Assessment of plant density for barley and wheat using UAV multi-spectral imagery for high-throughput field phenotyping

Norman Wilke [1,*], Bastian Siegmann [1], Johannes A. Postma [1], Onno Muller [1], Vera Krieger [1], Ralf Pude [2,3] and Uwe Rascher [1]

Abstract

Cereal plant density is a relevant agronomic trait in agriculture and high-throughput phenotyping of plant density is important for the decision-making process in precision farming and breeding. It influences the water as well as the fertilization requirements, the intraspecific competition, and the occurrence of weeds or pathogens. Using spatially high-resolution images (0.02 cm)recent studies have determined plant density using machine-learning approaches and feature extraction. However the accuracy and practical applicability decreased when only lower resolution images were available. In this study, we present an approach that uses the linear relationship between plant density manually counted in the field and fractional cover derived from a RGB and a multispectral camera equipped on an unmanned aerial vehicle (UAV). We assumed that at an early seedling stage fractional cover is closely related to the number of plants. Spring barley and spring wheat experiments, each with three genotypes and four different sowing densities, were examined. The practicability and repeatability of the methodology were evaluated with an independent experiment consisting of 42 winter wheat genotypes. This experiment mainly differed for genotypes, sowing density and season. At BBCH stage 13 plants were large enough to determine fractional cover also from the lower resolution image data. The empirical regression models using multispectral images with a ground sampling distance (GSD) of 0.69 cm were also suitable to determine plant density with a high prediction accuracy for barley and wheat ($R^2 > 0.91$, mean absolute error (MAE) < 28 plants). In addition, prediction accuracy only slightly declines for multispectral image data having 1.4 cm GSD or RGB image data having 0.6 cm GSD (MAE < 35 plants m^{-2}).
. In the independent experimental field the prediction accuracy of UAV estimated plant density showed an R^2 value of 0.83 and an MAE of less than 21 plants m^{-2} verifying empirical regression model robustness across conditions. Furthermore, manual measurements of 11 randomly selected plots proved sufficient for a user-based training of the regression model ($R^2 = 0.83$, MAE < 23 plants m^{-2}) adapted to the independent experimental field.
The method and the use of UAV image data enable high-throughput phenotyping of cereal plant density with uncertainties of less than 10 %. The practicability, repeatability and robustness of the developed approach were demonstrated in this study.

Keywords: plant density, germination rate, barley, wheat, high-throughput phenotyping, UAV.

1. Introduction

The non-invasive assessment of plant traits is becoming increasingly important in agriculture. Novel image processing methods in combination with new sensors and autonomous small aircraft will change crop phenotyping and agricultural crop production in the coming years [1–3]. In particular, unmanned aerial vehicles (UAVs) enable a flexible and cost-effective acquisition of high-spatial resolution image data [4–6]. Several plant traits have already been determined non-invasively, such as leaf area index [7–9], canopy height [10–12], biomass [13–15] and lodging [16, 17]. These plant traits are vital for applications in the fields of precision agriculture, breeding research, insurance applications or crop modeling.

The assessment of plant density under field conditions is also part of this ongoing development. Water and fertilization requirements are dependent on this parameter for agricultural management purposes. In addition, plant density has an impact on intraspecific competition and the occurrence of

weeds or pathogens [18–22]. A homogenous and ideal plant density is an important prerequisite for efficient crop production and potential yield [23–27].

Plant density depends on sowing density and germination rate. Most yield prediction models typically consider sowing density as a trait rather than plant density. Although the germination rate is often corrected for in sowing density this only takes into account the genotypic variation in germination rate determined under optimal conditions and not the influence of abiotic (temperature, moisture, soil, nutrients and frost), biotic (pests and diseases) and management factors (sowing variations, e.g. depth) [28–34]. Yield predictions can be improved by substituting estimated plant density for sowing density. However, such an implementation in modeling or decision-making is only feasible and realistic with a working high-throughput phenotyping approach [35, 36]. The current practice of representing the conditions in a field or a plot is to count individual plants within several smaller subsample areas. This approach requires agricultural experts, is time-consuming, expensive and not representative in the case of spatial variability within the field.

To replace these laborious ground data collection methods, the plant density has already been determined using different methods for maize [37, 38], potatoes [39, 40] and sugar beet [41]. The planting structures of these cultivars with evenly spaced seedlings prevent an overlap among neighboring plants in early developmental stages and allow a relatively easy quantification of the germination rate. For cereals, however, mechanical seed drilling with a non-even seed distribution is the standard practice. Plants emerge very close to each other with narrow leaves overlapping among neighboring plants and single plants develop in multiple tillers further complicating individual plant identification especially at later stages. Therefore, the aforementioned methods developed for crops with a clear row structure cannot be reliably applied to cereal crops.

Previous studies determined the plant density of cereals from images recorded by RGB cameras mounted on an UAV, field bicycle or monopod [21, 42, 43]. Machine-learning techniques based on feature extraction were used to estimate plant density. High-resolution image data (0.02 cm) were required in these studies to enable a good prediction accuracy with feature extraction [21, 42, 43]. In this context, Jin et al. [43] have already demonstrated a distinctly decreasing accuracy using image data with a lower spatial resolution (0.1 cm). The prediction accuracy was also dependent on the sowing density with decreasing performance in the case of higher sowing density and probability of overlapping plants [42, 43]. In summary, the aforementioned approaches only permit a low throughput and adaptability to high sowing frequency.

