An active-optical reflectance sensor in-field testing for the prediction of winter wheat harvest metrics

Marko Milan Kostić,[1] Nataša Ljubičić,[2] Vladimir Aćin,[3] Milan Mirosavljević,[3] Maša Budjen,[2] Miloš Rajković,[3] Nebojša Dedović[1]

[1]University of Novi Sad; [2]BioSense Institute, Novi Sad, Vojvodina; [3]Institute of Field and Vegetable Crops, Novi Sad, Vojvodina, Serbia

Abstract

The ambition of this study was to justify the possibility of wheat trait prediction using a normalized difference vegetation index (NDVI) from a newly developed Plant-O-Meter sensor. Acquired data from Plant-O-Meter was matched with GreenSeeker's, which was designated as a reference. The experiment was carried out in the field during the 2022 growing season at the long-term experimental field. The experimental design included five different winter wheat genotypes and 20 different NPK fertilizer treatments. The GreenSeeker sensor always gave out NDVI values that were higher than those of the Plant-O-Meter by, on average, 0.029 (6.36%). The Plant-O-Meter sensor recorded similar NDVI values (94% of the variation is explained, P<0.01). The Plant-O-Meter's NDVIs had a higher CV for different wheat varieties and different sensing dates. For almost all varieties, GreenSeeker exceeded Plant-O-Meter in predicting yields for the early (March 21st) and late (June 6th) growing seasons. $NDVI_{GreenSeeker}$ data improved yield modeling performance by an average of 5.1% when compared to $NDVI_{Plant\text{-}O\text{-}Meter}$; in terms of plant height prediction, $NDVI_{GreenSeeker}$ was 3% more accurate than $NDVI_{Plant\text{-}O\text{-}Meter}$ and no changes in spike length prediction were found. A compact, economical and user-friendly solution, the Plant-O-Meter, is straightforward to use in wheat breeding programs as well as mercantile wheat production.

Correspondence: Marko Milan Kostić, Agricultural Engineering, University of Novi Sad, Novi Sad, Serbia.
E-mail: markok@polj.uns.ac.rs

Key words: multispectral sensors; NDVI; modelling; proximal sensors; cereals.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: none.

Received: 10 April 2023.
Accepted: 12 September 2023.

Journal of Agricultural Engineering 2024; LV:1559
doi:10.4081/jae.2024.1559

Introduction

Wheat is traditionally one of the most grown crops worldwide, with an estimated production area of 220 million hectares in 2022 and an average yield of 3.5 Mg ha^{-1} (World Agricultural Production, 2022). Also, in the region of the Pannonian Plain, including Serbia, wheat is the most significant winter cereal crop, providing a notable alternative for other spring crops such as soybean, maize, or sunflower. In Serbia, total wheat production increased by 19.8% in 2021, compared to 2020 (Statistical yearbook of the Republic of Serbia, 2022), with the intention of continuing this in 2022 and maintaining the average yields. As a result of improvements in both plant breeding and crop management, grain yield has more than doubled in the Pannonian Plain and Serbia in recent decades. In the recent past, many aspects and reasons have been identified, including positive effects on weed control (particularly perennial species), soil structure remediation during crop rotation and increased profitability. Although wheat grain yields in Serbia are higher than the average global yield (ranked 32 out of 124), they are often limited by unfavorable climate conditions and lower than in other production regions, especially Western Europe (Jaćimović, 2012).

The last two years have seen a massive disturbance in the global food market caused by various circumstances (Ukraine crisis and COVID) initiating the price of wheat to rise rapidly ($189/ton in March 2020; $406.7/ton in February 2022, according to "Wheat Prices – 40 Year Historical Chart"). At the same time, according to the Fertilizers Price Index (n.d.), the price of nitrogen fertilizers, which are highly demanded by cereals in general due to the physiological role of N in plants, has almost tripled in the last two years. The emerging input/output ratio of wheat production (and other crops) may open a wide scope for the adoption of advanced technologies for better nitrogen (N) management and higher yield performance. Moreover, agriculture is experiencing a revolution triggered by emerging technologies that seem highly promising, as they will allow higher yields, quality and profitability. According to the results of Raun *et al*. (2002), the use of optical sensors as a diagnostic tool has improved the efficiency of in-season nitrogen application (nitrogen use efficiency) by 15%, which helped in recommending the optimal amount of nitrogen fertilizer. Achieving a better production economy through the implementation of precision agriculture (PA) is well recognized in developed countries, where the acceptance of advanced technologies far exceeds the technologies applied in Serbia. The prevailing decision-making system in open-sky farming in Serbia is still based on traditional patterns, personal non-objective "experience, intuition and habits from the past", despite the introduction of sophisticated mobile systems and equipment that have the improved ability to operate in a spatially predefined manner (not just uniformly). Such a stochastic way of managing a temporally dynamic and spatially heterogeneous field resource leads to results that are not compatible with the concept of PA, the Green Agenda, or the Nitrates Directive.

Plants are natural biosensors that indicate deficiencies in soil through changes in turgor and decreased chlorophyll pigment activity. Monitoring of soil-plant conditions provides critical information for improving resource utilization efficiency and developing a site-specific database of the relationship between soil conditions and plant growth for intelligent and sustainable agriculture systems (Yin *et al.*, 2021). The human eye perceives vegetative changes in the RGB (440-690 nm) wavelength range with insufficient sensitivity and the impossibility of impartial quantification of the condition; hence, a timely intervention is needed in the field. Active optical reflectance sensing has demonstrated effectiveness in particular areas for producing prescriptions for N fertilizer that increase N usage efficiency, but locally produced algorithms have not been evaluated concurrently across a wide region (Bean *et al.*, 2018). If crop yield could be predicted with a certain degree of accuracy, N fertilizer applications could be customized for each site, taking into account the crop's needs and the N status of the soil, to maximize potential yield (Crain *et al.*, 2012). The normalized difference vegetation index (NDVI), which measures the sharp distinction between photosynthetic light absorption in the visible portion of the spectrum and reflectance in the near-infrared region (NIR), is a well-known vegetation index (Raun *et al.*, 2002; Tagarakis *et al.*, 2022). Various spectral bands can be used in different ways to check the health of plants. The normalized simple ratio of two bands (NIR and RED) was found to be useful, easy to get and has a great potential for crop monitoring. Normalization has also been proposed to mitigate the effects of sensor deterioration (Bannari *et al.*, 1995). The algebraic expression of the level of photosynthetic activity in the range of discrete values (-1, 1) is applied to discretize the readings for the purposes of analysis and further use of NDVI in management systems. Throughout history, numerous proximal sensors have proven useful in agriculture. The early products, including the Soil Plant Analysis Development meter (Ang *et al.*, 2020) (Konica Minolta Inc., Osaka, Japan), had a few automations and required manual data entry. The options for mapping and real-time spatially variable rate fertilizer applications are now available with fully automated sensing systems like Yara N-sensor (Raper *et al.*, 2013) (Yara International ASA, Oslo, Norway), GreenSeeker (Xia *et al.*, 2016) (Trimble Inc., CA, USA), Crop Circle (Cao *et al.*, 2017) (Holland Scientific, NE, USA) and CropScan (Sankaran *et al.*, 2019) (CROPSCAN, Inc. Rochester, MN, USA), among others. On the other hand, the complexity of agriculture and limited ability of farmers to adopt sophisticated technologies impose a need for continuous development of rapid, cost-affordable and reliable data acquisition systems, followed by automated data processing and online data transfer. Thus, the primary objective of this study was to evaluate the performance of a new Plant-O-Meter active multispectral proximal sensor in a long-term wheat trial and to compare it to the GreenSeeker portable device, a commonly used commercial crop sensor. The specific objectives of this study were to: i) determine the relationship between NDVI measurements from the two sensors; ii) determine the specific growth stage at which the sensors provide a more reliable estimate of end-of-season yield under the specific climatic conditions of the Vojvodina region; iii) determine the ability of Plant-O-Meter to estimate end-of-season yield traits from mid-season canopy measurements, compared to the hand-held GreenSeeker sensor.

