by mass spectrometry using an isotope ratio mass spectrometer with an ANCA SL 20–20 preparation unit (Europe Scientific, Crewe, UK). The N uptake of the total aboveground plant (kg N ha^{-1}) was calculated as DW x N content.

Samples were also taken at indicative developmental stages from the plots for destructive biomass data collection in Experiment 2 conducted in 2009 and 2016 at BBCH 37 and in 2018 at BBCH 45 and in Experiment 3 at BBCH 32, 39, and 65. The samples were then dried and analyzed to determine their N content. A detailed description of Experiment 2 is given in Prey and Schmidhalter (2019) and Li et al. (2012) as well as of Experiment 3 in Westermeier and Maidl (2019).

Table 4
List of indices and sensor platforms (UAV and Phenotrac IV vehicle-based sensor platform). The original bands were approximated depending on the technique used.

Index	Equation	Sensor platform	Reference
NDVI (NDVI_UAV and NDVI_Pheno)	$\frac{R790-R660}{R790+R660}$	UAV and Phenotrac IV	Rouse et al. (1974)
NIR/GREEN	R780/R550	UAV	Mistele and Schmidhalter (2008)
NIR/RED	R780/R670	UAV	Gitelson et al. (2003)
NIR/REDEDGE	R780/R735	UAV	De Souza et al. (2021)
NDRE	$\frac{R790-R720}{R790+R720}$	UAV	Barnes et al. (2000)
REIP	$700+40\left(\frac{\left(\frac{R670+R780}{2}\right)-R700}{R740-R700}\right)$	Phenotrac IV	Guyot et al. (1988)
R760/R730	R760/R730	Phenotrac IV	Erdle et al. (2011)
R780/R740	R780/R740	Phenotrac IV	Mistele and Schmidhalter (2010)

2.4. Statistics and calculations

The data were analyzed using Microsoft® Excel® 2019 MSO (16.0.14701.20240), and R Core Team (2021).

Based on the Akaike Information Criterion (AIC) (Webster and McBratney, 1989), either linear or polynomial (second-order) regressions were calculated between the N uptake (dependent variable = Y) and the indices and single bands (independent variable = X), respectively. The coefficient of determination (R^2) was calculated as a measure of the goodness of fit. R^2 shows the portion of the explained variance in the model with respect to the total variance (Eq. 1):

$$R^2 = \frac{explained\ \ deviation\ \ sum\ \ of\ \ squares}{total\ \ deviation\ \ sum\ \ of\ \ squares\ \ to\ \ be\ \ explained} = \frac{\sum_{i=1}^{n}(\hat{y}i-\bar{y})^2}{\sum_{i=1}^{n}(yi-\bar{y})^2} \quad (1)$$

where $\bar{y}$ is the arithmetic mean of all observed y_i and $\hat{y}_i$ is the estimator (regression function) of each observed y_i. Fig. 2 illustrates these relationships graphically.

As a further measure of the predictive quality of the regression, the confidence interval (95% level) for the estimated value (regression estimator) and the prediction interval (95% level) for the observed value are commonly used statistics. A detailed description can be found in Köhler et al. (2012) and Bleymüller et al. (2008). To obtain a more specific evaluation of the error of the entire model, the root mean square error (RMSE) was calculated as follows:

$$\mathrm{RMSE}\ [\mathrm{kg\ N\ ha^{-1}}] = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

where $\hat{y}_i$ are the predicted and y_i the measured values for the N uptake and n the number of samples. The advantage of RMSE is that it uses the same data unit as that of the variable to be explained. To enable better comparability of RMSE with other data sets, the RMSE values were standardized as a percentage and calculated as follows (modified according to Loague and Green, 1991):

$$\mathrm{RMSE}\ [\%] = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \times \frac{100}{\bar{O}} \quad (3)$$

where P_i and O_i are the predicted and observed values and n the number of samples. $\bar{O}$ represents the mean of the observed data. The RMSE values were then classified and evaluated according to Westermeier and Maidl (2019). RMSE values < 10% were considered excellent, 10–20% good and > 30% sufficient.

Another generally accepted error of the entire model is the mean absolute error (MAE), calculated as follows:

$$\mathrm{MAE}\ [\mathrm{kg\ N\ ha^{-1}}] = \frac{\sum_{i=1}^{n}|P_i - O_i|}{n} \quad (4)$$

where P_i and O_i are the predicted and observed values and n the number of samples. For a detailed description of RMSE and MAE, the reader is referred to Willmott (1984).

In addition to the commonly used statistical error measures for regressions, confidence intervals were calculated that tolerate a model error that is acceptable from an agronomical point of view. The magnitude of the error depends mainly on the N uptake of wheat, which

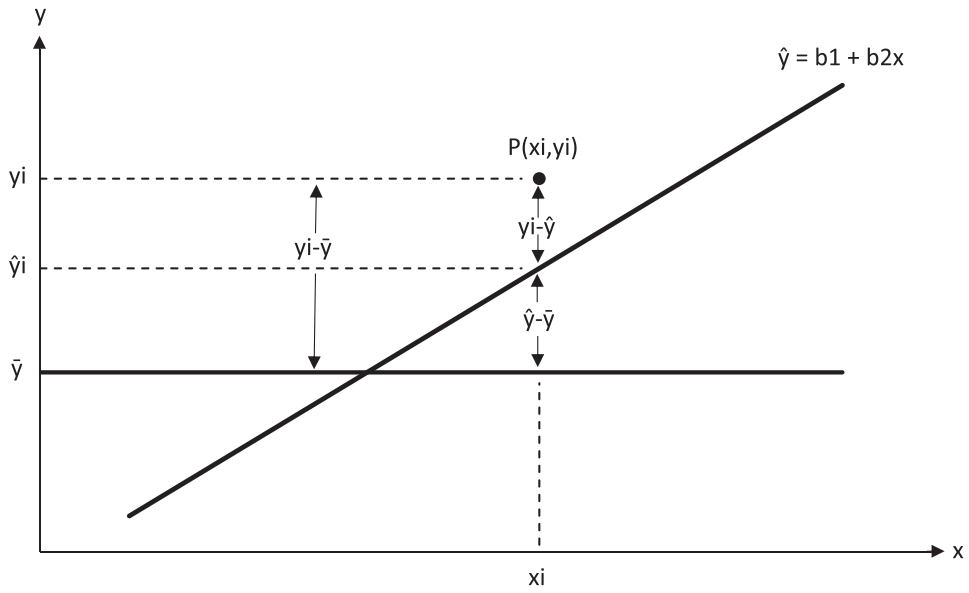

Fig. 2. Diagram showing the decomposition of the deviation sum of squares. x and y represent the independent and dependent variables with P as an example of a single measured value. $\bar{y}$ is the arithmetic mean of all observed y_i and $\hat{y}_i$ is the estimator (regression function) of each observed y_i (according to Bleymüller et al., 2008).

is primarily determined by the growth stage. Two growth stage intervals were defined: those at which N fertilization can still be performed (up to BBCH 50) and those that allow an assessment of the preceding N fertilization (after BBCH 50). Since the N uptake after BBCH 50 can reach values well above 200 kg ha^{-1}, a wider interval is tolerable, whereas up to BBCH 50, the error should be smaller, because the N uptake is much lower and further N fertilization decisions can be made based on the spectral detection. Identifying these limits was based on screening all available data sets with a subsequent post hoc analysis. This error is calculated using:

$$\text{agronomical error [kg N ha}^{-1}] = \hat{y}_i \pm \text{error of } y_i \quad (5)$$

where $\hat{y}_i$ is the estimator of the regression function and the errors of y_i used were assumed to be 10 and 15 kg ha^{-1} N uptake.