In the present study, the plant density determination of cereals is based on fractional cover estimated from UAV multispectral image data. At the early seedling stage, plant density does not affect individual plant size. Only the area of overlapping leaves increases, especially with higher plant density. However, even at high plant densities, we expected that there would be a relatively little surface area overlapping as leaves are narrow. We assumed that a higher value of fractional cover at the early seedling stage also indicates a larger number of plants and a linear relation**.** Using this hypothesis, we trained empirical regression models for barley and wheat based on reference measurements acquired in the field and UAV derived fractional cover values. The procedure does not require the assignment of single leaves to particular seeds. This enabled the determination of plant density from spatial image data with potentially lower resolution leading to higher practicability of the methodology. Sankaran et al. [44] had already made use of pixel values of a calculated vegetation index for germination assessment with a good correlation to ground truth observations. However, the visual ground truth rating of germination in their study with 10 % increments was only a rough estimation of plant density. Moreover, the development of an empirical regression model and the transferability of the model to unknown data was not investigated.

The following study aims to answer the following questions, which can be divided into factors influencing the accuracy of the empirical regression model (i to v) and the evaluation of transferability of the method to an independent test site (v, vi): (i) What is the ideal plant developmental stage for data collection? (ii) What spatial resolution is necessary to achieve both high prediction accuracy and high throughput? (iii) What influence do plant characteristics (e.g., leaf arrangement, species,

genotype) have on the methodology? (iv) Is an RGB camera sufficient or is it necessary to use a multispectral camera with five spectral bands? (v) What accuracy can be obtained when the previously trained empirical regression model is transferred to an independent site and environment? (vi) How many reference measurements are necessary for a user-based training of an empirical regression model at an independent site?

2. Material and Methods

2.1 Study site and experimental design

The study was conducted at the agricultural research station Campus Klein-Altendorf (50°37'N, 6°59'E), which is affiliated to the Faculty of Agriculture of the University of Bonn. Three experimental fields were investigated on the research campus at 66 m above sea level. The soil can be classified as a luvisol.

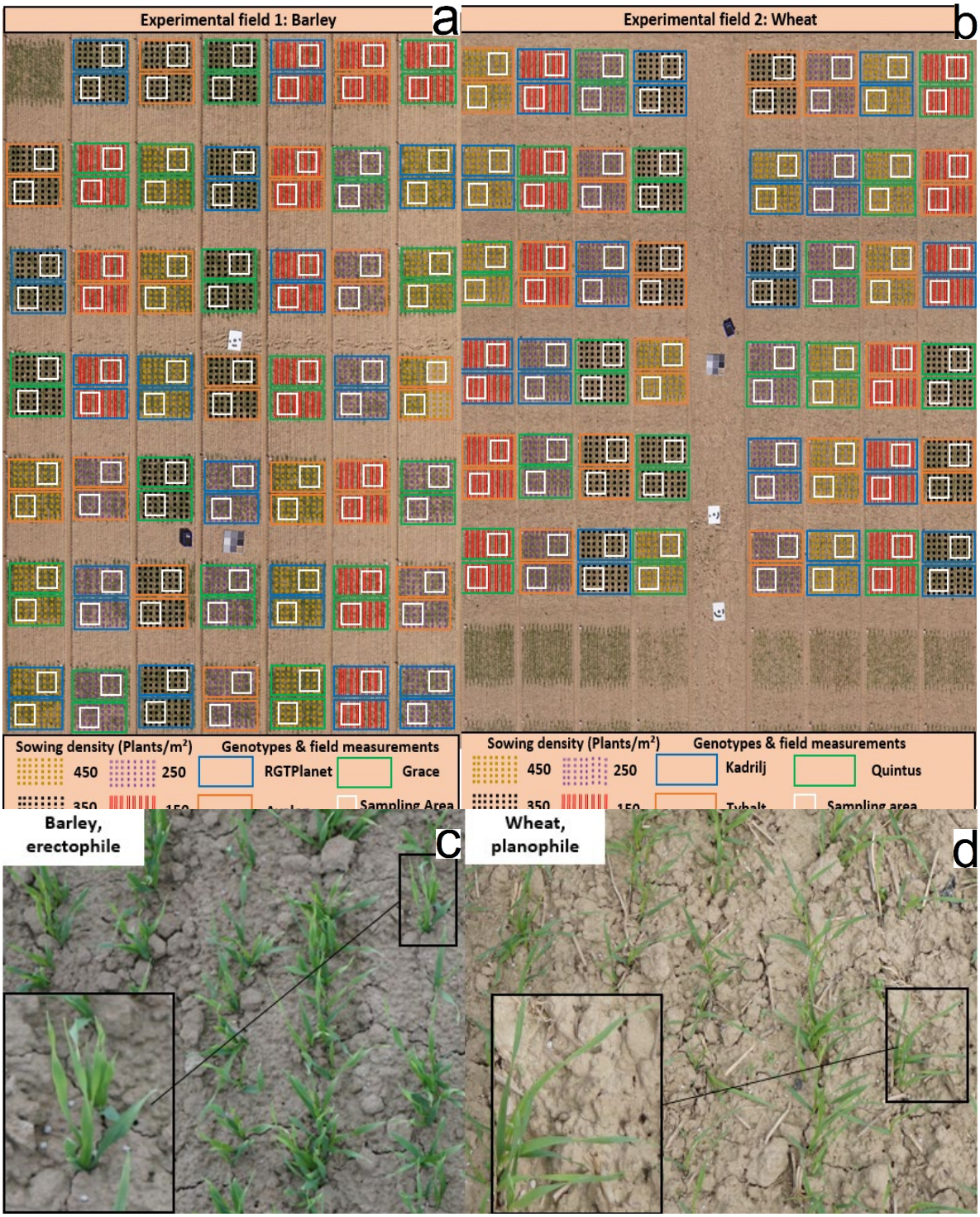


Fig. 1. Experimental fields 1 and 2, consisting of breeding plots with four sowing densities and three genotypes and a different leaf arrangement of each species **a** Experimental field 1: Spring barley; **b** Experimental field 2: Spring wheat; **c** Barley plants with mainly erectophile leaf arrangement; **d** Wheat plants with mainly planophile leaf arrangement