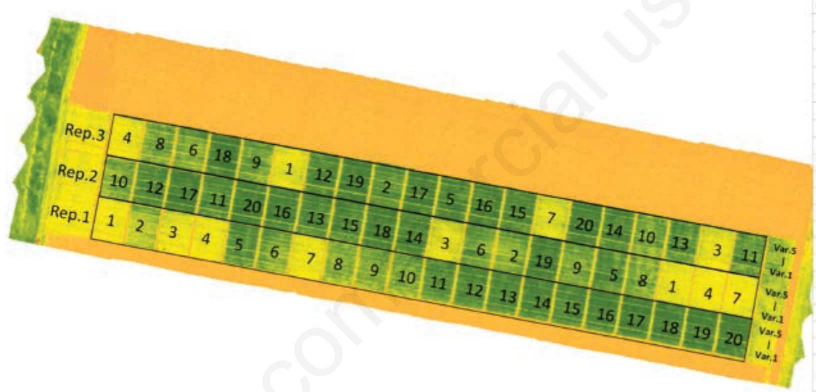
Materials and Methods

Field selection and experimental setup

A comparative test of sensors in real field conditions must be managed in such a way that the sensor's advantages can be objectively highlighted, while any uncertainty arising from either poor trial setup or operator bias are minimized. At present, there is no standardized in-field test procedure to verify the validity of optical reflectance sensor measurements. Due to the high soil organic matter content, crops in the on-farm trials on chernozem soil type respond poorly to the applied soil treatments, which could jeopardize the effect of the experimental block setup. The objective of field selection was to ensure that plant variability was attributed to the treatments of the experiment and not due to spatial variability in soil properties. Accordingly, sensor testing was conducted in the fields of the Institute of Field and Vegetable Crops at Rimski šančevi (Serbia), which are under long-term (Long-term Field Experiments in Europe, n.d.) consistent NPK treatment with a certain plot-to-plot discrepancy. The field trial (45.3326123°N, 19.8300244°E) was established during the 2021-2022 growing season. The soil type at the site is Haplic Chernozem Aric (WRB, 2014), which dominates in Vojvodina region and is characterized as highly fertile (approximately 43% of the total arable land, (Jovanović *et al.*, 2013). Five domestic high-yielding cultivars of winter wheat (*Triticum aestivum* L.), widely grown in Serbia, were included in the study: NS Igra, NS Rajna, NS Futura, NS Epoha and NS Obala. The varieties were chosen for their prospective high-yielding traits, the difference in plant height among varieties and other grain quality traits. NS Igra is a new high-yielding cultivar, characterized by a lower habitus and good disease resistance. NS Obala and NS Rajna are cultivars with awns, characterized by high yield stability. NS Futura is a cultivar with a higher plant height and premium quality. NS Epoha is an awnless wheat cultivar with a stable yield and high grain quality. The experiment was laid out as a split-plot design with fertilizer treatments as the main factors and winter wheat cultivars as sub-factors. Each experimental plot was considered as an independent block (replicated three times) that consisted of five wheat varieties in a consistent order with randomized fertilizer treatments (20 combinations of NPK fertilizer rates, 100 subplots, Figure 1). For practical reasons, the fertilizer treatments were coded from 1 to 20. The subplots were initially established at 3×13.5 m (3 m seeder width), from which 1.5×12.5 m was harvested (the combine harvester width was 1.5 m) to prevent the border effect. Given the limited resources, the design of the experiment was planned in view of the labor costs that could be allocated per unit of observed area, while gathering enough information to enable an objective sensor comparison. In order to guarantee uniform soil conditions across the plots, the soil preparation procedures comprised conventional tillage by means of a moldboard plow at a depth of 20 cm and a final soil consolidation using a combined cultivator. A 25-row wheat seeder with a double-disc furrow opener (Amazone D8-30 Super) was used for sowing in fall (October). It was calibrated to achieve a seeding rate of 500-550 germinated seeds per square meter. In the trials, the standard pest control strategy was applied: early insecticide application after plant emergence (Alfa-cipermetrin, 100 g L^{-1}); herbicide application in the early spring (metsulfuron metil, 600 g kg^{-1}) concurrent with the first fungicide treatment (Propikonazol, 250 g L^{-1}); and the final pesticide application for ear protection at the beginning of flowering (Tebukonazol, 133 g L^{-1} + Prohloraz, 267 g L^{-1}). Harvest was conducted at the beginning of July when the plants had attained full maturity and the grain moisture content was around 14%. In order to eliminate marginal effects, only the central part of the subplots was harvested to determine the yield of each variety using a small plot combine harvester (Witnersteiger Delta). The samples were weighed in the field and put in bags with labels before being processed.

During the experiment conducted on March 21st, April 6th and 18th, May 9th and 20th and June 6th, corresponding to the growth stages BBCH 22, 31, 37, 52, 73 and 77, respectively, NDVI was measured by means of two active proximal sensors: the pocket version GreenSeeker (Trimble Inc., CA, USA) and the improved version of the active multispectral optical sensor, Plant-O-Meter (BitGear, Serbia), whose prototype was introduced by Kitić *et al*. (2019). GreenSeeker, a self-illuminated sensor, emits light and measures reflectance at 660 nm (R) and 770 nm (NIR) and calculates NDVI. The capability of the pocket version of GreenSeeker was confirmed in the study by Crain *et al*. (2012) who did comprehensive testing on wheat and maize. The GreenSeeker sensor was held about 60 cm above the wheat canopy and parallel to the row direction. Measurements of in-field reflectance were manually done, by taking an average reading from the measurement area in each plot.

The Plant-O-Meter (Figure 2) is an active sensor that has a built-in multispectral source and light sources that emit light at six indicative wavelengths: 465 nm (blue), 535 nm (green), 630 nm (red), 740 nm (red edge) and 850 nm (NIR). This sensor detects the light reflected from the canopy of plants and provides raw data measurements for user-defined indices, that can be calculated based on its ability to independently record reflectance for each band. Plant-O-Meter establishes a connection with any Android smartphone, as well as it logs and processes data using that device's storage and processing resources. In addition, the data are georeferenced via the smartphone's GPS receiver. Also, it can be extended with a wired or wireless connection for different communication interfaces and protocols. The frequency of data acquisition was 1 Hz, which corresponds to approximately 10-15 measurements per plot. GreenSeeker was chosen as a reference sensor in this study, due to its proven capabilities as well as its widespread commercial and scientific applications (Tagarakis *et al.*, 2022). Although the sensors use the same principle for detection, there are some differences between them. As stated by Kitić *et al*. (2019), who carried out initial Plant-O-Meter tests, a broader surface may be covered by the Plant-O-Meter's elliptical-shaped beam, since it



Code	Fert. treat.
1	Ø
2	N_2
3	P_2
4	K_2
5	N_2P_2
6	N_2K_2
7	P_2K_2
8	$N_1P_1K_1$
9	$N_1P_2K_1$
10	$N_1P_2K_2$
11	$N_2P_1K_1$
12	$N_2P_2K_1$
13	$N_2P_2K_2$
14	$N_2P_3K_1$
15	$N_2P_3K_3$
16	$N_3P_1K_1$
17	$N_3P_2K_1$
18	$N_3P_2K_2$
19	$N_3P_3K_2$
20	$N_3P_3K_3$
1,2,3	50, 100, 150 kg NPK/ha

Figure 1. Unmanned aerial vehicle-derived normalized difference vegetation index of the experimental field on May 9th, 2022 (anthesis growth stage). Indices 1, 2 and 3 of the NPK combination refer to the content of specific nutrients (N, P and K) as follows: 50, 100 and 150 kg ha^{-1}.

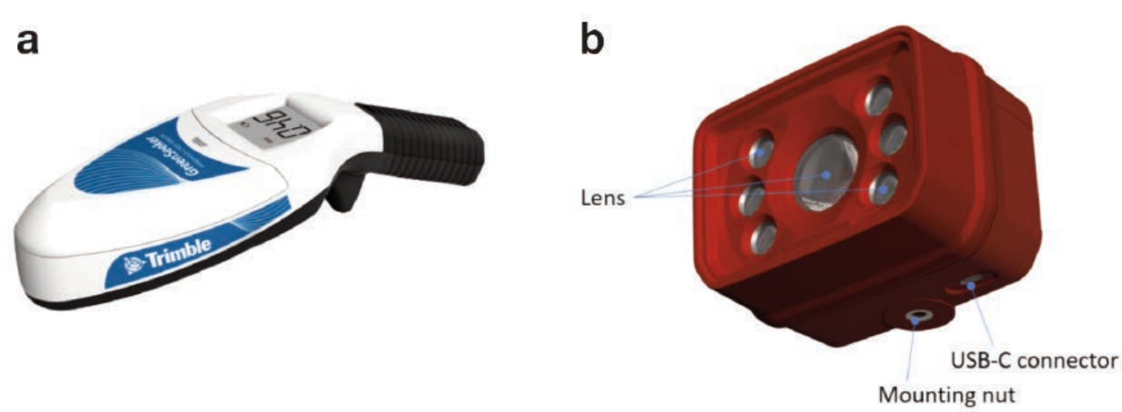


Figure 2. Handheld active proximal sensors used in the study: (**a**) GreenSeeker; (**b**) Plant-O-Meter (https://www.plant-o-meter.com/).

has a wider view angle than the GreenSeeker (72°). In addition, the GreenSeeker's minimum operating distance ranged from 60 to 90 cm, depending on the target's reflectance characteristics; targets with high reflectance required larger distances, due to saturation. The minimum operating distance for Plant-O-Meter was 50 cm, which is less than the GreenSeeker's minimum operating distance range of 60-90 cm.

GreenSeeker NDVI formula: Plant-O-Meter NDVI formula:

$$NDVI_{GreenSeeker}=\frac{NIR_{780}\text{-}RED_{660}}{NIR_{780}+RED_{660}} \qquad NDVI_{Plant\text{-}O\text{-}Meter}=\frac{NIR_{810}\text{-}RED_{630}}{NIR_{810}+RED_{630}}$$

The effects of plot treatments on the differences in soil physical condition were assessed by comparing the mean values of the observed soil properties. Basic statistical indicators were used in the following analysis. Duncan's test with a 95% level of confidence was applied to determine the statistically significant differences between the data. Pearson's coefficient (R) was used to determine the linearity between the observed parameters. The statistical analyses were performed using Statistica 12 software (Dell Software, TX, USA). In Statistica, a straightforward best-fit regression approach was used to model wheat traits by using NDVI data from each sensor during the trial and for each measurement.

The evaluation of the difference between sensor readings was done by calculating the relative Root Mean Square Error (RMSE):

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(P_i\text{-}\tilde{O}_i)^2}{n}}\cdot\frac{100}{\tilde{O}}, \tag{1}$$

where n is the number of observations, P_i is the predicted value, $\tilde{O}_i$ is the calculated value and $\tilde{O}$ is the calculated mean value. The best method was the one which had the lowest RMSE.

Results

Weather and climate data from the experimental site

The beginning of the wheat growing season (October-November 2021) was characterized by a relatively sufficient water supply for plant germination. According to the findings from previous studies in the Vojvodina region (Jaćimović, 2012), a highly positive correlation was obtained between the yield and the amount of rainfall during November. Weather conditions from the sowing phase until February are crucial for vernalization and water accumulation in the soil profile. A lower amount of precipitation from February to March indicated a lack of water in the soil. Additionally, around this time, the first post-emergence application of nitrogen fertilizer was carried out to produce adequate nutrient supplies for the purpose of constructing the organic matter of the plants. The beginning of stem elongation and fast growth occurs around the end of February. At this stage, any water scarcity reduces the yield potential. When compared to the historical climatic average, the average monthly temperatures for the 2021-2022 growth season were similar. Since the temporal weather regimes did not significantly deviate from historical records, the assumption was that a sensor's recordings reflected wheat development dynamics, which might correspond to most growing seasons, thus, the importance of the drawn deductions could be more universal. As a result, the significance of the conclusions drawn from these recordings may be more widely applicable in terms of the potential of sensor utilization in wheat trait predictions.