Correlations according to Pearson (r) were calculated between the N content and DW and all spectral indices and bands used. Next, the influence of the N content on the correlation of DW with all spectral indices and bands was determined. This was done by eliminating the influence of the N content using the partial correlation coefficient of $x1$ and $x2$ given ξ which is calculated as follows:

$$r_{x1,\ x2|\xi} = \frac{r_{x1,x2} - r_{x1,\xi} \cdot r_{x2,\xi}}{\sqrt{(1 - r_{x1,\xi}^2) \cdot (1 - r_{x2,\xi}^2)}} \quad (6)$$

where $r_{x1,x2}$ are the correlation coefficients between $x1$ and $x2$, $r_{x1,\xi}$ are those between $x1$ and ξ and $r_{x2,\xi}$ are those between $x2$ and ξ (Backhaus et al., 2006). If the N content does not influence the correlation of DW and all spectral indices and bands, the partial correlation coefficient is at the same level as the Pearson correlation coefficient (without the covariate). Increasing influence is indicated through a lower partial correlation coefficient. The difference between the Pearson and the partial correlation coefficient is equal to the influence of the covariate (N content).

The Nitrogen Nutrition Index (NNI) was calculated according to Lemaire and Gastal (1997) to characterize the nitrogen supply of wheat at each date*treatment combination, and is expressed as:

$$\text{NNI} = \frac{\text{Nact}}{\text{Nc}} \quad (7)$$

where N_{act} (%) is the actual measured nitrogen content and N_c (%) the critical nitrogen content of the biomass shoot dry matter. For N_c, Justes et al. (1997) developed the following estimation equation:

$N_c = 5.35$ (shoot dry weight [t ha^{-1}]$^{-0.442}$).

based on investigations of winter wheat when the shoot dry weight was in the range 1.55–12 t ha^{-1} and growth stages ranged from BBCH 30–65. NNI values above the line indicate no limitation of biomass production as there was sufficient N content, whereas NNI values below the line indicate a biomass limitation due to insufficient N content.

3. Results

3.1. Weather conditions and plant development

The following weather and plant growth conditions refer to Experiment 1. The weather in 2020 displayed some significant deviations from the long-term average (mean 1981–2010) in some months, which had an effect on the wheat growth (Fig. 3). The temperature was above the long-term average between January and April inclusive and also in August. The largest difference was observed in February and April. A precipitation deficit was observed in January, March to May inclusive, and July, while in February, June, and August, there was an increased amount of precipitation. The most pronounced differences were in February, May, and June (CDC, 2020).

At BBCH 37, the DW, N content, and NNI values were in the ranges 1.9–3.1 t ha^{-1}, 1.5–2.7%, and 0.4–0.7, respectively. The aboveground N uptake ranged from 28.2 to 70.3 kg N ha^{-1}. At BBCH 61, the DW, N content, and NNI values were in the ranges 4.7–8.0 t ha^{-1}, 1.0–2.2%, and 0.6–1.7, respectively while the N uptake ranged from 47.2 to 170.9 kg N ha^{-1}.

N-fertilization increased the N content at BBCH 37 and 61, but only increased DW at BBCH 61. The NNI showed only a moderate increase at BBCH 37 across the N levels, whereas at BBCH 61, a more significant increase in the NNI could be observed with increasing N-fertilization (data not shown).

3.2. Spectral detection of N uptake using different statistical measures of goodness, and the influence of the N levels

Using data from all three experiments (1−3), Table 5 presents an evaluation of the regression models using selected statistical measures of

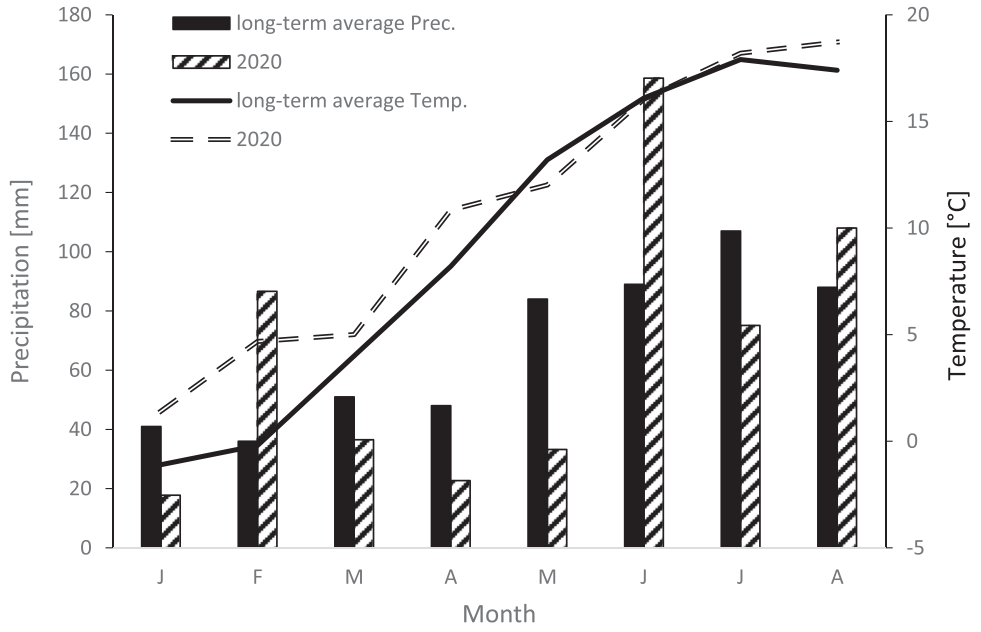

Fig. 3. Monthly weather conditions (January to August) at the experiment site in 2020 compared to the long-term average (1981–2010). The temperature (Temp.) is shown as lines and the precipitation (Prec.) as bars (CDC, 2020).

Table 5
Comparison of statistical measures of goodness relating to spectral recording of N uptake [kg N ha^{-1}] for wheat (Experiments 1–3) at different stages of development. Data sets labeled [a] include the N levels 0–240 kg N ha^{-1} with 60 kg increments. Data sets labeled [b] include N levels 0–420 kg N ha^{-1} in 60 kg increments. $\hat{y}_i$ is the estimator of the polynomial (poly) and linear (lin) regression function at the point x_i. Asterisks indicate the significance level (* ≙ $p < 0.05$, ** ≙ $p < 0.01$ and *** ≙ $p < 0.001$).