Two of them consisted of breeding plots (1.4×3 m) arranged in a randomized block design planted with monocultures of three spring barley genotypes (Avalon, Grace, RGTPlanet, Field 1, Fig. 1a) and three spring wheat genotypes (Quintus, Kadrilj, Tybalt, Field 2, Fig. 1b).
The genotypes were sown in four different densities (150, 250, 350 and 450 seeds m^{-2}) with eight repetitions (Fig. 1 a,b). Each repetition was divided into a sampling unit (SU) of one square meterresulting in 96 samples per species.
The seeds were sown in rows with a space of 10.4 cm for barley and wheat on 09 April. The experiments were treated with 80 kg N/ha on 19 April, herbicides were applied after data acquisition.
Spring wheat and spring barley differed strongly in their leaf arrangement. The leaf orientation of barley was erectophile with mainly vertical leaves, whereas wheat in this study had a planophile orientation with mainly horizontal leaves (Fig. 1 c,d).

Experimental field 3 was a breeding trial with 42 winter wheat genotypes arranged in a randomized block design (1.4 × 3 m). Three repetitions of each genotype were sown with a density of 460 seeds m^{-2} in rows with a space of 11 cm on 16 October (n =126). Fertilization as well as herbicides were applied after data acquisition and do not have an influence on the scene in this study. This third experimental field was used to test the repeatability and practicability of the methodology. The experiment differed in genotype, sowing density, row spacing, sowing date and season. Furthermore, the moisture content and thus the color of the soil varied among the experimental fields. The soils of experimental field 1 and 2 were dry and thus characterized by a light brown color (Fig. 2a). During wintertime, the soil of experimental site 3 was much darker because of the higher soil moisture.

2.2 Field validation

For all plots in the three experimental fields, the individual plant numbers were counted in the field within one square meter. A metal frame covering one square meter was used to facilitate high plant counting accuracy in the field (Fig. 2a). The corners were marked with pink sticks to highlight the area of interest for the UAV image data analysis.

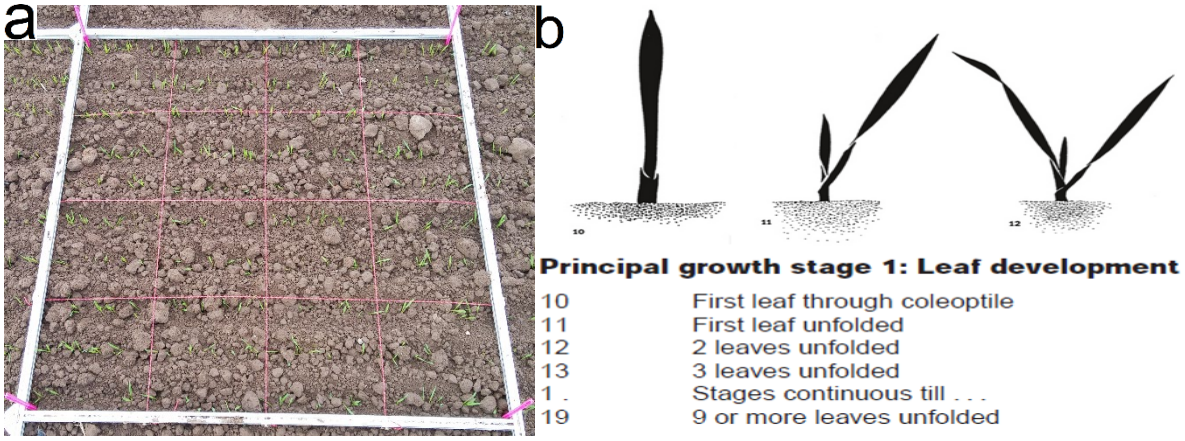


Fig. 2 a Example of the field validation setup. Metal frame (1x1 m) used for field plant counting with pink sticks at the corners to facilitate the identification of the region of interest in the UAV data; **b** Graphical illustrations of the principal growth stages in the leaf development of cereal plants according to the scale of the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) (modified according to Meyer [45]).

Plant counting in the field was conducted according to the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) scale [45] at stage 11 when the first leaf was unfolded

and the tip of the second leaf became visible (Fig. 2b). BBCH 11 has two main advantages as a reference point. First, the developmental stage enables to consider delayed development of seeds (BBCH 10) during the reference measurement. Second, at this stage the plant only consists of one unfoldedleaf and therefore it is still possible to separate the individual plants, even if the seeds have germinated closely side by side.

2.3 Unmanned aerial vehicle platform and sensor

In this study, we used the Falcon 8 octocopter (Ascending Technologies GmbH, Krailing, Germany) for data acquisition. The octocopter was equipped with a Sony Alpha 6000 (Sony Europe Limited, Weybridge, Surrey, UK) RGB camera (24 megapixels, 6000 × 4000 pixels). A fixed lens with a focal length of 20 mm was mounted on the camera, resulting in a spatial resolution of 0.20 cm at a flight altitude (FA) of 10 m above ground level. Moreover, we used a MicaSense RedEdge multispectral camera (Micasense, Seattle, USA), which contains five spectral bands to additionally cover the red edge (RE) and near infrared (NIR) spectral range. Image data at an FA of 10 m above ground level resulted in a spatial resolution of 0.69 cm.

The multispectral camera was connected to an incoming light sensor that recorded solar irradiance during image data acquisition. Together with a radiometrically calibrated reference panel acquired before and after the flight, the irradiance measurements were used for the radiometric correction of the image data. Both cameras as well as the DSL were integrated in a gimbal to compensate for pitch and roll movements of the UAV platform. Depending on the wind speed, the power of the battery was sufficient for a flight duration between 10 and 15 minutes.