Overall statistics of wheat traits

Figure 3 represents the results of the analysis of the acquired data for each variety and fertilizer treatment. Relative values were calculated for each trait to provide simultaneous characterization. The data were put on a scale from 0 to 100% while preserving their relative relationships so that they could be easily compared and analyzed. This reduced bias in statistical analyses, as it ensured that different variables were given equal weight in the analysis. Considering the confidence intervals of the mean values of the evaluated traits, it is clear that the included factors (variety and fertilizer) had distinctive contributions to the variations of the

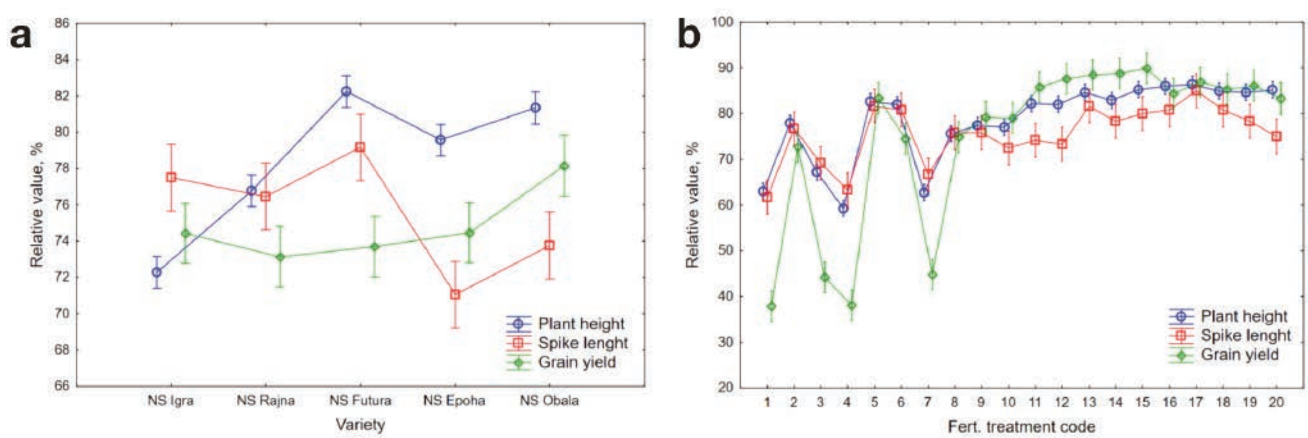


Figure 3. Effect of controlled factors: (**a**) variety; (**b**) fertilizer treatment. The error bars represent 95% confidence intervals.

observed wheat indicators. It can be concluded from Figure 3 that the variety as a controlled factor did not dominate in grain yield variability (4.69-5.03 t ha^{-1}). Only NS Obala provided a statistically significant difference in grain yield with respect to the others. The wheat variety had a more significant impact on plant height and spike length. The analysis of the impacts of fertilizer application revealed evident variations in wheat phenotypic characteristics. Table 1 gives an overview of the main statistical indicators that are already presented in Figure 3. The wheat variety had a notable impact on plant height (63.60-72.38 cm) and spike length (5.68-6.33 cm), while a mild impact was achieved in the case of grain yield, which ranged from 4.69 to 5.03 tha^{-1}. Based on ANOVA analysis, average values of plant height for included varieties were classified into four categories, with the highest mean value for NS Futura (72.38 cm) and statistically different values recorded for NS Rajna (67.57 cm) and NS Igra (63.60 cm), while

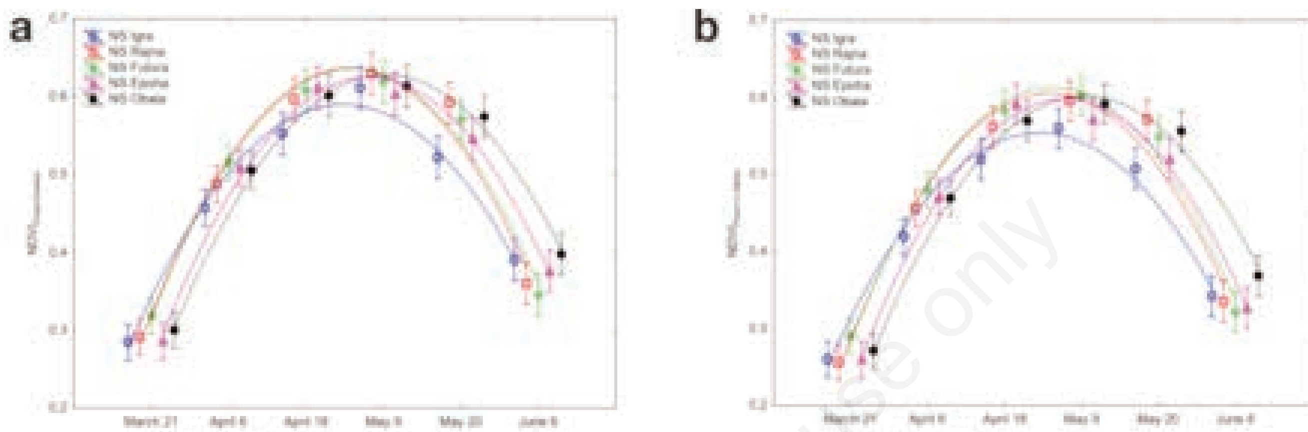


Figure 3. Effect of controlled factors: (**a**) variety; (**b**) fertilizer treatment. The error bars represent 95% confidence intervals.

Table 1. Descriptive statistics of plant indicators for the tested wheat varieties.

Variable	Variety	Valid N	Mean	95% conf. intervals		Min	Max	CV	Skew.	Kurt.
				Sample based data						
Grain yield, t ha^{-1}	NS Igra	60	4.78^{b}	4.49	5.06	1.62	6.14	26.75	-1.26	0.26
	NS Rajna	60	4.69^{b}	4.41	4.98	1.41	6.20	27.56	-1.08	-0.06
	NS Futura	60	4.73^{b}	4.45	5.00	1.43	6.20	26.18	-1.28	0.38
	NS Epoha	60	4.78^{b}	4.51	5.05	1.90	6.09	25.19	-1.19	0.03
	NS Obala	60	5.03^{a}	4.79	5.27	1.94	6.42	21.63	-1.52	1.68
Plant height, cm	NS Igra	60	63.60^{d}	62.42	64.78	47.00	72.00	8.36	-1.26	1.1^{2}
	NS Rajna	60	67.57^{c}	65.74	69.39	44.00	79.00	12.13	-1.04	0.5^{2}
	NS Futura	60	72.38^{a}	70.20	74.57	44.00	88.00	13.57	-1.11	0.38
	NS Epoha	60	70.02^{b}	68.06	71.97	44.00	82.00	12.56	-0.90	0.17
	NS Obala	60	71.58^{a}	69.61	73.56	46.00	83.00	12.40	-1.20	0.60
Spike length, cm	NS Igra	60	6.20^{a}	6.01	6.39	4.00	8.00	13.94	-0.45	0.24
	NS Rajna	60	6.12^{a}	5.93	6.30	4.00	8.00	13.79	0.21	0.73
	NS Futura	60	6.33^{a}	6.15	6.52	4.00	8.00	13.03	-0.31	-0.02
	NS Epoha	60	5.68^{b}	5.54	5.83	4.00	7.00	11.31	0.24	-0.22
	NS Obala	60	5.90^{c}	5.67	6.13	4.00	8.00	17.16	-0.12	-0.61
				Sensor based data						
$NDVI_{GreenSeeker}$	NS Igra	300	0.46^{c}	0.44	0.48	0.18	0.81	33.94	0.28	-1.06
	NS Rajna	300	0.48^{ab}	0.46	0.50	0.19	0.81	35.82	0.21	-1.40
	NS Futura	300	0.49^{a}	0.47	0.50	0.19	0.78	34.13	0.17	-1.40
	NS Epoha	300	0.48^{b}	0.46	0.49	0.20	0.77	33.18	0.06	-1.42
	NS Obala	300	0.49^{a}	0.47	0.50	0.24	0.75	30.24	0.01	-1.30
$NDVI_{Plant-O-Meter}$	NS Igra	300	0.43^{c}	0.41	0.44	0.15	0.79	34.72	0.28	-1.03
	NS Rajna	300	0.45^{ab}	0.44	0.47	0.18	0.78	38.22	0.23	-1.38
	NS Futura	300	0.46^{a}	0.45	0.48	0.17	0.77	36.29	0.20	-1.35
	NS Epoha	300	0.45^{b}	0.43	0.46	0.17	0.76	36.18	0.15	-1.36
	NS Obala	300	0.46^{a}	0.45	0.48	0.20	0.76	32.43	0.07	-1.28

Note: the mean values in the columns for the same letter do not differ from each other by the Duncan's test at 5% probability.

data for NS Epoha (70.02 cm) and NS Obala (71.58 cm) were similar to those of NS Rajna and NS Futura. The grain yield heterogeneity expressed, as the coefficient of variation (CV) ranged from 21.63 to 27.56%. Also, measured values for Futura showed the highest variation of plant height (13.57%), while Igra showed a reasonably lower variation (8.36%). Spike length differed based on the statistical values (Table 1), where the highest average value (6.33 cm) was obtained for the NS Futura variety and the highest CV was settled for the NS Obala variety (17.16%). According to Duncan's test, the spike length of the other varieties, with the exception of NS Igra, were statistically different, whereas NS Epoha had the smallest average value (5.68 cm).

Table 1 also shows the results of the sensor-based data analysis relative to the wheat varieties. The highest mean $NDVI_{GreenSeeker}$ was obtained for NS Futura and NS Obala (0.49). $NDVI_{GreenSeeker}$ for NS Rajna and NS Epoha (0.48) were slightly lower, falling into the same ANOVA group. A significantly lower average $NDVI_{GreenSeeker}$ value was acquired for NS Igra (0.46). An identical data grouping was obtained by ANOVA in the case of the $NDVI_{Plant\text{-}O\text{-}Meter}$, although with lower average values, ranging from 0.43 to 0.46. By comparing the CV values of the NDVI recorded with tested sensors, it is clear that they have a very similar range of values ($CV_{GreenSeeker}$=30-34%; $CV_{Plant\text{-}O\text{-}Meter}$=32-38%; with a bit higher variation of Plant-O-Meter NDVI values). The skewness and kurtosis are within the range (±2) proposed by Curran *et al*. (1996), which indicates normally distributed data.