Experiment	Year	Variety	BBCH	Index	Type of Regression	R^2	RMSE [kg N ha^{-1}]	RMSE [%]	MAE [kg N ha^{-1}]	Proportion of data points (%) outside the agronomically relevant interval [$\hat{y}_i \pm 10$ kg N uptake ha^{-1}]	Proportion of data points (%) outside the agronomically relevant interval [$\hat{y}_i \pm 15$ kg N uptake ha^{-1}]
1[a]	2020	Apostel	37	REIP	poly	0.52 * **	6.6	13.1	5.2	17.9	2.6
1[a]	2020	Apostel	61	REIP	poly	0.83 * **	13.5	12.7	10.5	42.5	25.0
2[a]	2009	Nongda318	37	REIP	lin	0.44 * **	9.9	26.2	8.1	33.3	20.0
2[a]	2016	Diskus	37	REIP	poly	0.94 * **	8.4	11.8	6.3	20.0	15.0
2[a]	2016	Rumor	37	REIP	poly	0.95 * **	7.9	10.5	5.9	30.0	10.0
2[a]	2018	Diskus	45	REIP	poly	0.71 * **	12.5	16.5	10.8	53.3	33.3
2[a]	2018	Rumor	45	REIP	poly	0.72 * **	8.1	12.7	6.7	28.6	0.0
3[a]	2017	Reform	32	REIP	lin	0.80 * **	5.3	12.2	4.5	0.0	0.0
3[a]	2017	Reform	39	REIP	poly	0.95 * **	10.9	7.1	8.6	35.0	17.5
3[a]	2017	Reform	65	REIP	poly	0.85 * **	23.7	12.3	18.3	59.4	48.4
2[b]	2009	Nongda318	37	REIP	poly	0.51 * **	12.7	26.9	11.2	58.3	20.8
2[b]	2016	Diskus	37	REIP	poly	0.93 * **	11.7	12.3	9.1	43.8	18.8
2[b]	2016	Rumor	37	REIP	poly	0.91 * **	13.4	13.2	10.0	32.3	25.8
2[b]	2018	Diskus	45	REIP	lin	0.68 * **	14.1	16.3	12.3	54.2	33.3
2[b]	2018	Rumor	45	REIP	lin	0.58 * **	13.9	18.6	11.0	45.8	25.0

goodness. Polynomial regression models predominated for all data sets. Linear regression models were only observed at early growth stages. Across all, the models were highly significant. The values of R^2, RMSE, MAE, and data points outside the agronomically relevant intervals increased with increasing growth stages. However, regardless of this, the MAE was always lower than the RMSE. For the agronomically relevant intervals, the proportion of the data points within the interval increased significantly as the interval size increased ($\hat{y}i \pm 10$ vs. ± 15 kg N uptake ha^{-1}). The proportion of data points falling outside the ±10 kg N uptake ha^{-1} interval was on average approximately 30% for developmental stages up to BBCH 50 and approximately 50% from BBCH 50 onwards. In contrast, the proportion of data points outside the ± 15 kg N uptake ha^{-1} interval for developmental stages up to BBCH 50 was on average approximately 12% and from BBCH 50 onward 37%. In addition, the number of data points outside the interval limits below 10 kg N uptake ha^{-1} increased sharply (data not shown). High R^2 values did not imply a low proportion of data points falling outside the agronomically relevant interval. The greater influence in this context was due to the stage of development.

In addition, data from extreme N levels (300–360–420 kg N ha^{-1}) were included in the analysis (indicated by [b]). This led in part to changes in the different types of regression models. R^2 was slightly lower (except in one case), while the RMSE and MAE values were slightly higher. The proportion of data points outside the agronomically relevant interval ($\hat{y}i \pm 10$ kg N uptake ha^{-1}) also increased. This was also true for the wider interval ($\hat{y}i \pm 15$ kg N uptake ha^{-1}) except with one data set, which remained at the same level.

Based on three example data sets, Fig. 4 illustrates the different intervals around the fitted regression line. The width of the interval was smallest for the confidence interval (95%), increased for the

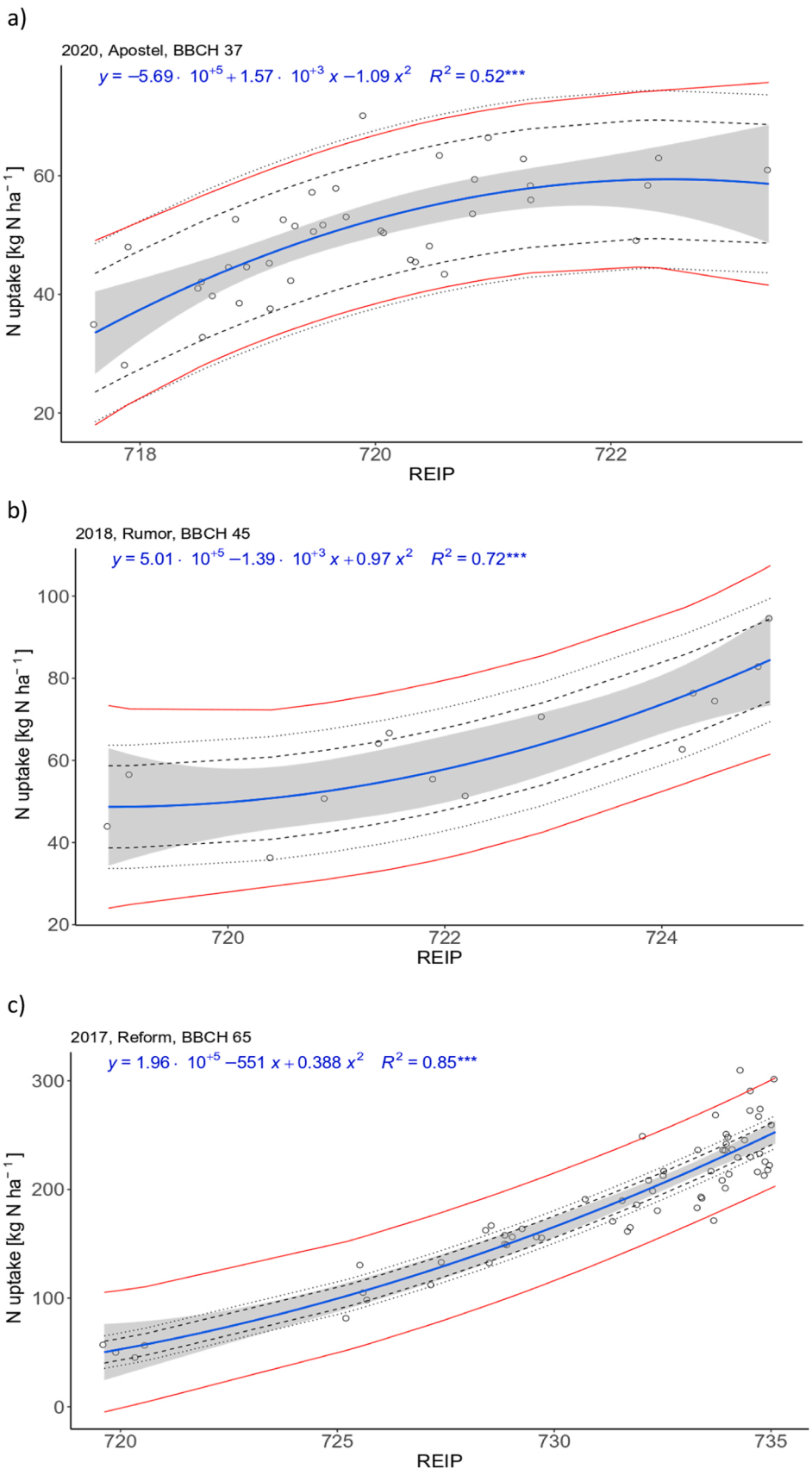

Fig. 4. Three example data sets, showing different stages of development (a-c), are used to compare different intervals to detect the N uptake [kg N ha^{-1}] using the index REIP with polynomial regression analysis. Asterisks indicate the significance level (*** ≙ $p < 0.001$). The grey areas show the confidence intervals (95%). The red lines show the limits of the prediction intervals (95%). The dashed ($\hat{y}i \pm 10$ kg N uptake ha^{-1}) and dotted ($\hat{y}i \pm 15$ kg N uptake ha^{-1}) lines show the limits of the agronomically relevant intervals. The year of the experiment, cultivar and growth stage are indicated at the top of the figures.

agronomically relevant intervals (depending on the specified error), and was largest for the prediction interval (95%).