2.4 Unmanned aerial vehicle data acquisition

A planned waypoint pattern with 60% across and 80% along track overlap was used for image data acquisition. The RGB images were recorded at a shutter speed of 1/1000 seconds and stored in JPEG format. The images of the five spectral bands acquired by the multispectral camera were separately saved as TIFF files.

We covered the experimental fields 1 and 2 four times. Table 1 provides an overview of the different acquisition dates, growth stages, flight altitudes and illumination conditions. At the beginning of data acquisition, barley was already at BBCH stage 12, while wheat was at stage 11. For the subsequent acquisition dates, the plants were at the next BBCH stage (Table 1). For both crops, we recorded data at BBCH stages 12, 13 and 14, which were expected to be the most appropriate growth stages for plant density assessment.

Table 1. Data acquisition of the experimental fields 1 (barley) and 2 (wheat) with information on BBCH stages of the plants, sensors used, flight altitude and illumination conditions for the data point. S: Direct solar illumination during data acquisition; (S) Minor clouds but predominantly direct solar illumination; O: Overcast sky; (O) Short periods of direct solar illumination but mainly overcast sky.

Nr.	Date	BBCH Stage		Sensor	Flight altitude (m)	Illumination conditions	
		Barley	Wheat			Barley	Wheat
1	22.04.2018	12	11	RGB	10	(S)	S
					30	O	
				Multispectral	10	(S)	
					20	S	
2	26.04.2018	13	12	RGB	10	O	(O)
					30		O
				Multispectral	10	(O)	S
					20	O	O
3	30.04.2018	14	13	RGB	10	S	O
					30		

				Multispectral	10 20	O
4	04.05.2018	15	14	RGB	10 30	S
				Multispectral	10 20	

In addition to the different spectral and spatial properties of the RGB and multispectral images, data were acquired from two different altitudes with both cameras. This allowed us to investigate which spatial and spectral resolutions were optimal for each BBCH stage. The FAs of both cameras and the corresponding ground sampling distance (GSD) are summarized in Table 2.

Table 2. Spatial resolution of the UAV images depending on the sensor and the flight altitude

Sensor	RGB		Multispectral	
Flight altitude (m)	10	30	10	20
Ground sampling distance (cm/pixel)	0.20	0.59	0.69	1.38

Furthermore, illumination conditions varied between direct solar illumination, completely overcast sky and the combination of both conditions at the different dates of data acquisition (Table 1). The first and fourth data sets were acquired under predominantly direct solar, while the second and third data sets were recorded under predominantly overcast conditions.

For experimental field 3, a data set was acquired on 17 November, at BBCH stage 12, with the multispectral sensor from an FA of 20 m above ground level (GSD = 1.38 cm) under overcast conditions. For these data acquisition parameters, flight altitude and sensor were selected based on the results of the best-performing wheat empirical regression model in relation to the lower GSD.

2.5 Data processing

Structure from motion (SfM) algorithms were used to process the individual UAV images with Agisoft Metashape software (Agisoft LLC, Saint Petersburg, Russia, version 1.5.5). The algorithms identify corresponding images by means of feature recognition. Using a certain number of overlapping images, the software recreates their orientation in a spatial three-dimensional (3D) structure [46, 47]. The primary product of the reconstruction is a 3D point cloud. The secondary and main product for this study is the two-dimensional orthomosaic which was used for further analysis. For the multispectral imagery, the radiometric calibration provided by MicaSense integrated into Agisoft Metashape was included in the processing workflow [48]. The solar irradiance acquired by the upward-looking sensor of the camera in combination with the reference panel allowed the radiometric calibration of the image data and the calculation of spectral reflectance for each pixel in the orthomosaics.

The excess green minus excess red (ExGR) [49] vegetation index was calculated for the RGB camera data (Equation 1) and the normalized difference vegetation index (NDVI) [50] for the multispectral data (Equation 2).

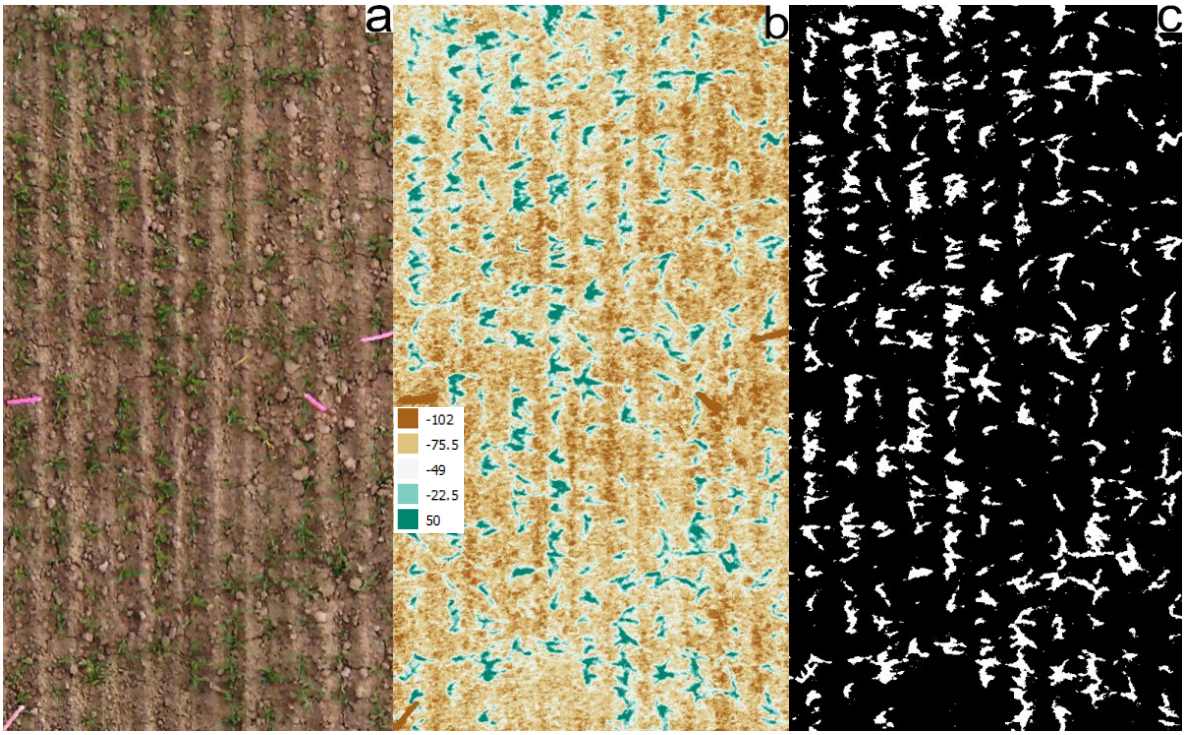
$$NDVI = \frac{Ref_{NIR} - Ref_{RED}}{Ref_{NIR} + Ref_{RED}} \quad (1)$$