Comparison of sensors' output

Figure 4 shows the NDVI values for wheat varieties and especially for each measurement date. The 95% confidence interval limits are indicated, based on which statistically significant differences can be recognized for individual measurement dates between the varieties. As Alvar-Beltrán *et al*. (2020) suggested, a polynomial regression was used to describe the general NDVI trend of each variable. Visual inspection of Figure 5 reveals that the NDVI shifts across the sensing dates for both sensors. The maximum average $NDVI_{GreenSeeker}$ values were obtained on May 9th (raising vegetative stage) for the NS Rajna variety (0.63), while NS Epoha had the lowest value (0.60). Plant-O-Meter provided different scales of NDVI on May 9th; the highest NDVI was obtained for NS Futura (0.60) and the lowest for NS Igra (0.56). The NDVI data set for both sensors generated on March 21st shows weak sensitivity for

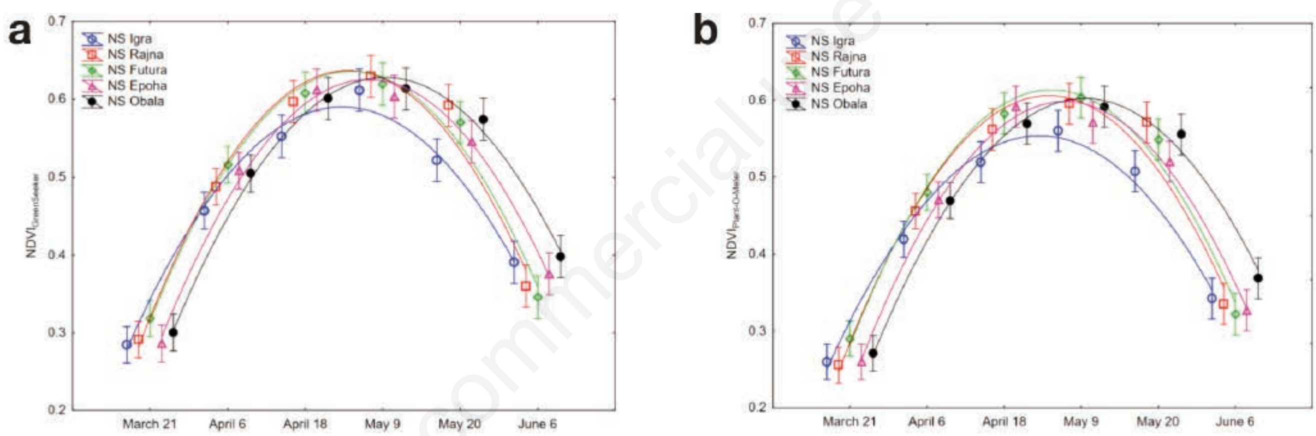


Figure 4. Normalized difference vegetation index average values (the vertical bars denote 0.95 confidence intervals) over the observation period for GreenSeeker (**a**) and Plant-O-Meter (**b**).

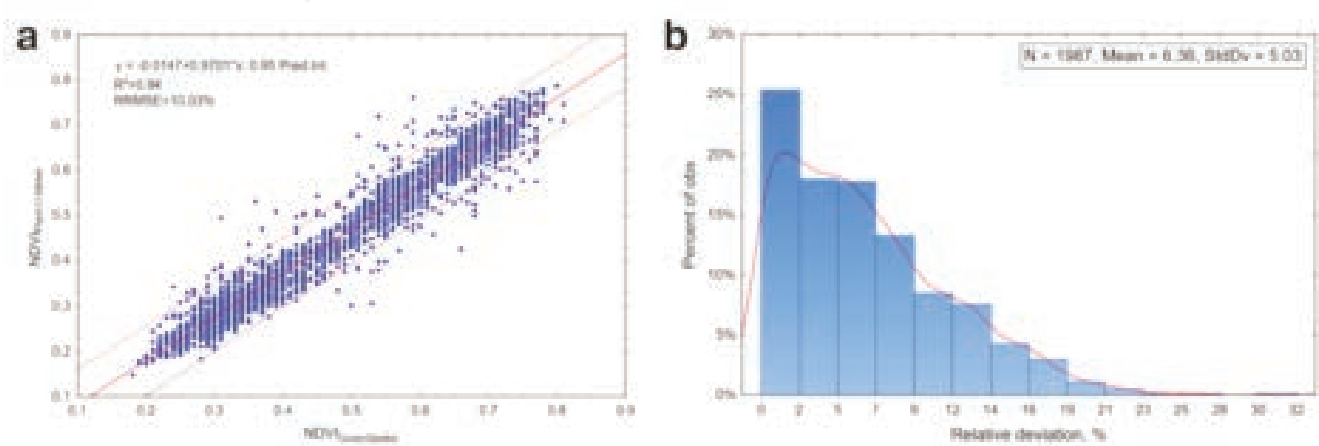


Figure 5. Relationship between the normalized difference vegetation index readings from the GreenSeeker and the Plant-O-Meter (**a**) and frequency distribution of the difference (**b**).

variety differentiation in accordance with confidence intervals for width and position. On April 6th, April 18th, May 20th and June 6th, there was a distinct variation in NDVI among different varieties. With a few exceptions, the data characteristics provided by the tested sensors were very similar. By observing Figure 4 and comparing mean NDVI values for sensing dates, Plant-O-Meter provided a little bit higher difference in average NDVI between varieties on May 9th, while GreenSeeker got similar advances on May 20th and June 6th. In general, the GreenSeeker sensor had systematically higher output values over the entire sensing period: they were on average 0.029. Expressed in absolute values, it was a difference of 0.02 to 0.03 obtained for March 21st. From the first sensing date to April 18th, the above mentioned difference decreased and, then, increased.

Figure 5 shows the relationship between the NDVI values of tested sensors by plotting entire data sets from different sensing dates on the graph. The Plant-O-Meter sensor readings collected over the observation period in winter wheat canopies consistently reproduced similar NDVI values compared to the GreenSeeker (94% of the variance is explained, $P<0.01$, RMSE = 10.03%). The variability in the data set was expected, because of the operator's inconsistent handling and the sensor's manufacturing quality (sensing element design, signal conditioning, sensor materials, embedded firmware, *etc*.). Moreover, the NDVI values from the GreenSeeker handheld sensor were 0.03 points higher than those from the Plant-O-Meter. Furthermore, Figure 5b depicts the difference between sensor readings with a frequency bar that illustrates the distribution of the relative disagreement of compared values. The average relative difference obtained for all sensing sessions is 6.36%, while the maximum is 21.49%. The majority of the values (90%) fall between 0 and 12%, with 5% being the most frequently calculated value, which is quite satisfactory. The results shown in Figure 5 suggest that there was a significant interaction between sensors. Given the experiment design, in which replications were considered as independent blocks and wheat varieties as independent factors with uncertain contributions to the sensors× readings, the $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ were compared for each wheat variety and sensing date, respectively (Figure 6). In order to model a trend in the stability of Pearson's R value over the sensing period, the correlations between sensors were plotted based on the performed measurements. Figure 6 clearly shows that there is a distinct divergence of R values along the replications, which proves the approach of data analysis. The correlations obtained for the first replication of the wheat trial show a very strong and consistent relationship among data from the used sensors (0.96), which is substantially different from the average R value drawn for the second and third replications (0.9 and 0.9, respectively), based on Duncan's test of significance. Concerning the wheat varieties, no significant difference was obtained but certain disagreements are evident. The smallest average R and widest confidence intervals were calculated for NS Obala (R = 0.88), whereas the highest average correlation was achieved for NS Futura (0.96) with the narrowest confidence intervals. The different correlation strengths seen among the data from the sensors were most likely caused by the fact that the NS Obala and NS Futura wheat varieties have different morphologies (NS Obala is an awned variety and NS Futura is an awnless variety). The R values from other varieties are the following: $R_{NS\ Igra}$, $R_{NS\ Rajna}$=0.92, $R_{NS\ Epoha}$=0.94. Sensing date

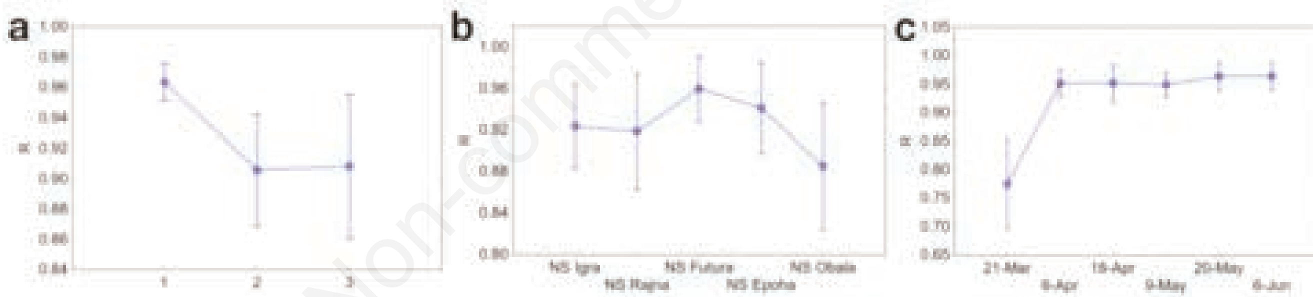


Figure 6. Correlations between $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ for replications (**a**), varieties (**b**) and sensing dates (**c**). The vertical bars denote 0.95 confidence intervals. NDVI, normalized difference vegetation index.

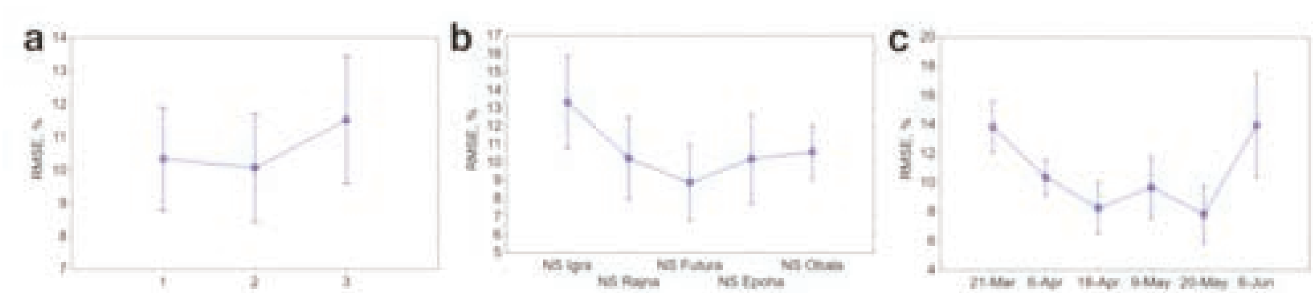


Figure 7. Comparison charts of root mean square error of linear models for replications (**a**), varieties (**b**) and sensing dates (**c**). The vertical bars denote 0.95 confidence intervals.

had a remarkable impact on the correlations between the sensors' data. The obvious lower synchronization of the sensor readings occurred on March 21st, yielding an average correlation of 0.73. Sensor data that were collected later provided a statistically reasonable and higher level of confidence (0.95-0.96 on average).

RMSE as a quality estimator of linear correlation between the observed NDVIs is shown in Figure 7. The influence of replications on the linear fitting error was not statistically significant and varied from 10.07 to 11.6% accordingly. The smallest deviation of residuals for linear models between $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ was achieved for variety NS Futura (8.87%), while the highest was recorded for NS Igra (13.25%), with an absence of statistical significance among relative RMSE values. Over the sensing dates, the highest RMSE was acquired on June 6th (13.95%), which is very close to the RMSE recorded on March 21st (13.82%). The lowest RMSE was reported for May 20th (7.79%). This implies that a lower relative RMSE was obtained for a higher NDVI, probably due to a smaller internal heterogeneity of subplots in terms of vegetative growth, with magnified differences among subplots (treatments).