3.3. Comparison of sensor platforms and spectral indices

Table 6 compares the suitability of the indices recorded by the UAV and the Phenotrac IV carrier vehicle platform for the N uptake in wheat at two developmental stages. The AIC criterion resulted in more polynomial regressions for the Phenotrac IV and more linear regressions for the UAV at BBCH 37, while linear regressions predominated at BBCH 61 for both sensor platforms. Highly significant models were overall observed. The highest R^2 and the lowest RMSE and MAE values were obtained at BBCH 37 by the NIR/REDEDGE index of the UAV and at BBCH 61 by the REIP index of the Phenotrac IV. At BBCH 37, the NDRE index of the UAV showed the lowest proportion of data points outside the agronomically relevant interval ($\hat{y}_i \pm 10$ kg N uptake ha^{-1}), while at BBCH 61, the Phenotrac IV indices REIP, R760/R730, and R780/R740 performed best. At BBCH 37, all indices of the UAV showed no data points outside the agronomically relevant interval ($\hat{y}i \pm 15$ kg N uptake ha^{-1}), while at BBCH 61, the Phenotrac IV index R780/R740 performed best.

3.4. Influence of N content and DW on spectral indices and bands

Table 7 shows the Pearson correlation coefficients between selected target traits and spectral indices and bands on the two measurement dates for Experiment 1. At BBCH 37, the strongest correlations for N content were observed across all indices and bands (highly significant) with a minimum and maximum *r* of 0.42 and 0.77, respectively. A weak correlation for DW was observed with a minimum and maximum *r* of 0.14 and 0.23, respectively. The correlations for N content and DW were positive for the indices and negative for the individual bands, with the exception of the NIR band.

The strength of the correlations generally increased for both traits at BBCH 61, and all were highly significant. Across all spectral indices and bands, the N content correlated more strongly with the spectral indices than DW, with a minimum and maximum *r* of 0.74 and 0.93, respectively. For DW, minimum and maximum *r* values of − 0.52 and 0.70 were observed. The alternation of positive and negative correlations was as in BBCH 37.

The use of N content as a covariate at BBCH 37 only resulted in a slight reduction in the *r* between DW and the spectral indices and bands. The minimum and maximum reductions of the correlation coefficient were 0% and 2%, respectively. A stronger influence of N content was observed at BBCH 61. On this measurement date, the minimum and maximum reduction of the *r* were 33% and 49%, respectively. Comparing the spectral indices and bands used, no significant difference could be observed at BBCH 37 whereas at BBCH 61 the deviations were slightly increased for the index NDVI and individual bands in the visible and red edge wavelength range.

4. Discussion

4.1. Weather conditions and plant development

In Experiment 1, wheat was in the tillering stage at the end of March 2020, as is typical for this site. Abundant precipitation in February resulted in cool soil temperatures and allowed only late passage of the field, leading to a relatively late initial N-application. Subsequently, the weather was sunny with strong cold winds. These weather conditions caused unproductive losses in the soil water and slowed early wheat development in spring. The subsequent level of rainfall, which was

Table 6
Comparison of sensor platforms and indices relating to spectral detection of N uptake [kg N ha^{-1}] in wheat (Experiment 1) at different developmental stages using different statistical measures of goodness. The best values of each statistic at each stage of development are highlighted. $\hat{y}_i$ is the estimator of the polynomial (poly) and linear (lin) regression function at the point x_i. Asterisks indicate the significance level (* ≙ $p < 0.05$, ** ≙ $p < 0.01$ and *** ≙ $p < 0.001$).

Sensor platform	BBCH	Index	Type of Regression	R^2	RMSE [kg N ha^{-1}]	RMSE [%]	MAE [kg N ha^{-1}]	Proportion of data points (%) outside the agronomically relevant interval [$\hat{y}_i \pm 10$ kg N uptake ha^{-1}]	Proportion of data points (%) outside the agronomically relevant interval [$\hat{y}i \pm 15$ kg N uptake ha^{-1}]
Phenotrac IV	37	NDVI	lin	0.31 * **	7.9	15.7	6.4	17.9	2.6
Phenotrac IV	37	REIP	poly	0.52 * **	6.6	13.1	5.2	17.9	2.6
Phenotrac IV	37	R760/R730	poly	0.46 * **	6.9	13.8	5.5	12.8	2.6
Phenotrac IV	37	R780/R740	poly	0.48 * **	6.8	13.6	5.4	10.3	2.6
UAV	37	NDRE	lin	0.57 * **	6.2	12.3	5.2	**10.0**	**0.0**
UAV	37	NDVI	lin	0.47 * **	6.9	13.7	5.6	17.5	**0.0**
UAV	37	NIR/GREEN	lin	0.46 * **	6.9	13.8	5.5	15.0	**0.0**
UAV	37	NIR/RED	lin	0.45 * **	7.0	14.0	6.0	15.0	**0.0**
UAV	37	NIR/REDEDGE	poly	**0.59 * ****	**6.1**	**12.1**	**5.0**	12.5	**0.0**
Phenotrac IV	61	NDVI	poly	0.71 * **	17.8	16.7	13.4	52.5	35.0
Phenotrac IV	61	REIP	poly	**0.83 * ****	**13.5**	**12.7**	**10.5**	**42.5**	25.0
Phenotrac IV	61	R760/R730	lin	0.81 * **	14.6	13.7	11.1	**42.5**	25.0
Phenotrac IV	61	R780/R740	lin	0.82 * **	14.2	13.3	10.9	**42.5**	**22.5**
UAV	61	NDRE	lin	0.75 * **	16.7	15.7	13.2	57.5	37.5
UAV	61	NDVI	poly	0.64 * **	19.8	18.6	15.5	62.5	37.5
UAV	61	NIR/GREEN	lin	0.70 * **	18.0	16.9	13.7	60.0	32.5
UAV	61	NIR/RED	lin	0.64 * **	20.0	18.8	15.6	57.5	37.5
UAV	61	NIR/REDEDGE	lin	0.75 * **	16.5	15.5	12.8	55.0	35.0

Table 7
Pearson correlation coefficients between selected target traits and spectral indices and bands (corresponding to UAV bands). Asterisks indicate the significance level (* ≙ $p < 0.05$, ** ≙ $p < 0.01$ and *** ≙ $p < 0.001$). [1]) Partial correlation with N content as a covariate.

Sensor platform	Spectral index/band	BBCH 37			BBCH 61		
		N content	DW	DW[1])	N content	DW	DW[1])
Phenotrac IV	NDVI	0.59 * **	0.15	0.13	0.85 * **	0.63 * **	0.20
Phenotrac IV	REIP	0.70 * **	0.21	0.22	0.93 * **	0.69 * **	0.34 *
Phenotrac IV	R760/R730	0.65 * **	0.19	0.18	0.91 * **	0.70 * **	0.36 *
Phenotrac IV	R780/R740	0.67 * **	0.18	0.18	0.91 * **	0.70 * **	0.37 *
UAV	NDRE	0.77 * **	0.23	0.26	0.88 * **	0.67 * **	0.29
UAV	NDVI	0.71 * **	0.18	0.18	0.83 * **	0.58 * **	0.11
UAV	NIR/GREEN	0.69 * **	0.21	0.22	0.85 * **	0.65 * **	0.26
UAV	NIR/RED	0.69 * **	0.18	0.18	0.81 * **	0.61 * **	0.21
UAV	NIR/REDEDGE	0.77 * **	0.23	0.26	0.88 * **	0.68 * **	0.31
UAV	GREEN	-0.70 * **	-0.21	-0.22	-0.81 * **	-0.56 * **	-0.09
UAV	RED	-0.72 * **	-0.17	-0.17	-0.80 * **	-0.52 * **	-0.03
UAV	REDEDGE	-0.70 * **	-0.19	-0.19	-0.78 * **	-0.52 * **	-0.04
UAV	NIR	0.42 * *	0.14	0.12	0.74 * **	0.62 * **	0.29

significantly below average, was probably not sufficient for the N-fertilizer to have an effect in the first and second N-applications. In addition, the residual mineral nitrogen level was low. Harmsen (1984) pointed out that if top dressed nitrogen is not moved into deeper soil layers and, the topsoil also dries out, the plant will not be able to use any potentially available nitrogen, which will lead to an N deficiency. This is first of all because the N concentration in the biomass decreases while the growth rate remains the same, and secondly, because the growth rate decreases and the accumulation of dry matter is reduced (Justes et al., 1997). In our case, the existing biomass was further reduced, evidenced by the reduction of secondary tillers. Averaged over a period of years, the biomass at this site was significantly lower in 2020 (data not shown). Above-average rainfall in June, however, allowed subsequent N uptake and favored the grain filling phase.