$$ExGR = (2\,\text{Green} - \text{Red} - \text{Blue}) - (1.4\,Red - Green) \quad (2)$$

The vegetation indices finally allowed the fractional cover assessment by applying a threshold that divided the pixels into two classes; foreground (plant pixels) and background (soil pixels). The threshold was automatically determined by the variance between two classes based on the Otsu

method [51]. The use of Otsu thresholding is an established method based on the aforementioned variance between two classes to distinguish vegetation from the background [39, 52–55]. The procedure and segmentation performance are exemplarily illustrated in Fig. 3. Shapefiles based on the aforementioned SU were finally used to calculate the fractional cover (number of plant pixels per square meter) for the calculated vegetation indices (ExGR, NDVI) in the processed orthomosaics (Fig. 3).

Fig. 3. a Section of an UAV RGB orthomosaic at BBCH stage 13 **b** Calculated ExGR in the region of interest **c** Segmentation of plants (white) and background (black) using the ExGR and Otsu thresholding **(c)**.



2.6 Statistical analysis

The manually counted plant density in combination with the determined fractional cover per SU allowed to develop linear empirical regression models for plant density assessment. This method assumes a linear relationship between two variables: A dependent variable like fractional cover and an independent variable (variable of interest) like plant density in this study. The prediction of the plant parameter is based on a fitting function between the independent and dependent variable [56] For this, , a k-fold cross-validation was used to estimate the model and test the error rate. Firstly, we randomly divided the data sets of experimental fields 1 and 2, each with a sample size of 96, 48 times into a calibration and validation data set for the empirical regression model development. Two-thirds were used to train the model (calibration), whereas the remaining third was used to evaluate the model performance (validation). Across the models, we determined the performance using R [57] based on the averaged coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE). The advantage of the MAE over the RMSE is that it gives the same weight to all errors. Therefore, the MAE is more robust to outliers. In contrast, the commonly used RMSE weights errors with larger absolute values more than errors with smaller absolute values and thus is more sensitive to outliers [58, 59].

In order to identify which settings provide the highest model accuracies and smallest errors, three-way ANOVAs were conducted to verify the MAE including a Tukey post-hoc test. ANOVAs were fitted for each species separately. The following model in R notation was used: MAE ~ (Vegetation Index * BBCH * GSD).

Secondly, a previously calibrated empirical regression model obtained from experimental field 2 was transferred to the third experimental field according to Equation 3. In this context, the best-

245 performing wheat model for the low GSD imagery and multispectral camera was chosen (Table 4,
246 BBCH 12, NDVI, 20 m; Slope = 0.17, Intercept = -15.82). By transferring a previously calibrated em-
247 pirical regression model to an independent data set with a different experimental design (genotypes,
248 sowing densities) and environmental conditions (season, soil color) it was possible to verify the
249 model robustness and the repeatability of the methodology at the same time.

$$UAV\ estimated\ plant\ density = \frac{\text{fractional cover} - \text{intercept}}{\text{slope}} \quad (3)$$

250 In addition, we again randomly divided the data, consisting of 126 samples, 63 times into a calibration
251 and validation data set for the third experimental field. Each time either 1/2, 1/3, 1/4, 1/6, 1/8, 1/11,
252 1/15 or 1/25 of the total number of samples was used for model calibration. Thus, it was possible to
253 identify the minimum number of reference measurements required to build up a user-based empiri-
254 cal regression model adapted to a specific experimental layout and environment. Each time, the re-
255 maining part of the data set was used to validate the performance of the model based on the averaged
256 R^2, RMSE and MAE. To identify significant deviations between the different sample sizes used for
257 the model calibration, one-way ANOVAs, including a Tukey post-hoc test, were conducted to verify
258 the MAE.
259

3. Results

3.1 Experimental field 1: Barley

The results of all individual models predicting the plant density of spring barley from UAV fractional cover values are summarized in Table 3. It can be seen that the spatial resolution had a substantial impact on ExGR model performance at the early stage of leaf development with only two unfolded leaves per plant (BBCH 12). The high GSD (0.20 cm) provided a model performance with an R^2 of 0.92 and an MAE of 26 plants m^{-2}. With decreasing GSD (0.59 cm), the correlation declined (R^2 = 0.70) with regression error metrics (RMSE, MAE) of more than 55 plants m^{-2} (Table 3). In contrast, the spatial resolution had less impact on model performance at BBCH stage 12 using the NDVI to predict the plant density. The results of the low GSD (1.38 cm) showed only slightly lower model performance (R^2 = 0.90, MAE = 29 plants m^{-2}) in comparison to the model constructed with the high GSD data (0.69 cm). This model provided the best performance with the highest R^2 of 0.92 and an MAE of 24 plants m^{-2} (Table 3).