The variability of the NDVI is expressed using the CV. This parameter is thought to be useful for understanding how well the sensor can predict yield traits. According to Raun *et al*. (2005), the optimal moment to sense and apply in-season N fertilizer is when the variability of the NDVI readings is the maximum one. They claimed that treating crops when they are at their highest variability of NDVI is expected to be the most effective method. Figure 8 depicts the mean CV values of NDVI from the tested sensors. Figure 8a shows the general trend of the CV over the trial period. From the obtained results, the highest CV values for both NDVI data sets were recorded on April 18th, during the stem elongation stage, when rapid growth occurs (Figure 8a). This is a time when farmers in Serbia usually perform the spring top-dress N application as an optional fine-tuning correction of Sharma *et al*. (2015) who discovered that the majority of sensor reflectance originates from the soil surface, since the leaf surface area determined by NADIR-aimed scans is minimum during the early development stages.

In total averages, NDVI data derived from Plant-O-Meter contributed to a higher CV (9.15%) with respect to the CV (7.04%) obtained for GreenSeeker×s NDVIs. The variability of NDVI from Plant-O-Meter was higher for all sensing dates (Figure 8a) and all tested wheat varieties (Figure 8b). When comparing the CV of NDVI by the sensing dates, the statistically significant influence of the sensor type on the CV of NDVI was recognized for March 21st ($CV_{GreenSeeker}$ = 6.17%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.93%) and June 6th ($CV_{GreenSeeker}$ = 6.25%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.4%). The lowest level of CV for both data sets appeared in a time frame between May 9th ($CV_{GreenSeeker}$ = 6.41%; $CV_{Plant\text{-}O\text{-}Meter}$ = 7.59%) and May 20th ($CV_{GreenSeeker}$ = 5.84%; $CV_{Plant\text{-}O\text{-}Meter}$ = 7.21%). In the last sensing session, the Plant-O-Meter showed significantly higher sensitivity. With reference to the variety, certain differences in terms of CV were observed. It is obvious from Figure 8b that Plant-O-Meter delivered NDVI made a stronger distinction between varieties in terms of CV. According to the confidence intervals, significantly different average CVs of $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ were observed for NS Igra ($CV_{GreenSeeker}$ = 8.27%; $CV_{Plant\text{-}O\text{-}Meter}$ = 11.35%), NS Rajna ($CV_{GreenSeeker}$ = 7.04%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.21%) and NS Epoha ($CV_{GreenSeeker}$ = 5.84%; $CV_{Plant\text{-}O\text{-}Meter}$ = 8.53%). Furthermore, NS Epoha had the lowest $CV_{GreenSeeker}$ (5.83%), while NS Futura had the lowest $CV_{Plant\text{-}O\text{-}Meter}$ average value (3.07%).

Sensor evaluation in wheat trait modeling

Tables 2-7 demonstrate the capacity of daily NDVI values to predict end-of-season harvest metrics across all experimental variants. In this section of the comparison, all the NDVI data that had been gathered were used. The results of regression analysis are shown in Tables 2-7, where the regression coefficient of determination (R^2) is calculated as an evaluator of model quality. The *Supplementary Tables 1-6* show all the regression plots with the fitted models that correspond to the results from Tables 2-7. The evaluation of a sensor-based yield prediction involved three different approaches, including raw reflectance data. In order to test the robustness of sensor readings and to segment the effects of the aforementioned factors on the quality of regression models derived from empirical data (20 fertilizer levels, 3 replications), the correlations were separately created for each variety and each sensing date. Furthermore, the R^2 of the regression models was determined by comparing the average NDVI from all dates for specific varieties and wheat traits evaluated in a single plot. In order to determine the effects of the date on the strength of correlations for each

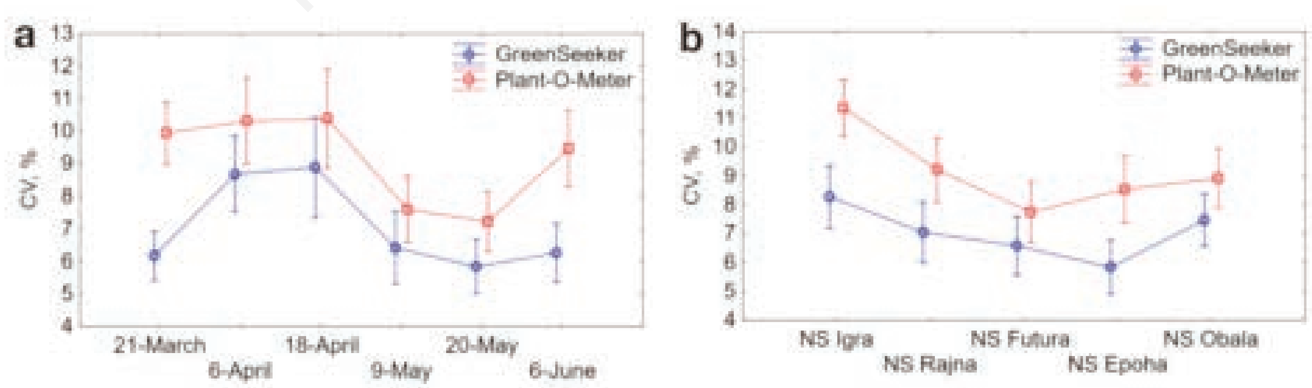


Figure 8. Coefficient of variation with 95% confidence intervals of normalized difference vegetation index: (a) effect of sensing date on CV; (b) effect of wheat variety on CV. The letters on the graph represent the various classes determined by ANOVA at a 0.05 level of significance. CV, coefficient of variation.

sensor, the measured values of wheat biological parameters and NDVI that were compared were changed, so that data from different varieties were merged into one date. Thus, the variability of the data caused by variety was generalized over the treatments and replications. Finally, NDVI values were compared with wheat traits on a general basis by grouping the data into treatments (replications) and omitting the sensing dates and varieties as factors, in order to give a one-way sensor-to-sensor comparison. With this data analysis approach, the variation of the output results decreased, thus increasing the accuracy of the regression models.

By using the GreenSeeker sensor, the raw values were the best for explaining grain yield variations, with an R^2 of 0.93 for NS Igra (May 9th) and NS Epoha (May 20th), while $NDVI_{Plant\text{-}O\text{-}Meter}$ provided the best yield predictions on May 9th for NS Epoha (R^2=0.93) and on May 20th for NS Obala (R^2=0.93). The weakest positive correlation (R^2=0.20) for all plots and both sensors was achieved for NS Rajna (May 20th), which generally coincides with the results from Table 3, showing Plant-O-Meter NDVI. The

Table 2. Relationship between NDVIGreenSeeker and grain yield.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$
NS Igra	0.75	0.85	0.75	0.93	0.90	0.82	0.94
NS Rajna	0.78	0.87	0.64	0.81	0.20	0.82	0.82
NS Futura	0.69	0.78	0.81	0.50	0.84	0.62	0.86
NS Epoha	0.74	0.86	0.92	0.92	0.93	0.83	0.94
NS Obala	0.65	0.50	0.76	0.80	0.89	0.78	0.88
$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$	0.83	0.87	0.89	0.89	0.91	0.86	0.96

R^2, relationship.

Table 3. Relationship between $NDVI_{Plant\text{-}O\text{-}Meter}$ and grain yield.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$
NS Igra	0.56	0.89	0.80	0.89	0.79	0.78	0.94
NS Rajna	0.65	0.83	0.66	0.79	0.17	0.57	0.88
NS Futura	0.75	0.76	0.81	0.49	0.84	0.45	0.87
NS Epoha	0.52	0.83	0.82	0.93	0.92	0.68	0.95
NS Obala	0.34	0.60	0.65	0.82	0.93	0.70	0.86
$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$	0.77	0.89	0.91	0.92	0.89	0.73	0.96

R^2, relationship.

Table 4. Relationship between $NDVI_{GreenSeeker}$ and plant height.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$
NS Igra	0.48	0.57	0.43	0.68	0.70	0.68	0.71
NS Rajna	0.66	0.65	0.50	0.71	0.12	0.69	0.72
NS Futura	0.56	0.68	0.70	0.57	0.81	0.63	0.79
NS Epoha	0.61	0.74	0.75	0.78	0.79	0.75	0.80
NS Obala	0.51	0.36	0.55	0.66	0.74	0.63	0.71
$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$	0.73	0.81	0.83	0.88	0.87	0.80	0.90

R^2, relationship.

Table 5. Relationship between $NDVI_{Plant\text{-}O\text{-}Meter}$ and plant height.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$
NS Igra	0.43	0.67	0.52	0.71	0.69	0.72	0.77
NS Rajna	0.44	0.65	0.53	0.68	0.17	0.53	0.75
NS Futura	0.55	0.70	0.73	0.53	0.82	0.58	0.82
NS Epoha	0.34	0.70	0.66	0.76	0.80	0.63	0.78
NS Obala	0.31	0.44	0.52	0.71	0.74	0.59	0.70
$R^2_{(\frac{\text{fert. treat.}\times\text{rep.}}{\text{variety}})}$	0.68	0.82	0.85	0.87	0.87	0.79	0.91

R^2, relationship.

apparently low R^2 obtained for NS Rajna on May 20th might be caused by a human error that emerged during the NDVI data recording and its association with a specific plot. For that reason, any further explanation is considered as redundant. GreenSeeker outperformed Plant-O-Meter in early season (March 21st) and late season (June 6th) yield prediction in almost all varieties. Looking at the correlations between grain yield and $NDVI_{GreenSeeker}$, that included only variety as a factor, by associating the measurements from different dates to certain plots, the highest R^2 was achieved for NS Epoha and NS Igra (0.94), while the lowest one was recorded for NS Rajna (0.82). In the overall comparison, a significantly high positive correlation (R^2=0.96) was obtained. The comparison of Tables 2-3 revealed that, for a specific date and variety, $NDVI_{GreenSeeker}$ produced better results in grain yield modeling than $NDVI_{Plant\text{-}O\text{-}Meter}$. The Plant-O-Meter performed better in generalized analysis when date and variety were disregarded using one-way analysis.