At BBCH 37, the previous two N-applications raised the N content but not the DW. Despite an increase in N content due to N-fertilization, NNI values < 1 showed a severe limitation in biomass production due to N deficiency (data not shown). Variations in N content led to variations in N uptake. The generally weak expression of the treatment was due to the low precipitation at this site, as indicated previously. Other authors have pointed out that the effect of N-fertilization on N uptake varies from year to year (López-Bellido and López-Bellido, 2001; Delogu et al., 1998).

The effects of N-fertilization, especially in the first two applications, were demonstrated by Maidl et al. (1998) at a comparable site for wheat. N-fertilization at the beginning of vegetation had a positive effect on the number of fertile tillers per m^2, while N-fertilization during stem elongation had a positive effect on the number of grains per ear. Small differences in biomass observed at BBCH 37 are more likely attributable to soil heterogeneities than to N-fertilization. Crain et al. (2013) observed coefficients of variation of 15–40% for DW in transect measurements at the end of tillering and at early stem elongation. Soil variability is given as the main reason for this variability, which in turn affects the water holding capacity and plant nutrient supply, as well as other variables affecting plant growth. Especially with limiting soil moisture, small variations in soil texture and soil cause increasingly large effects on plot-to-plot variability (Ceccarelli and Grando, 1996). At BBCH 61, a differentiation was observed in both N content and DW. At most N levels, NNI values were > 1, indicating no N deficiency (data not shown).

4.2. Statistical measures of goodness that describe the measurement quality for the spectral detection of N uptake and the influence of factor levels

Many previous studies use regressions to describe the relationship between spectral indices and the N uptake of wheat (Erdle et al., 2011; Li et al., 2012; Prey and Schmidhalter, 2019; Westermeier and Maidl, 2019). The most commonly used statistical measure of goodness is R^2, while others, such as RMSE and MAE, also describe the measurement quality of the regression.

The R^2 level depends, among many other influencing factors, on the differentiation of the N uptake. Differentiation can occur early on (cf. Table 5, 2016 Rumor and Diskus, BBCH 37), depending on the effect of the N fertilizer, but is usually more evident at later developmental stages, where higher R^2-values are frequently observed. This is probably because the proportion of the explained deviation sum of squares increases more than the total deviation sum of squares to be explained. Even if evaluations refer to the same developmental stage, additional statistical measures with the same unit as the target trait should be used. While the effect of MAE is to average the absolute deviations, RMSE averages the squared deviations. By squaring the errors, measured values that are further away from the regression equation are weighted more heavily, resulting in RMSE values being larger than MAE values. However, both average the model error, which led to an under- as well as overestimations of measured values. In this study, we attempted to define error limits from an agronomical perspective. The selected error limits ($\hat{y}i \pm 10$ and ± 15 kg N uptake ha^{-1}) are within an acceptable range, including other potential sources of error in N-fertilization, such as the amount applied, distribution accuracy, and small-scale crop variability. Error limits also depend on the development stage and the site-specific yield potential. Assessing a model against such limits can be done by calculating the data points that are within or out of range, this provides the dataset-based probability of capturing the N uptake within that error range. The model is larger for the agronomically relevant interval than for the RMSE and MAE measures. This suggests that the residuals are scattered over a wide range, leading to a more positive representation of the model error for RMSE and MAE values than agronomically relevant limits.

The influence of extreme N levels on the statistics was also investigated. Only data sets collected at a relatively early stage of development were available for this purpose. As indicated previously, further differentiation, particularly of high N levels, was not possible, which led to rather lower R^2 values but increased RMSE, MAE, and the proportion of data points outside the agronomical limits. Evaluating the models based on these statistics distorts the quality of the spectral detection of N uptake from an agronomical and application-oriented point of view.

4.3. Consideration of agronomical aspects in the spectral assessment of N uptake

In this study, regression function errors were defined to assess the spectral detection of the N uptake ($\hat{y}i \pm 10$ and ± 15 kg N uptake ha^{-1}). These interval limits depend on the site-specific N uptake in conjunction with subsequent nitrogen fertilization and the respective development stage (Fig. 4). Accordingly, high-yielding sites with N-fertilization

adjusted for high N uptake allow wider interval limits, as do late developmental stages.

N-fertilization applications in winter wheat can be done up to BBCH 50. If they are based on spectral information, the detection error should be as small as possible. However, if the chosen error limits are too small, the probability of the measured values lying outside these intervals increases significantly. This trade-off should be considered with respect to the given data set. We assume that a probability of at least 80% of the data points will be within the interval, resulting in a ± 15 kg N uptake ha^{-1} interval being the smallest one feasible in this study. At the early stages of stem elongation, a ± 10 kg N uptake ha^{-1} interval may also be a reasonable choice (cf. Table 5, 2017 Reform, BBCH 32). However, this needs to be further evaluated. At later development stages, when N-fertilization is no longer possible, larger probability and/or error limits are possible in the spectral assessment of N-fertilization. In addition, within the chosen developmental stages, higher aboveground biomass did not result in a higher proportion of data points outside the agronomic error limits. Recommendations for judging the quality or limits of the spectral assessment of N-fertilization should not be based solely on R^2.

Although the indices used are suitable for the spectral detection of N uptake, an unexplained scatter was determined in the models, which can probably be attributed to three main factors. On the one hand, the vertical nitrogen distribution in the wheat plant has to be considered, however, this cannot be fully detected by spectral means, despite the available choice of single sensor angles. In this study, destructively obtained N content values were determined as the mean of the entire plant, which cannot be fully mimicked by spectral detection due to the vertical gradients in the nitrogen content of the plants (Li et al., 2013) and can therefore lead to differences. The reference area of the destructively collected data represents a further potential significant source of error, although fairly large areas were sampled in this study. Furthermore, the spectral value of the plot represents an average value, which is calculated from strongly scattered individual values.

4.4. Comparison of sensor platforms and spectral indices

To compare different sensor platforms and spectral sensors, a UAV equipped with a multispectral sensor and a ground-based vehicle (Phenotrac IV) with a hyperspectral sensor were employed. At BBCH 37, the UAV, and at BBCH 61, the Phenotrac IV platform were able to capture the N uptake of wheat comparatively slightly better with differences however being small.