Table 3. Statistical relationship between manually counted plant density and the fractional cover of barley at different BBCH stages, VIs, FAs and GSDs. Averaged statistical measures were computed across the empirical regression models randomly divided into calibration and validation data sets. Different letters indicate significant differences between MAE ($p < 0.05$). VI = vegetation index; FA = flight altitude; GSD = Ground sampling distance; R^2 = averaged coefficient of determination; RMSE = averaged root mean square error; MAE = mean absolute error; ExGR = excess green red; NDVI = normalized difference vegetation index.

BBCH	VI	FA	GSD	Illumi-	Calibration (n = 64)			Validation (n = 32)				
		(m)	(cm)	nation	Slope	Intercept	R^2	Slope	Intercept	R^2	RMSE	MAE
12	ExGR	10	0.20	(S)	0.051	-1.666	0.91	1.010	-0.553	0.92	34	26ab
		30	0.59	O	0.085	-6.632	0.68	1.031	-6.796	0.70	75	55h
	NDVI	10	0.69	(S)	0.089	-10.107	0.93	1.017	-5.217	0.92	32	24a
		20	1.38	S	0.105	-15.201	0.89	1.009	-3.436	0.90	37	29bc
13	ExGR	10	0.20	O	0.077	-1.556	0.92	1.007	-3.432	0.92	34	26ab
		30	0.59	O	0.139	-12.822	0.87	1.001	-0.766	0.89	40	31c
	NDVI	10	0.69	(O)	0.108	-7.756	0.90	1.010	-1.392	0.90	38	26ab
		20	1.38	O	0.153	-15.301	0.90	1.012	-1.727	0.90	37	26ab
14	ExGR	10	0.20	S	0.113	0.422	0.80	0.984	4.957	0.78	56	44g
		30	0.59	S	0.177	-22.158	0.84	0.987	5.034	0.83	48	34d
	NDVI	10	0.69	O	0.130	-3.319	0.85	1.018	-4.930	0.85	47	35d
		20	1.38	O	0.162	-8.280	0.83	1.022	-5.527	0.84	49	38ef
15	ExGR	10	0.20	S	0.111	24.883	0.84	0.998	0.789	0.83	50	41f
		30	0.59	S	0.147	21.045	0.82	0.999	-2.178	0.83	50	39ef
	NDVI	10	0.69	S	0.163	4.893	0.85	0.996	0.970	0.86	45	36de
		20	1.38	S	0.184	5.947	0.84	1.001	-0.997	0.85	47	39ef

At BBCH stage 13, the best validation performance was observed for the ExGR with uncertainties of less than 10% for both GSDs (GSD = 0.20 cm, R^2 = 0.92, MAE = 26; GSD = 0.59 cm, R^2 = 0.89, MAE = 31), also illustrated in Fig. 4. Moreover, BBCH stage 13 was as well the optimal development stage for the multispectral data and low GSD (NDVI, 0.69 cm), which led to an R^2 of 0.90 and MAE of 26 plants m^{-2} (Table 3). The regression line showed a good match to the 1:1 line, with almost no over- or underestimation of the predicted plant density (Fig. 4, a-d).

The model performance of both vegetation indices declined in the subsequent growth stages (BBCH 14, 15), where the leaves grow in size and the plants develop further leaves. The R^2 of the regression