As the trial was set up in a long-term experimental field that had been managed in the same way for decades, sensor data were convenient for modeling good relationships between NDVI and plant height (Tables 4-5). Unlike grain yield models, plant height data were fitted by using linear rather than polynomial regression analysis (*Supplementary Tables 3-4*). Linear regression models were also used by Lu *et al*. (2017). The wheat variety and the sensing date had the biggest impact, independently on the sensor equipment. In contrast to NS Igra and NS Rajna (Table 4), where NDVI fitted better a quadratic polynomial function, NDVI data from both sensors showed the best fit to plant height when a linear model was used for NS Epoha. The GreenSeeker sensor's measurements provided the best R^2 of 0.78 for NS Epoha (dated May 9th) and reasonably lower explanation confidence for NS Igra (R^2=0.43, March 21st). While $NDVI_{Plant\text{-}O\text{-}Meter}$ readings for NS Futura had the highest correlation (R^2=0.8) on May 20th and the lowest one (R^2=0.31) on March 21st. The common feature of both sensors is reflected in the weaker relationship between NDVI and plant height in the early stage of plant development, while in the next development stages, this relationship gradually strengthens. For all wheat varieties, GreenSeeker outperformed Plant-O-Meter in the early season (March 21st) plant height forecast. By connecting the measurements from various dates to specific plots, it was possible to determine the correlations that solely took into account the variety. Corresponding to the plant height regression analysis, a generalized approach that included variety as a factor promoted the preference of $NDVI_{Plant\text{-}O\text{-}Meter}$ (R^2=0.70-0.82) in the prediction of plant height compared to $NDVI_{GreenSeeker}$ (R^2=0.71-0.80). On the other hand, if the sensing date was considered as a factor, GreenSeeker's data would better fit with plant height (R^2=0.73-0.88) rather than Plant-O-Meter's data, (R^2=0.68-0.87). From the sensor perspective, it is possible to see the difference in the average correlations among the varieties, where $NDVI_{Plant\text{-}O\text{-}Meter}$ showed a better final R^2 (0.70-0.82) rather than $NDVI_{GreenSeeker}$ (0.71-0.80). Furthermore, Plant-O-Meter data were more closely fitted to the desired model (R^2=0.91) rather than GreenSeeker data (R^2=0.9) in a generalized correlation analysis between NDVI and plant height.

R^2 is shown in Tables 6 and 7 as a result of comparisons between sensor readings and spike length data. Regression parameters indicate weak relationships for both data sets, especially for date- and variety-specific analysis (R^2<0.4). On this level of observation, sensors equally behaved in terms of spike length prediction scope. Nonetheless, the obtained results show that there is some advantage in the prediction accuracy of the $NDVI_{Plant\text{-}O\text{-}Meter}$ on May 6th. Valid improvements in R^2 are derived by using a one-way approach with sensing date as a category. Plant-O-Meter recordings partially achieved a higher model confidence with this analysis setup, particularly on April 6th (R^2=0.42), April 18th (R^2=0.43),

Table 6. Relationship between $NDVI_{GreenSeeker}$ and spike length.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{fert.\ treat. \times rep.}{variety})}$
NS Igra	0.22	0.23	0.11	0.23	0.18	0.10	0.22
NS Rajna	0.10	0.09	0.04	0.15	0.02	0.17	0.13
NS Futura	0.13	0.19	0.21	0.05	0.20	0.11	0.21
NS Epoha	0.06	0.15	0.14	0.18	0.18	0.17	0.17
NS Obala	0.22	0.25	0.33	0.40	0.37	0.39	0.41
$R^2_{(\frac{fert.\ treat. \times rep.}{variety})}$	0.36	0.38	0.38	0.46	0.42	0.42	0.47

R^2, relationship.

Table 7. Relationship between $NDVI_{Plant\text{-}O\text{-}Meter}$ and spike length.

Variety/date	March 21st	April 06th	April 18th	May 9th	May 20th	June 6th	$R^2_{(\frac{fert.\ treat. \times rep.}{variety})}$
NS Igra	0.22	0.22	0.15	0.19	0.18	0.18	0.25
NS Rajna	0.05	0.11	0.05	0.17	0.02	0.20	0.16
NS Futura	0.15	0.21	0.21	0.06	0.19	0.13	0.22
NS Epoha	0.04	0.09	0.13	0.17	0.18	0.19	0.16
NS Obala	0.11	0.28	0.32	0.33	0.39	0.35	0.39
$R^2_{(\frac{fert.\ treat. \times rep.}{variety})}$	0.35	0.42	0.43	0.44	0.43	0.45	0.49

R^2, relationship.

May 20th (R^2=0.43) and June 6th (R^2=0.45). With reference to the variety, $NDVI_{Plant\text{-}O\text{-}Meter}$ produced better results for NS Igra (R^2=0.25), NS Rajna (R^2 = 0.16) and NS Futura (R^2=0.22), while $NDVI_{GreenSeeker}$ showed an advantage for NS Epoha (R^2=0.17) and NS Obala (R^2=0.41). Overall regression, by omitting sensing dates and varieties as different ranks, expressed Plant-O-Meter to be a bit more reliable sensor for spike length prediction with a moderate level of confidence (49%), although GreenSeeker reached a similar score (47%).

Discussion

For the purpose of Plant-O-Meter sensor validation testing in outdoor conditions, the experiment was set up in a field, which was undergoing long-term soil treatment with the intention to minimize side effects on sensor measurements. Small field area helped to minimize uncontrollable spatial variations, reduce the data collection time, avoid possible interruptions in measurement quality and follow the technical specifications of the used equipment (impact of temperature, battery capacity, sensor calibration, *etc*.), as well as the subsequent processing of the data. The included factors (variety and fertilizer) showed unique contributions to the variances for the observed wheat indicators, as shown in Figure 4, which is a prerequisite for effective data analysis and subsequent modeling using NDVIs. Due to the mutual genetic origin of the included wheat genotypes, it can be concluded that variety as a controlled factor did not dominate in grain yield variability. The wheat variety had a significant impact on plant height and spike length. Wheat properties noticeably varied after fertilizer application, according to an analysis of the effects. As the agronomical aspects of the performed treatments were not the focus of this study, a more in-depth examination of them was omitted.

The variability of NDVI shown in Table 1 can be associated with different sources. One can be related to the performed treatment, which is expected to be as extensive as possible, even if it includes the uncontrolled impact of operator handling, sensor hardware imperfection, external lighting side effects, *etc*. ANOVA produced the same data grouping for the measurements from the tested sensors, indicating that GreenSeeker and Plant-O-Meter have equivalent linearity and repeatability. This general overview revealed the consistency between NDVI values for different varieties when comparing tested sensors. The results in Table 1 did not uncover the Plant-O-Meter characteristics in wheat trait modeling, so further analysis was undertaken.

Visual analysis of Figure 4 reveals NDVI variations across the sensing dates for both sensors. Figure 4 shows how NDVI data values change with the stages during the wheat growing season. This pattern was also observed by Magney *et al*. (2016), who looked at how useful the daily NDVI data can be for monitoring crop phenology. The saturation of NDVI occurred on May 9th (raising), when the highest values were recorded both for GreenSeeker and Plant-O-Meter. In the peak of vegetative growth, the saturation effect signifies a non-linear asymptotic flattening or loss of sensitivity, of the curve between NDVI *vs*. biomass. The red band, whose light is strongly absorbed by pigments in plants, is the primary cause of NDVI saturation. When vegetation becomes very dense, the presence of supplementary pigments within a leaf leads to a sustained and relatively unchanging level of reflectance. These pigments also have a modest capacity to light in the blue band. Even if the pigment concentration reaches a certain level, more pigments will lead to a low blue band reflectance. The positive slope of the NDVI curve defines the period of vegetation development and any deviation is attributed to the uneven field conditions. NDVI variations over the downslope curve occur from the perspective of variety ripening. Sensor characteristics in the maturation stages of wheat are not valuable for the early detection of nitrogen deficiency but they can be beneficial for grain yield and biomass production end-of-season prediction (Panek and Gozdowski, 2020).

By showing the whole data sets from various sensing dates on the graph, Figure 5 illustrates the link between the NDVI values of the tested sensors. The NDVI values continuously replicated by the Plant-O-Meter sensor readings over the observation period in winter wheat canopies were comparable to those of the GreenSeeker (94% of the variation is explained, $P<0.01$, RMSE = 10.03%). This might be an excellent result for a device that is designed and manufactured in a developing country such as Serbia to be commercially used on a global scale. However, more comprehensive testing must be conducted to include different crops, seasons and regions. The Plant-O-Meter's stability with respect to the reference GreenSeeker suggests that no calibration is required before use, which means that the Plant-O-Meter could be used for extended periods of time without any concern about the quality of the readings. Kitić *et al*. (2019) also confirmed the reading stability but under laboratory conditions. Even if there was some variation in the stability data, it was most likely caused by the operator and the ability of the GreenSeeker and Plant-O-Meter sensors to accurately assess a canopy. The majority (90%) of the numbers lie between 0 and 12%, with 5% being the most often estimated result, which is quite acceptable. The findings demonstrate that various operators and sensors can provide outcomes that are comparable, despite the evident variations. This is crucial for a prototype model, especially if it will be produced in large quantities and used by several operators. However, the operators should be taught and given enough room to gain experience, in order to feel comfortable and competent operating the pocket sensors.

The correlations between sensors were shown based on the measurements that were taken to simulate a trend in the stability of Pearson's R-value across the sensing time. The data analysis method (Figure 7) shows that there is a noticeable divergence of R-values along the replications. It also shows that replications and sensing dates were the primary factors that influenced the correlations between the sensors' values. The correlation between $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ was the highest in the first replication of the wheat treatment (0.96), while it was the lowest in the second and third ones (0.96). On March 21st, when the tillering stage ended, a significantly lower R was noted (0.73), compared to the later wheat development stages (0.95-0.96). The level of the synergetic impact of measurement inconsistency induced by subjective operator handling error during the sensing process and intrinsic variations at the plot and subplot level could explain the apparent disagreement between raw sensor data in a specific scope. Crain *et al*. (2012) investigated the aforementioned factors, which included sensing angle variability and GreenSeeker height. They discovered some specific interactions between the operator and the logged data. Proximal sensing requires a small distance between the sensor and the inspected object, so that the sensor itself must be hand-held or vehicle-mounted. Eitel *et al*. (2009) and Huete (1987) observed that significant parameters influencing crop canopy reflectance are sensor view angle, atmospheric conditions, canopy architecture and plant background. In this study, the man-operating concept was practiced. Every measurement session took about 1.5 hours, which can be tiring, especially if the additional effort of the operator for constant sensor leveling and canopy height maintenance is taken into account. During testing, the use of sensors caused battery discharge, which may have affected the stability of the readings over time. Two operators were in charge of field works during the testing, indicating the possibility of subjective error. The experience that emerged from the given data suggests that sensor mounting platforms that protect records from disturbance caused by height (scene size) and exposure angle (scene size, leaf/background ratio) should be included in a future study. The lower R coefficient from March 21st was likely due to the fact that each plot's plant canopy covers the soil in a different way. Uneven sensor movement by the operator might have also caused differences between the data from GreenSeeker and Plant-O-Meter.