The main difference between the sensors is their bandwidth. However, after calculating the NDVI of the Phenotrac IV spectral information with the same bandwidths as used for the UAV NDVI, no differences in sensitivity were observed (data not shown). Due to the low biomass and the resulting presumed influence of soil heterogeneities, the index SAVI (Huete, 1988) was also calculated from the hyperspectral data (Phenotrac IV). The SAVI delivered no improvement in sensitivity (data not shown) therefore we conclude that the measurements were not markedly affected by soil signals. With the Phenotrac's IV FOV of 24° and the measuring distance to the canopy of 0.8 m, both nadir and off-nadir recording of the canopy take place. Aparicio et al. (2004) observed an influence of off-nadir measurements on the spectral indices. This was due to the relatively higher reflectance in the visible range compared to the near-infrared wavelength range caused by the higher influence of stems. This particularly occurs in canopies with a high leaf area index (LAI). At low LAI, the reflection in the near-infrared wavelength range is more pronounced and leads to higher NDVI values for off-nadir measurements. At BBCH 37, slightly higher NDVI values from the Phenotrac IV were observed compared to the NDVI of the UAV (data not shown). Further, for indices calculated from near-infrared reflectance, polynomial regressions prevailed (Table 6). This indicates that in given cases index values were higher at the same N uptake. This is in line with the findings by Mistele and Schmidhalter (2010), showing that with off-nadir measurements, more biomass is found in the sensor`s field of view, and so the signal intensity increases. Gnyp et al. (2015) reported advantages of off-nadir compared to nadir measurements in detecting N uptake at early and late stages of wheat development. They also observed saturation effects for both measurement geometries. In contrast, no saturation effects were observed for any index in this study. This is rather exceptional at this site, but can be explained by the strong reduction in biomass production obtained in the experimental year. At the same site, Mistele et al. (2004) observed saturation effects for all indices in a two-year nitrogen increase experiment with wheat in detecting N uptake. This effect increased at later development stages. However, the authors pointed out that biomass was more important than the developmental stage. Zheng et al. (2018) also determined advantages of UAV over ground-based spectral measurements for pre-heading stages in detecting the N content of rice. The authors attribute such differences to the different measuring areas of the sensors. This, however, can be ruled out in this study, since the areas measured were comparable. In contrast, Gnyp et al. (2016) found no difference between UAV and ground-based wheat measurements in detecting N uptake within a single growing season.

Regardless of the sensor platform, indices combining the REDEDGE and NIR bands performed best at BBCH 37 and 61 across all given statistics. This is in line with other studies (Mistele and Schmidhalter, 2010; Erdle et al., 2011; Gnyp et al., 2016; Prey and Schmidhalter, 2019). In addition, testing other commonly used spectral indices showed no improvement in the assessment of the N uptake.

4.5. Influence of N content and DW on spectral reflectance and indices

Pearson correlation coefficients were calculated between the N content and DW traits and spectral indices and bands from data obtained in Experiment 1 (Table 7). The strength of the correlation indicates the spectral detectability of the traits contributing to N uptake.

Across the two measurement dates, the correlation coefficients of all indices and bands were constantly higher for N content than for DW. In contrast, Elsayed et al. (2018) observed much weaker correlation coefficients for N content at early and late tillering stages compared to DW. Lilienthal (2014) reported no relationship between spectral indices and N content, whereas Prey and Schmidhalter (2019) observed both significant and non-significant relationships for N content across years and developmental stages. At different developmental stages across two growing seasons, Erdle et al. (2011) found slightly higher coefficients of determination (R^2) across all indices for the assessment of DW in wheat compared to N content. However, R^2 was high for both traits, and N content always increased with increasing N-fertilization, although significant differences were observed between growing seasons. DW did not always increase with increasing N-fertilization, especially between the N levels 160–220 kg N ha^{-1}. The different results suggest that the spectral detection of N content and DW may vary depending on the location, year, and variety. The main cause seems to be the annual variation in plant growth due to differences in weather conditions, mainly characterized by variations in precipitation, height and distribution. In the year of investigation, 2020, limiting soil moisture resulted in a greater differentiation of the N content compared to DW, thus improving the spectral detectability of this trait.

Next, N content was used as a covariate to assess the influence on the correlation between DW and the spectral indices and bands (Table 7). At BBCH 37, the influence of the N content on DW was low, despite a relatively high correlation of the N content with the spectral indices and bands used. This can be explained by the generally low correlation of DW with the spectral indices and bands due to the lack of N-fertilization and, in turn, differentiation of this trait. At BBCH 61, the N content significantly influenced the correlation of all indices and bands with DW, particularly in the visible and red edge wavelength range. DW differed more at this developmental stage than at BBCH 37. The sensitivity of the just mentioned spectral ranges is in line with other studies

indicating the promoting influence of nitrogen on pigments which in turn influence light absorption (Thomas and Gausman, 1977; Fernandez et al., 1994; Gitelson et al., 2003; Lilienthal, 2014). However, the N content still showed a significant influence on N uptake, which was probably due to the effect of lower dilution caused by the generally weaker development of DW (Mistele and Schmidhalter, 2008). This assumption is further supported by the high NNI values (data not shown).

When biomass differentiation is insufficient, N content can be an important additional parameter. However, differentiating N content as a basic trait of N uptake does not appear sufficient for spectral differentiation of N uptake. Furthermore, it should be pointed out that N content and DW at the canopy level represent the sum of the individual plant organs. For DW, N content, and N uptake, the leaf is the organ of the plant that lends itself best to spectral detection (Prey et al., 2020).

5. Conclusion

The results of this study suggest, that from an agronomical perspective it is not sufficient to use only R^2 as a statistical measure for judging the quality of spectral N uptake assessments, as done in nitrogen fertilization experiments, and so it needs to be augmented with additional information. The R^2 values are influenced by the differentiation in N uptake, which is in turn influenced by both the variation in N-fertilization and the developmental stage, which, again, may affect the interpretation. Intrinsic confidence intervals defined as error limits based on agronomical considerations improve the validity of regression models. Up to BBCH 50, where N-fertilization actions can be varied, a $\pm$ 15 kg N uptake ha^{-1} error limit was deemed applicable to the evaluation of multiple experiments over several years with different sensors. In addition, extreme N levels in field experiments may bias regression models. UAV-based multispectral as well as hyperspectral vehicle-based reflectance were equally suitable in detecting the N uptake of wheat. Regardless of the sensor platform, indices combining REDEDGE and NIR bands performed best. For the spectral detection of the N uptake, the differentiation of biomass was more crucial than the differentiation of the N content. Besides common statistical measures in the evaluation of reflectance-based detection of the N uptake in wheat, agronomically-supported detection limits should be further considered.

CRediT authorship contribution statement

Paul Heinemann and Urs Schmidhalter conceived and designed the experiment. Paul Heinemann conducted the experiment, analyzed the data, wrote the original draft. Stephan Haug carried out the statistical supervision. Paul Heinemann, Stephan Haug and Urs Schmidhalter reviewed and edited the manuscript. Urs Schmidhalter was responsible for project administration and funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported as part of the GreenWindows4_0 project by funds of the Federal Ministry of Food and Agriculture 516 (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the 517 Federal Office for Agriculture and Food (BLE) under the innovation support program.

References

Aasen, H., Bolten, A., 2018. Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers—from theory to application. Remote Sens. Environ. 205, 374–389. https://doi.org/10.1016/j.rse.2017.10.043.

Ali, M.M., Al-Ani, A., Eamus, D., Tan, D.K.Y., 2017. Leaf nitrogen determination using non-destructive techniques- A review. J. Plant Nutr. 40 (7), 928–953. https://doi.org/10.1080/01904167.2016.1143954.

Bayerische Landesanstalt für Landwirtschaft (LfL), 2018, Leitfaden für die Düngung von Acker- und Grünlandflächen (Gelbes Heft).