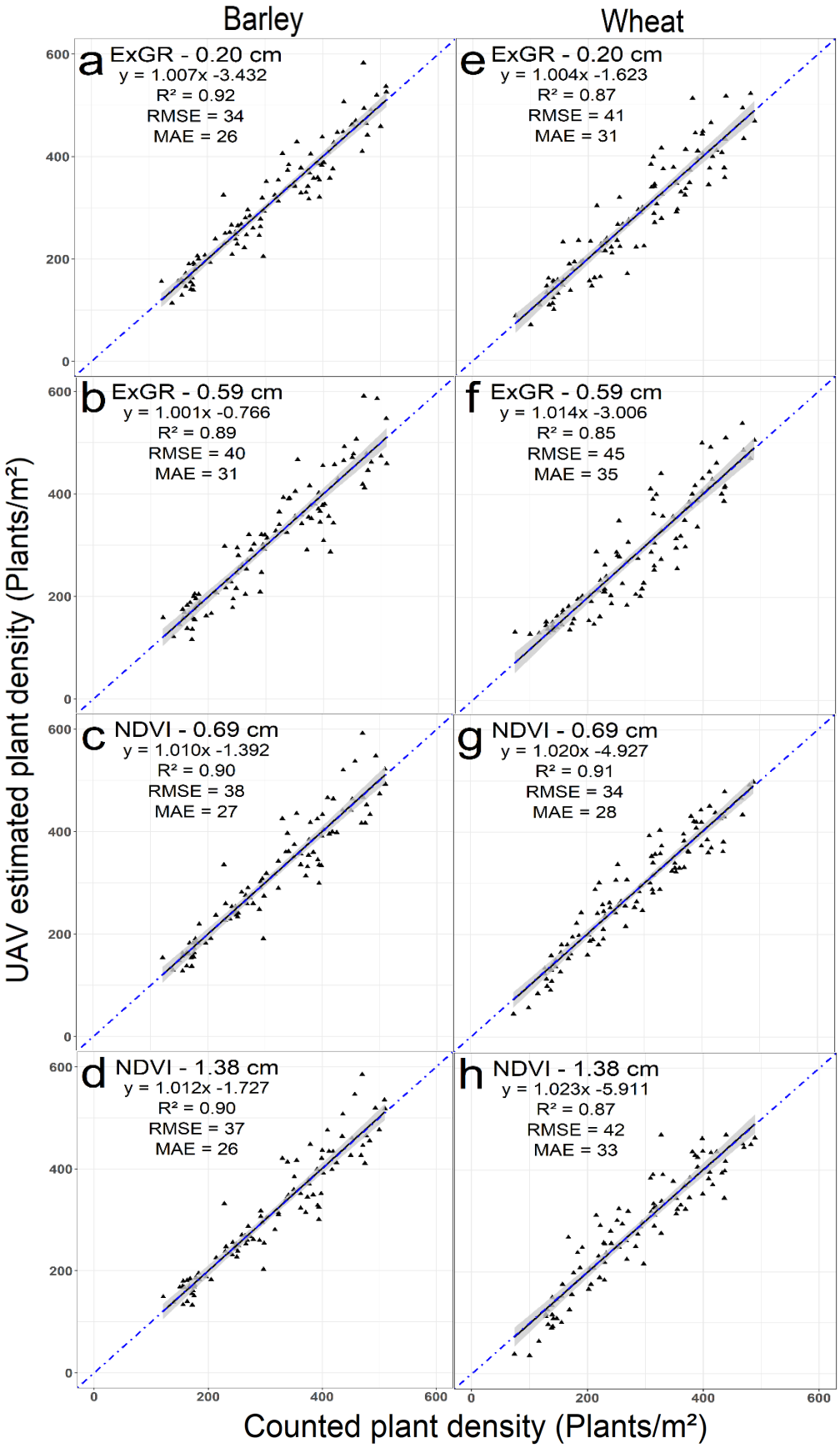
289 models, varying between 0.78 and 0.86, was still high, but the regression error metrics (RMSE, MAE)
290 of 34-44 plants m^{-2} indicated lower model accuracy (Table 3).
291 In general, a higher BBCH stage led to a higher slope in the calibration models (Table 3). This was
292 influenced by the fact that areas with a higher number of plants have a proportionally stronger in-
293 crease in fractional cover over time compared to areas with a lower number of plants. The slopes in
294 the validation models with values around one and intercepts of around zero for almost all BBCH
295 stages demonstrated the high prediction accuracy of the regression models.

296 **3.2 Experimental field 2: Wheat**
297 The model performance at early leaf development stages of spring wheat (BBCH 11, 12) was also
298 influenced by the spatial resolution. The ExGR calculated from the high GSD (0.20 cm) data led to a
299 significantly higher prediction accuracy compared to the models based on the low GSD (0.59 cm)
300 data (Table 4). The spatial resolution has less impact on model performance when the NDVI was used
301 for plant density assessment. The NDVI provided similar model accuracies for data sets recorded
302 from different GSDs within a BBCH stage. Even at BBCH stage 12, the model for multispectral data
303 and low GSD imagery (NDVI, 1.38 cm) already led to the best performance (R^2 = 0.89, MAE < 29
304 plants m^{-2}) with uncertainties of less than 10 % (Table 4). However, it can be observed that the model
305 made use of the image data with a lower GSD (1.38 cm) led to better results compared to the model
306 based on data with the higher GSD (0.69 cm). Considering the illumination conditions, the high GSD
307 data was acquired under sunny, while the low GSD data was recorded under cloudy conditions.

308 **Table 4.** Statistical relationship between manually counted plant density and the fractional cover of
309 wheat at different BBCH stages, VIs and FAs. Averaged statistical measures were computed across the
310 empirical regression models randomly divided into calibration and validation data sets. Different letters
311 indicate significant differences between MAE ($p < 0.05$).

BBCH	VI	FA	GSD	Illumi-	Calibration (n = 64)			Validation (n = 32)				
		(m)	(cm)	nation	Slope	Intercept	R^2	Slope	Intercept	R^2	RMSE	MAE
11	ExGR	10	0.20	S	0.029	1.394	0.81	1.007	-3.532	0.81	52	40fg
		30	0.59	S	0.036	-1.706	0.59	1.007	-4.064	0.60	87	67j
	NDVI	10	0.69	S	0.086	-7.358	0.83	1.000	0.414	0.83	50	39fg
		20	1.38	S	0.078	-11.406	0.80	1.010	-2.157	0.81	54	42g
12	ExGR	10	0.20	(O)	0.082	-8.013	0.83	1.013	-4.587	0.84	48	38ef
		30	0.59	O	0.089	-14.445	0.63	1.039	-11.496	0.63	86	62i
	NDVI	10	0.69	S	0.116	-5.648	0.84	1.011	-4.485	0.85	45	34cd
		20	1.38	O	0.171	-15.628	0.89	1.004	-3.493	0.89	37	29ab
13	ExGR	10	0.20	O	0.148	-4.750	0.87	1.004	-1.623	0.87	41	31bc
		30	0.59	O	0.219	-26.645	0.86	1.014	-3.006	0.85	45	35de
	NDVI	10	0.69	O	0.160	4.893	0.91	1.020	-4.927	0.91	34	28a
		20	1.38	O	0.182	6.681	0.87	1.023	-5.911	0.87	42	33bc
14	ExGR	10	0.20	S	0.103	17.981	0.83	0.960	10.483	0.82	50	38fg
		30	0.59	S	0.133	16.955	0.78	0.956	10.650	0.76	59	46h
	NDVI	10	0.69	S	0.180	1.230	0.90	1.019	-4.790	0.90	36	28a
		20	1.38	S	0.200	9.290	0.86	1.017	-4.812	0.87	43	34cd

312



313 **Fig. 4.** Linear regression between manually counted and UAV estimated plant density for different vege-
314 tation indices and GSDs obtained for BBCH stage 13 (n = 96). The UAV estimated plant density represents
315 an averaged number of plants computed by the randomly divided models for **(a-d)** barley and **(e-h)** wheat.
316 Black lines represent regression lines with 95% confidence intervals; blue lines represent 1:1 line.