Figure 8 provides a representation of RMSE as a quality estimate of the linear correlation between the observed NDVI values. The sensing date gave a statistically significant contribution to RMSE even if the wheat variety and trial replication did not influence the relationship between NDVI from GreenSeeker and Plant-O-Meter, respecting the commonly used threshold for statistical significance ($P \leq 0.05$). This implies that a lower RMSE was obtained for a higher NDVI, probably due to smaller internal heterogeneity of subplots in terms of vegetative growth with magnified differences among subplots (treatments).

The coefficient of variation was used to represent the level of variation in the NDVI. This parameter is considered as crucial for determining how well the sensor can forecast yield characteristics. Based on the results, it was possible to discover that the CV values for both NDVI records were maximum on April 18th, *i.e.*, during the stem elongation stage, which is characterized by fast growth (Figure 8). Uneven emergence of the plants in the early stages of their growth, variations in residue kind and cover, as well as the variations in other surface characteristics, are likely the main causes of the significant coefficient of variation.

By comparing the NDVI variation, there were a few discernible changes found in terms of CV. NDVI data acquired from Plant-O-Meter led to a higher CV (9.15%), when compared to that reported for GreenSeeker's NDVIs (7.04%). This resulted when overall averages were considered. The GreenSeeker data were most likely less influenced by noise rather than the Plant-O-Meter ones, resulting in higher relative stability reading. Noise reduces measurement accuracy and resolution, limiting the minimum quantity of measurements that can be performed with a specified degree of uncertainty (Vig and Walls, 2000). Although the NDVI values produced by GreenSeeker were consistently higher than the NDVI values produced by Plant-O-Meter, $NDVI_{GreenSeeker}$ did not promote higher variation. This finding suggests that the relative accuracy of the data may be more informative rather than the absolute data accuracy. Following the completion of the regression analysis, a response will be provided about this issue. The sensor type had a statistically significant impact on the CV of NDVI on March 21st ($CV_{GreenSeeker}$ = 6.17%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.93%) and June 6th ($CV_{GreenSeeker}$ = 6.25%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.4%). The coefficient of variation was the lowest one for both data sets over the period between May 9th ($CV_{GreenSeeker}$ = 6.41%; $CV_{Plant\text{-}O\text{-}Meter}$ = 7.59%) and May 20th ($CV_{GreenSeeker}$ = 5.84%; $CV_{Plant\text{-}O\text{-}Meter}$ = 7.21%). With the increasing biomass in the scanning area, a low variability of NDVI is expected, presumably due to the well-known saturation effect or loss of sensibility (Yue *et al.*, 2019). In the last sensing session, Plant-O-Meter showed a significantly higher sensitivity. Even if in the late development stage nitrogen management is not feasible at all, predictions of certain wheat properties can be successfully performed (Jin *et al.*, 2017). Plant-O-Meter is able to measure NDVI, which provides a clearer differentiation among wheat varieties in terms of CV (Figure 8b). For NS Igra ($CV_{GreenSeeker}$ = 8.27%; $CV_{Plant\text{-}O\text{-}Meter}$ = 11.35%), NS Rajna ($CV_{GreenSeeker}$ = 7.04%; $CV_{Plant\text{-}O\text{-}Meter}$ = 9.21%) and NS Epoha ($CV_{GreenSeeker}$ = 5.84%; $CV_{Plant\text{-}O\text{-}Meter}$ = 8.53%), significantly different average CV values of $NDVI_{GreenSeeker}$ and $NDVI_{Plant\text{-}O\text{-}Meter}$ were observed. Additionally, NS Futura had the lowest $CV_{Plant\text{-}O\text{-}Meter}$ average value, while NS Epoha had the lowest $CV_{GreenSeeker}$ value (5.83%). Overall sensor data variability does not confirm its performance in the prediction of wheat traits, as correlation analysis will be the final criterion. When properties are measured in the natural environment by using sensing methods, the small-scale heterogeneity of soil or plants contributes to short-term changes in the sensor signal; this is described as the noise effect (Kerry *et al.*, 2010). The operator and the precision of the measurements taken with the Plant-O-Meter and GreenSeeker sensors are likely due to some variation in the stability data. The findings demonstrate that various operators can produce comparable outcomes, despite the discrepancies that were noticed. Even if the Plant-O-Meter and GreenSeeker sensors had comparable performance, there were instances of noticeable distinction between the NDVI values at a level of α=0.05. The variations in NDVI occurred at low values of this index, even if the Plant-O-Meter sensor, in the vast majority of cases, had confidence levels of 95% or more, that were within GreenSeeker's acceptable range. The significance of the existing deviation between the two data sets can be analyzed from the perspective of its impact on the accuracy of the fertilizer rate. In the case of wheat, N recommendations would only deviate by 3-5 kg N ha^{-1} from the actual rate, considering that the Plan-O-Meter NDVI lags by an average of 0.03. According to the sensor-based nitrogen calculator (https://www.nue.okstate.edu/SBNRC/mesonet.php), even with a difference of 0.05 in NDVI, the suggested N rate would differ from the required N rate by 8-10 kg N ha^{-1}. In real field conditions, the spinning disc centrifugal fertilizer spreader, which is widely used by farmers all over the world, causes more distortion of the target rate (10-35%), particularly at extremely low or high rate settings (Parish, 2002), due to the spreader patterns varying with changes in rate setting up. According to Lawrence and Yule (2007), urea application using a spinning disc spreader was successful only 24% of the time within the planned application rate. Therefore, even with minor deviations, a recommended rate from the Plant-O-Meter would typically be close enough to the needed rate, so that application error and other environmental factors might have a higher impact on crop development.

The ability of daily NDVI data to forecast harvest metrics at the end of the growing season is shown in Tables 2-7. The correlations were developed independently for each variety and each sensing date, in order to verify the reliability of sensor readings and analyze the impacts of the aforementioned factors on the quality of regression models built from empirical data. In general, the modeling of grain yield and plant height with sensor readings suggests strong relationships (in most cases, R^2>0.8), which are a consequence of the long-term trial effect on the stability of output (Tables 2-3). This statement is supported by Laurent *et al*. (2022), who uncovered that the standard deviation was three times higher in on-farm experiments, compared to small-plot trials conducted at agricultural experimental stations of research institutes or universities. Tables 2 and 3 show that the sensors were able to detect the majority of variations in grain yield of all wheat varieties with high confidence, even for specific sensing dates. The graphs in the supplementary material (*Supplementary Tables 1-2*) testify that all of the examined correlations between sensor data and wheat traits were positive. A number of studies on crop yield as their main research topic have used the linear model (Oglesby *et al.*, 2022) as the best mathematical way to describe the relationship between NDVI and grain yield. However, the polynomial function was also used (Varinderpal-Singh *et al.*, 2022). Given model curves have a tendency to flatten, which means that NDVI loses its ability to predict yield when plants reach a certain level of development. By comparing Tables 2 and 3, it is evident that $NDVI_{GreenSeeker}$ outperformed $NDVI_{Plant\text{-}O\text{-}Meter}$ in grain yield modeling for a particular date and variety. In nearly all varieties, GreenSeeker performed better than Plant-O-Meter in predicting grain yield early (March 21st) and late (June 6th) in the growing season. These two dates delineate R^2 for the sensors, where $NDVI_{GreenSeeker}$ was more relevantly efficient rather than NDVIs from Plant-O-Meter. An accurate early-season yield forecast has broad implications for farm resource management (*e.g.*, nitrogen and water management), economic trading (Yiqing *et al.*, 2017) and global food security. However, in generalized analysis, where date and variety were ignored by using one-way analysis, the Plant-O-Meter performed better. The captured characteristics could be associated with the discrepancy in Plant-O-Meter readings during the sensing dates, which was smoothed out in the one-way approach and diminished the advantage of GreenSeeker.

The modeling of plant height using sensor data mainly suggests moderate to strong connections as a consequence of the long-term trial effect on output stability (Tables 4-5). The lower association between NDVI and plant height in the early stages of plant growth, whereas this relationship eventually becomes stronger with plant development, reflects the similar characteristics of both sensors. Again, in the early season (March 21st) GreenSeeker performed better than Plant-O-Meter in plant height prediction for all cultivars. This could be very important due to the fact that early prediction of biomass yield can help stakeholders, energy managers and decision-makers working in the sustainable and renewable energy sectors to consider agriculture biomass for energy production at a larger scale (Saleem, 2022). $NDVI_{GreenSeeker}$ data were slightly more linearly oriented to the plant height data pattern, implying that GreenSeeker had slightly more favorable characteristics in terms of sensitivity for detecting plant height changes (*Supplementary Tables 3-4*). $NDVI_{Plant\text{-}O\text{-}Meter}$ showed a better R^2 (0.70-0.82), that was higher than $NDVI_{GreenSeeker}$, allowing to see the difference in the average correlations among the varieties from the sensor's point of view (0.71-0.80). GreenSeeker's data were better fitted with plant height (R^2=0.73-0.88) rather than Plant-O-Meter's ones if the sensing date is taken into account (R^2=0.68-0.87). Overall regression results show that GreenSeeker's data are slightly more suitable for predicting plant height rather than Plant-O-Meter ones.

Tables 6 and 7 show the coefficient of determination as a consequence of comparisons and modeling of spike length from sensor values. On this level of observation, sensors equally behaved in terms of spike length prediction scope (R^2<0.4). Nonetheless, the given tables show that there is some advantage in the prediction accuracy of the $NDVI_{Plant\text{-}O\text{-}Meter}$ on May 6th. Experts in plant breeding might be interested in this, because it could help them to understand how different wheat traits affect grain yield per unit area, which is becoming more important in harsh growing conditions. With a reasonable degree of confidence (49%), the generalized regression determined that Plant-O-Meter was a little more accurate as a sensor for predicting spike length rather than GreenSeeker (47%), even after excluding the sensing dates and varieties as separate rankings.