Backhaus, K., Erichson, B., Plinke, W., Weiber, R., 2006. Multivariate Analysemethoden: Eine anwendungsorientierte Einführung. Springer,, Auflage, p. 11.

FAO, ISRIC, and ISSS, 1998. World reference base for soil resources. In: FAO (ed.) World Soil Resources Rep, 84. FAO,, Rome, Italy.

Climate Data Center (CDC), 2020, Retrieved from ⟨https://www.dwd.de/DE/klimaumwelt/cdc/cdc_node.html⟩.

Aparicio, N., Villegas, D., Royo, C., Casadesus, J., Araus, J.L., 2004. Effect of sensor view angle on the assessment of agronomic traits by ground level hyper-spectral reflectance measurements in durum wheat under contrasting mediterranean conditions. Int. J. Remote Sens. 25 (6), 1131–1152. https://doi.org/10.1080/0143116031000116967.

Barmeier, G., Schmidhalter, U., 2017. High-throughput field phenotyping of leaves, leaf sheaths, culms and ears of spring barley cultivars at anthesis and dough ripeness. Front. Plant Sci. 8, 1920. https://doi.org/10.3389/fpls.2017.01920.

Barmeier, G., Hofer, K., Schmidhalter, U., 2017. Mid-season prediction of grain yield and protein content of spring barley cultivars using high-throughput spectral sensing. Eur. J. Agron. Vol. 90, 108–116. https://doi.org/10.1016/j.eja.2017.07.005.

Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R.J., Li, H., Moran, M.S., 2000. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Coincident Detection of Crop Water Stress, Nitrogen Status and Canopy Density Using Ground Based Multispectral Data. Proc. of the Fifth Int. Conf. on Precis. Agric., Bloomington, MN, USA., pp. 1–15

Bleymüller, J., Gehlert, G., Gülicher, H., 2008. Statistik für Wirtschaftswissenschaftler. Franz Vahlen 15 (Auflage).

Ceccarelli, S., Grando, S., 1996. Drought as a challenge for the plant breeder. Plant Growth Regul. 20, 149–155. https://doi.org/10.1007/BF00024011.

Chen, Z., Miao, Y., Lu, J., Zhou, L., Li, Y., Zhang, H., Lou, W., Zhang, Z., Kusnierek, K., Liu, C., 2019. In-season diagnosis of winter wheat nitrogen status in smallholder farmer fields across a village using unmanned aerial vehicle-based remote sensing. Agron 9, 619. https://doi.org/10.3390/agronomy9100619.

Crain, J.L., Waldschmidt, K.M., Raun, W.R., 2013. Small-scale spatial variability in winter wheat production. Commun. Soil Sci. Plant Anal. 44 (19), 2830–2838. https://doi.org/10.1080/00103624.2013.812735.

Darwinkel, A., 1983. Ear formation and grain yield of winter wheat as affected by time of nitrogen supply. Neth. J. Agric. Sci. 31, 211–225. https://doi.org/10.18174/njas.v31i3.16944.

De Souza, R., Buchhart, C., Heil, K., Plass, J., Padilla, F.M., Schmidhalter, U., 2021. Effect of time of day and sky conditions on different vegetation indices calculated from active and passive sensors and images taken from UAV. Remote Sens. 13, 1691. https://doi.org/10.3390/rs13091691.

Delogu, G., Cattivelli, L., Pecchioni, N., De Falcis, D., Maggiore, T., Stanca, A., 1998. Uptake and agronomic efficiency of nitrogen in winter barley and winter wheat. Eur. J. Agron. 9 (1), 11–20. https://doi.org/10.1016/s1161-0301(98)00019-7.

Ding, J., Ding, Y., Li, F., Wu, P., Zhu, M., Li, C., Zhu, X., Guo, W., 2021. Promoting pre-anthesis nitrogen accumulation in wheat to achieve high yield and nitrogen-use efficiency through agronomic measures. J. Plant Nutr. 1–13. https://doi.org/10.1080/01904167.2021.1918716.

Elsayed, S., Barmeier, G., Schmidhalter, U., 2018. Passive reflectance sensing and digital image analysis allows for assessing the biomass and nitrogen status of wheat in early and late tillering stages. Front. Plant Sci. 9. https://doi.org/10.3389/fpls.2018.01478.

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 Crops Res. 124, 74–84. https://doi.org/10.1016/j.fcr.2011.06.007.

Feng, W., Guo, B.B., Zhang, H.Y., He, L., Zhang, Y.S., Wang, Y.H., Zhu, Y.J., Guo, T.C., 2015. Remote estimation of above ground nitrogen uptake during vegetative growth in winter wheat using hyperspectral red-edge ratio data. Field Crops Res. 180, 197–206. https://doi.org/10.1016/j.fcr.2015.05.020.

Fernandez, S., Vidal, D., Simón, E., Solé-Sugrañes, L., 1994. Radiometric characteristics of Triticum aestivum cv, Astral under water and nitrogen stress. Int. J. Remote Sens. 15 (9), 1867–1884. https://doi.org/10.1080/01431169408954213.

Fitzgerald, G., Rodriguez, D., O`Leary, G., 2010. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index-The canopy chlorophyll content index (CCCI). Field Crops Res 116, 318–324. https://doi.org/10.1016/j.fcr.2010.01.010.

Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., Matinelli, L.A., Seitzinger, S.P., Sutton, M.A., 2008. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science Vol. 320, 889–892. https://doi.org/10.1126/science.1136674.

Gitelson, A.A., Viña, A., Arkebauer, T.J., Rundquist, D.C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies. Geophys. Res. Lett. 30 (5), 1248. https://doi.org/10.1029/2002GL016450.

Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J. Plant. Phys 160 (3), 271–282. https://doi.org/10.1078/0176-1617-00887.