Conclusions

In this study, a hand-held active proximal sensor called Plant-O-Meter was tested to see how well it could predict wheat traits, by comparing its features to those of the well-proven GreenSeeker. With a high level of confidence, we came to the following conclusions by using about 2,000 representative readings per sensor that were based on repeated measurements of each trial plot, it was possible to conclude that:

- the GreenSeeker sensor had consistently higher output values rather than Plant-O-Meter over the entire sensing period, i.e., on average 0.029 (6.36%); the differences between the compared values (90%) were between 0 and 12%; the different central operating wavelengths, the impact of the operator on sensor readings and the different sensing angles of Plant-O-Meter and GreenSeeker can help to explain this;
- although the Plant-O-Meter and GreenSeeker readings were analogous, they were noticeably different at the 0.05 level in some cases; the previously mentioned differences in NDVIs appeared when this index was low but, in most cases, the Plant-O-Meter sensor had confidence levels of 95%, which were within the limits of the GreenSeeker;
- the stability of the Plant-O-Meter, compared to the GreenSeeker, suggests that it does not need to be calibrated before use; therefore, the Plant-O-Meter could be used for a long time without worrying about the accuracy of the readings;
- the NDVI data from Plant-O-Meter distinguished wheat varieties more clearly in terms of CV rather than the NDVI ones from GreenSeeker;
- GreenSeeker performed better than Plant-O-Meter for predicting grain yield early (March 21st) and late in the growing season (June 6th), for almost all varieties; based on NDVIGreenSeeker data, the performance of yield modeling was better, on average, by 5.1%, compared to NDVIPlant-O-Meter; NDVIGreenSeeker was 3% more accurate rather than NDVIPlant-O-Meter in predicting plant height and there were almost no difference in predicting spike length;
- in the field, the Plant-O-Meter was easier to use, because it includes a mobile phone app that automatically records the measured data and synchronizes them with the cloud server; this option eliminates room for subjective error during measurement and data logging; moreover, the raw data can be retrieved from any place and post-processed with geospatial representation in geographic information systems; Plant-O-Meter's raw data recordings saved in digital format are easy to filter for outliers, reducing "pseudo" variations and their negative impact on data modeling; this can be a disadvantage for the common user-farmer, who usually does not have the appropriate skills to manipulate data;
- for final approval, the Plant-O-Meter proximal sensor must pass more tests, including trials with wide row crops, which are different in their canopy architecture and Leaf Area Index and have a high nitrogen demand, that can be managed in the middle of the growing season using this type of diagnostic approach.

Generally, the Plant-O-Meter sensor has commercial potential, because it is affordable (<500$) and user-friendly, along with its proven ability to accurately map crop status variations, which may make it a viable and economical option for small and medium-sized farmers who want to implement precision agriculture.

References

Alvar-Beltrán J., Fabbri C., Verdi L., Truschi S., Dalla Marta A., Orlandini S. 2020. Testing Proximal Optical Sensors on Quinoa Growth and Development. Remote Sens. 12:1958.

Ang Y., Norasma N., Ya C., Roslin N., Ismail M., Che'Ya N. 2020. Rice Chlorophyll Content Monitoring using Vegetation Indices from Multispectral Aerial Imagery. Pertanika J. Sci. Technol. 28:779-95.

Bannari A., Morin D., Bonn F., Huete A.R. 1995. A review of vegetation indices. Remote Sens. Rev. 13:95-120.

Bean G.M., Kitchen N.R., Camberato J.J., Ferguson R.B., Fernandez F.G., Franzen D.W., Laboski C.A.M., Nafziger E.D., Sawyer J.E., Scharf P.C., Schepers J., Shanahan J.S. 2018. Active-Optical Reflectance Sensing Corn Algorithms Evaluated over the United States Midwest Corn Belt. Agron. J. 110:2552-65.

Cao Q., Miao Y., Li F., Gao X., Liu B., Lu D., Chen X. 2017. Developing a new Crop Circle active canopy sensor-based precision nitrogen management strategy for winter wheat in North China Plain. Precision Agric. 18:2-18.

Crain J., Ortiz-Monasterio I., Raun B. 2012. Evaluation of a Reduced Cost Active NDVI Sensor for Crop Nutrient Management. J. Sens. 2012:1-10.

Curran P.J., West S.G., Finch J.F. 1996. The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis. Psychol. Methods. 1:14.

Eitel J.U.H., Long D.S., Gessler P.E., Hunt E.R., Brown D.J. 2009. Sensitivity of ground-based remote sensing estimates of wheat chlorophyll content to variation in soil reflectance. Soil Sci. Soc. Am. J. 73:1715-23.

Fertilizers Price Index. n.d. Available from: https://ycharts.com/indicators/fertilizers_index_world_bank (accessed 10.31.22).

Huete A.R. 1987. Soil-dependent spectral response in a developing plant canopy1. Agron. J. 79:61-8.

Jaćimović G. 2012. Optimization of mineral nutrition of wheat, depending on weather conditions during the year (Ph.D. Thesis). University of Novi Sad, Serbia.

Jin X., Liu S., Baret F., Hemerlé M., Comar A. 2017. Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. Remote Sens. Environ. 198:105-14.

Jovanović M., Pavić D., Mesaroš M., Stankov U., Pantelić M., Armenski T., Dolinaj D., Popov S., Ćosić Đ., Popović L., Frank A., Crnojević V. 2013. Water shortage and drought monitoring in Bačka region (Vojvodina, North Serbia) – setting-up measurement stations network. Geogr. Pannonica. 17:114-24.

Kerry R., Oliver M.A., Frogbrook Z.L. 2010. Sampling in Precision Agriculture, in: Oliver M.A. (eds.), Geostatistical Applications for Precision Agriculture. Springer Netherlands, Dordrecht. pp. 35-63.

Kitić G., Tagarakis A., Cselyuszka N., Panić M., Birgermajer S., Sakulski D., Matović J. 2019. A new low-cost portable multispectral optical device for precise plant status assessment. Comput. Electron. Agric. 162:300-08.

Laurent A., Heaton E., Kyveryga P., Makowski D., Puntel L.A., Robertson A.E., Thompson L., Miguez F. 2022. A yield comparison between small-plot and on-farm foliar fungicide trials in soybean and maize. Agron. Sustain. Dev. 42:86.

Lawrence H.G., Yule I.J. 2007. Estimation of the in□field variation in fertiliser application. New Zeal. J. Agric. Res. 50:25-32.

Long-term Field Experiments in Europe. n.d. Available from: https://lte.bonares.de/lte-details/153/ (accessed 11.6.22).

Lu J., Miao Y., Shi W., Li J., Yuan F. 2017. Evaluating different approaches to non-destructive nitrogen status diagnosis of rice using portable RapidSCAN active canopy sensor. Sci. Rep. 7:14073.

Magney T.S., Eitel J.U.H., Huggins D.R., Vierling L.A. 2016. Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. Agric. For. Meteorol. 2016:46-60.

Oglesby C., Fox A.A.A., Singh G., Dhillon J. 2022. Predicting in-season corn grain yield using optical sensors. Agronomy. 12:2402.

Panek E., Gozdowski D. 2020. Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. Remote Sens. App. Soc. Environ. 17:100286.

Parish R.L. 2002. Rate setting effects on fertilizer spreader distribution patterns. Appl. Eng. Agric. 18.

Raper T.B., Varco J.J., Hubbard K.J. 2013. Canopy-based normalized difference vegetation index sensors for monitoring cotton nitrogen status. Agron. J. 105:1345-54.

Raun W.R., Solie J.B., Johnson G.V., Stone M.L., Mullen R.W., Freeman K.W., Thomason W.E., Lukina E.V. 2002. Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agron. J. 94:815-20.

Raun W.R., Solie J.B., Martin K.L., Freeman K.W., Stone M.L., Johnson G.V., Mullen R.W. 2005. Growth stage, development, and spatial variability in corn evaluated using optical sensor readings. *Contribution from the Oklahoma Agricultural Experiment Station and the International Maize and Wheat Improvement Center (CIMMYT). J. Plant Nutr. 28:173-82.

Saleem M., 2022. Possibility of utilizing agriculture biomass as a renewable and sustainable future energy source. Heliyon. 8:e08905.

Sankaran S., Espinoza C.Z., Hinojosa L., Ma X., Murphy K. 2019. High-throughput field phenotyping to assess irrigation treatment effects in quinoa. Agrosyst. Geosci. Environ. 2:180063.

Sharma L., Bu H., Denton A., Franzen D., 2015. Active-optical sensors using red NDVI compared to red edge NDVI for Prediction of corn grain yield in North Dakota, U.S.A. Sensors 15:27832-53.

Statistical yearbook of Republic of Serbia. 2022. Statistical Office of the Republic of Serbia, Belgrade.

Tagarakis A.C., Kostić M., Ljubičić N., Ivošević B., Kitić G., Pandžić M. 2022. In-field experiments for performance evaluation of a new low-cost active multispectral crop sensor. In: Bochtis D.D., Lampridi M., Petropoulos G.P., Ampatzidis Y., Pardalos P. (Eds.), Information and Communication Technologies for Agriculture—Theme I: Sensors, Springer Optimization and Its Applications. Springer Int. Pub. Cham. pp. 305-25.

Varinderpal-Singh, Kunal, Kaur R., Mehtab-Singh, Mohkam-Singh, Harpreet-Singh, Bijay-Singh. 2022. Prediction of grain yield and nitrogen uptake by basmati rice through in-season proximal sensing with a canopy reflectance sensor. Precision Agric. 23:733-47.

Vig J.R., Walls F.L. 2000. A review of sensor sensitivity and stability. Proceedings of the 2000 IEEE/EIA International Frequency Control Symposium and Exhibition (Cat. No.00CH37052), Kansas City, MO, USA. pp. 30-3.

World Agricultural Production. 2022. Foreign Agricultural Service/USDA.

World Reference Base (WRB). 2014. World Reference Base for soil resources 2014: international soil classification system for naming soils and creating legends for soil maps. Available from: https://www.fao.org/3/i3794en/I3794en.pdf

Xia T., Miao Y., Wu D., Shao H., Khosla R., Mi G. 2016. Active optical sensing of spring maize for in-season diagnosis of nitrogen status based on Nitrogen Nutrition Index. Remote Sens. 8:605.

Yin H., Cao Y., Marelli B., Zeng X., Mason A.J., Cao C. 2021. Soil Sensors and Plant Wearables for Smart and Precision Agriculture. Adv. Mater. 33:2007764.

Yiqing C., Moore K., Pellegrini A., Elhaddad A., Lessel J., Townsend C., Solak H., Semret N. 2017. Crop yield predictions - high resolution statistical model for intra-season forecasts applied to corn in the US. Curr. Biol. 7:R126.

Yue J., Yang G., Tian Q., Feng H., Xu K., Zhou C. 2019. Estimate of winter-wheat above-ground biomass based on UAV ultra-high-ground-resolution image textures and vegetation indices. ISPRS J. Photogram. Remote Sens. 150:226-44.

Online supplementary material:

Table S1. Regression models for explaining grain yield using GreenSeeker normalized difference vegetation index data.

Table S2. Regression models for explaining grain yield using Plant-O-Meter normalized difference vegetation index data.

Table S3. Regression models for explaining plant height using GreenSeeker normalized difference vegetation index data.

Table S4. Regression models for explaining plant height using Plant-O-Meter normalized difference vegetation index data.

Table S5. Regression models for explaining spike length using GreenSeeker normalized difference vegetation index data.

Table S6. Regression models for explaining spike length using Plant-O-Meter normalized difference vegetation index data.