Gnyp, M.L., Panitzki, M., Reusch, S., 2015. Proximal nitrogen sensing by off-nadir and nadir measurements in winter wheat canopy. Proceedings of the 10th Europ. Conf. on Prec. Agric. Wageningen Academic Publishers,, Wageningen, pp. 43–50.
Gnyp, M.L., Panitzki, M., Reusch, S., Jasper, J., Bolten, A., Bareth, G., 2016. Comparison between tractor-based and UAV-based spectrometer measurements in winter wheat. Proc. 13th Intern. Conf. Prec. Agric., St. Louis, Mo., USA, 31 July-3 August 2016.
Guo, B.B., Zhu, Y.J., Feng, W., He, L., Wu, Y.P., Zhou, Y., Ren, X.X., Ma, Y., 2018. Remotely estimating aerial N uptake in winter wheat using red-edge area index from multi-angular hyperspectral data. Front. Plant Sci. 9, 675. https://doi.org/10.3389/fpls.2018.00675.
Guyot, G., Baret, F., Major, D.J., 1988. High spectral resolution: determination of spectral shifts between the red and the near infrared. Int. Arch. of Photogramm. Remote Sens 11, 750–760.
Harmsen, K., 1984. Nitrogen fertilizer use in rainfed agriculture. Fertil. Res. 5, 371–382. https://doi.org/10.1007/BF01049117.
Heil, K., Schmidhalter, U., 2003. Derivation of soil texture and soil water content from electromagnetic induction measurements. Program book of the joint conference. European Conf. on Precision Farming-European Conf. on Precision Livestock Farming,, Berlin, pp. 429–430.
Heinemann, P., Schmidhalter, U., 2021. Simplifying residual nitrogen (N_{min}) sampling strategies and crop response. Eur. J. Agron. 130, 126369 https://doi.org/10.1016/j.eja.2021.126369.
Hu, Y., Knapp, S., Schmidhalter, U., 2020. Advancing high-throughput phenotyping of wheat in early selection cycles. Remote Sens 12, 574. https://doi.org/10.3390/rs12030574.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.
Justes, E., Jeuffroy, M.H., Mary, B., 1997. Wheat, barley, and durum wheat. In: Lemaire, G. (Ed.), Diagnosis of the Nitrogen Status in Crops. Springer-Verlag, Berlin, Heidelberg, pp. 73–92. https://doi.org/10.1007/978-3-642-60684-7.
Köhler, W., Schachtel, G., Voleske, P., 2012. Biostatistik: Eine Einführung für Biologen und Agrarwissenschaftler. Springe Spektrum 5.
Ladha, J.K., Pathak, H., Krupnik, T.J., Six, J., van Kessel, C., 2005. Efficiency of fertilizer nitrogen in cereal production: retrospects and prospects. Adv. Agron. Vol. 87, 85–156. https://doi.org/10.1016/S0065-2113(05)87003-8.
Lemaire, G., Gastal, F., 1997. N uptake and distribution in plant canopies. In: Lemaire, G. (Ed.), Diagnosis of the Nitrogen Status in Crops. Springer-Verlag, Berlin, Heidelberg, pp. 3–43. https://doi.org/10.1007/978-3-642-60684-7.
Li, F., Mistele, B., Hu, Y., Yue, X., Yue, S., Miao, Y., Chen, X., Cui, Z., Meng, Q., Schmidhalter, U., 2012. Remotely estimating aerial N status of phenologically differing winter wheat cultivars grown in contrasting climatic and geographic zones in China and Germany. Field Crops Res. 138, 21–32. https://doi.org/10.1016/j.fcr.2012.09.002.
Li, F., Mistele, B., Hu, Y., Chen, X., Schmidhalter, U., 2014. Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. Eur. J. Agron. 52, 198–209. https://doi.org/10.1016/j.eja.2013.09.006.
Li, H., Zhao, C., Huang, W., Yang, G., 2013. Non-uniform vertical nitrogen distribution within plant canopy and its estimation by remote sensing: A review. Field Crops Res 142, 75–84. https://doi.org/10.1016/j.fcr.2012.11.017.
Lilienthal, H., 2014. Optical sensors in agriculture: principles and concepts. J. für Kult. 66 (2), 34–41. https://doi.org/10.5073/JFK.2014.02.01.
Loague, K., Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. J. Contam. Hydrol. 7 (1–2), 51–73. https://doi.org/10.1016/0169-7722(91)90038-3.
López-Bellido, R.J., López-Bellido, L., 2001. Efficiency of nitrogen in wheat under mediterranean conditions: effect of tillage, crop rotation and N fertilization. Field Crops Res 71, 31–46. https://doi.org/10.1016/S0378-4290(01)00146-0.
Maidl, F.-X., Sticksel, E., Retzer, F., Fischbeck, G., 1998. Effect of varied N-fertilization on yield formation of winter wheat under particular consideration of mainstems and tillers. J. Agron. Crop Sci. 180, 15–22. https://doi.org/10.1111/j.1439-037X.1998.tb00363.x.
Meier U. (2018). Growth stages of mono- and dicotyledonous plants: BBCH Monograph. Julius Kühn-Institut (JKI).
Mistele, B., Schmidhalter, U., 2008. Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. Eur. J. Agron. 29, 184–190. https://doi.org/10.1016/j.eja.2008.05.007.
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. https://doi.org/10.2134/agronj2009.0282.
Mistele, B., Gutser, R., Schmidhalter, U., 2004. Validation of field-scaled spectral measurements of the nitrogen status in winter wheat. In: Mulla, D. (Ed.), 7th Int. Conf. Prec. Agric. Other Prec. Res. Manag., Minneap., Minn., USA 1187–1195.
Olfs, H.-W., Blankenau, K., Brentrup, F., Jasper, J., Link, A., Lammel, J., 2005. Soil- and plant-based nitrogen-fertilizer recommendations in arable farming. J. Plant Nutr. Soil Sci. 168, 414–431. https://doi.org/10.1002/jpln.200520526.
Prey, L., Schmidhalter, U., 2019. Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. Sensors 19, 3712. https://doi.org/10.3390/s19173712.
Prey, L., Kipp, S., Hu, Y., Schmidhalter, U., 2019a. Nitrogen use efficiency and carbon traits of high-yielding european hybrid vs. line winter wheat cultivars: potentials and limitations. Front. Plant Sci. 9, 1988. https://doi.org/10.3389/fpls.2018.01988.
Prey, L., Germer, M., Schmidhalter, U., 2019b. Temporal and organ-specific responses in NUE traits to N fertilization, fungicide intensity and early sowing in winter wheat cultivars. Agron 9 (6), 313. https://doi.org/10.3390/agronomy9060313.
Prey, L., Hu, Y., Schmidhalter, U., 2020. High-Throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: optimizing the selection of vegetation indices and growth stages. Front. Plant Sci. 10, 1672. https://doi.org/10.3389/fpls.2019.01672.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (URL). ⟨https://www.R-project.org/⟩.
Robertson, G.P., Vitousek, P.M., 2009. Nitrogen in agriculture: balancing the cost of an essential resource. Annu. Rev. Environ. Resour. 34, 97–125. https://doi.org/10.1146/annurev.environ.032108.105046.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., Harlan, J.C., 1974. Monitoring the Vernal Advancement of Retrogradation of Natural Vegetation. Report No. E75-10354. NASA,, Washington, DC.
Swarbreck, S.M., Wang, M., Wang, Y., Kindred, D., Sylvester-Bradley, R., Shi, W., Varinderpal-Singh, V., Bentley, A.R., Griffiths, H., 2019. A roadmap for lowering crop nitrogen requirement. Trends Plant Sci. Vol. 24 (Issue 10), 892–904. https://doi.org/10.1016/j.tplants.2019.06.006.
Thomas, J.R., Gausman, H.W., 1977. Leaf reflectance vs. leaf chlorophyll and carotenoid concentrations for eight crops. Agron. J. 69 (5), 799–802. https://doi.org/10.2134/agronj1977.00021962006900050017x.
Webster, R., McBratney, A.B., 1989. On the Akaike information criterion for choosing models for variograms of soil properties. J. Soil Sci. 40, 493–496. https://doi.org/10.1111/j.1365-2389.1989.tb01291.x.
Westermeier, M., Maidl, F.-X., 2019. Comparison of spectral indices to detect nitrogen uptake in winter wheat. J. für Kult. 71 (8–9), 238–248. https://doi.org/10.5073/JfK.2019.08-09.02.
Willmott, C.J., 1984. On the evaluation of model performance in physical geography. In *Spatial statistics and models*. Springer,, Dordrecht, pp. 443–460. https://doi.org/10.1007/978-94-017-3048-8_23.
Zhang, J., Wang, W., Krienke, B., Cao, Q., Zhu, Y., Cao, W., Liu, X., 2021. In-season variable rate nitrogen recommendation for wheat precision production supported by fixed-wing UAV imagery. Prec. Agric. https://doi.org/10.1007/s11119-021-09863-2.
Zheng, H., Cheng, T., Li, D., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2018. Combining unmanned aerial vehicle (UAV)-based multispectral imagery and ground-based hyperspectral data for plant nitrogen concentration estimation in rice. Front. Plant Sci. 9, 936. https://doi.org/10.3389/fpls.2018.00936.