Almost the same prediction accuracy was observed for the NDVI model constructed using the low GSD data (1.38 cm) at BBCH stage 13 (Fig. 4), while the higher GSD model provided the highest R^2 of 0.91 and the lowest MAE of 28 plants m^{-2} of all models (Table 4). The best model performances for the ExGR and both GSDs were also observed when the third leaf of the plants was unfolded (BBCH 13). Both models had an R^2 higher than 0.85 and MAE of less than 35 plants m^{-2} (Fig. 4). Similar to barley, the regression lines showed a good match with the 1:1 lines and only less scattering was visible (Fig. 4, e-h). At BBCH stage 14, the accuracy for the ExGR models clearly decreased. In contrast, the NDVI models led to similar performances as at the BBCH stage 13 (Table 4).
As already observed for barley, higher BBCH stages led to a higher slope value for the calibration models (Table 4). In contrast, these slope values were higher for wheat compared to barley (Table 3). The slopes of the validation models had values close to one, which illustrated the high prediction accuracy of the generated models (Table 4). However, the performance evaluation for wheat led to a slightly lower R^2 and higher regression error metrics compared to barley.

3.3 Ideal parameters for data acquisition

The three-way ANOVA for the MAE was calculated for barley (Additional file 1: Table S1, Table S2) and wheat (Additional file 1: Table S3, Table S4) to identify the ideal parameters for plant density assessment and data acquisition. In the case of barley the vegetation index, the BBCH stage and the interaction between BBCH and GSD explained most of the variation in the methodology (Additional file 1: Table S1). The ANOVA results identified in general the NDVI as the vegetation index, BBCH 13 as the development stage and the high GSD as the ideal parameter for data acquisition and plant density assessment of barley (Additional file 1: Table S2)
For wheat, the vegetation index clearly explained most of the variation (Additional file 1: Table S3). The GSD also had a high impact on the plant density assessment. The interaction between BBCH and GSD was the third most important factor. Despite the investigated differences of barley and wheat, the ANOVA of wheat identified the same predictors (NDVI, BBCH 13, high GSD) for data acquisition and plant density assessment (Additional file 1: Table S4).

3.4 Experimental field 3: Investigate repeatability

In the case of the third experiment, two investigations were conducted to evaluate the practicability and repeatability of the developed methodology. Firstly, the best-performing empirical regression model for wheat and the low GSD (Table 4, BBCH 12, NDVI, 20 m; slope = 0.177, intercept = -15.628) was applied to the data of experimental field 3 according to Equation 3. The best-performing model of the lower GSD was considering instead of the ideal model performance to increase the practicability of the methodology. The experiments differed with regard to season, soil color, genotypes and sowing density. Especially with its 42 genotypes and the large number of variations in plant characteristics, the third experimental field was well suited for testing the repeatability and practicability of the previously developed regression model.
The transfer of the calibrated empirical regression model for UAV based plant density prediction yielded in a high prediction accuracy. The determined MAE was lower than 20 plants m^{-2} and the R^2 of 0.83 was relatively high (Fig. 5a). The regression line showed a good match with the 1:1 line and only lower values were slightly overestimated.
Secondly, the size of the calibration data subset was reduced stepwise to determine the minimum number of field measurements necessary for a user-based calibration of a new empirical regression model (Table 5). The starting calibration subset consisting of 63 samples also provided the highest accuracy with an MAE of less than 21 plants m^{-2} and an R^2 of 0.83. The prediction accuracy of the user-based calibration model based on the 63 training samples and the corresponding scatterplot (Fig. 5b) were almost similar compared to the transferred calibrated empirical regression model from experimental field 2 (Fig. 5a).

Table 5. Analysis results of the number of field measurements necessary to calibrate an empirical regression model to a specific layout or environment (n = 126). The calibration sample size was reduced stepwise from 63 to five reference measurements, while the remaining samples were used to validate the performance of the calibrated models. Averaged statistical measures were computed across the empirical regression models randomly divided into calibration and validation data sets for each sample size. Different letters indicate significant differences between MAE ($p < 0.05$).

Experimental field 3									
Calibration				Validation					
n	Slope	Intercept	R^2	n	Slope	Intercept	R^2	RMSE	MAE
63	0.169	-14.868	0.83	63	1.005	-2.246	0.83	25	21a
42	0.169	-14.955	0.83	84	1.008	-2.976	0.83	25	21a
32	0.169	-14.972	0.82	94	1.009	-2.885	0.83	25	21a
21	0.168	-14.446	0.81	105	1.024	-10.355	0.83	26	22a
16	0.168	-14.441	0.81	110	1.026	-11.117	0.83	27	22a
11	0.165	-12.837	0.81	115	1.063	-26.407	0.83	28	23ab
8	0.165	-13.404	0.78	118	1.086	-35.029	0.83	31	26bc
5	0.161	-11.325	0.79	121	1.154	-66.075	0.83	36	29c

The continuous reduction of the calibration sample size from 63 to 11 samples led to no significant changes in model performance (MAE < 23 plants m^{-2}). The slope and intercept values of the validation regression functions slightly increased with decreasing size of the calibration samples (Table 5). However, the residual deviations and the 95 confidence intervals were quite uniformly independent of the calibration sample sizes (Fig. 5b-d). The last two reduction steps from 11 to only eight and five samples led to significantly lower model performances (MAE > 26 plants m^{-2}) (Table 5). Additionally, the corresponding scatterplots (Fig. 5e,f) illustrated an underestimation of lower values and an overestimation of higher values.

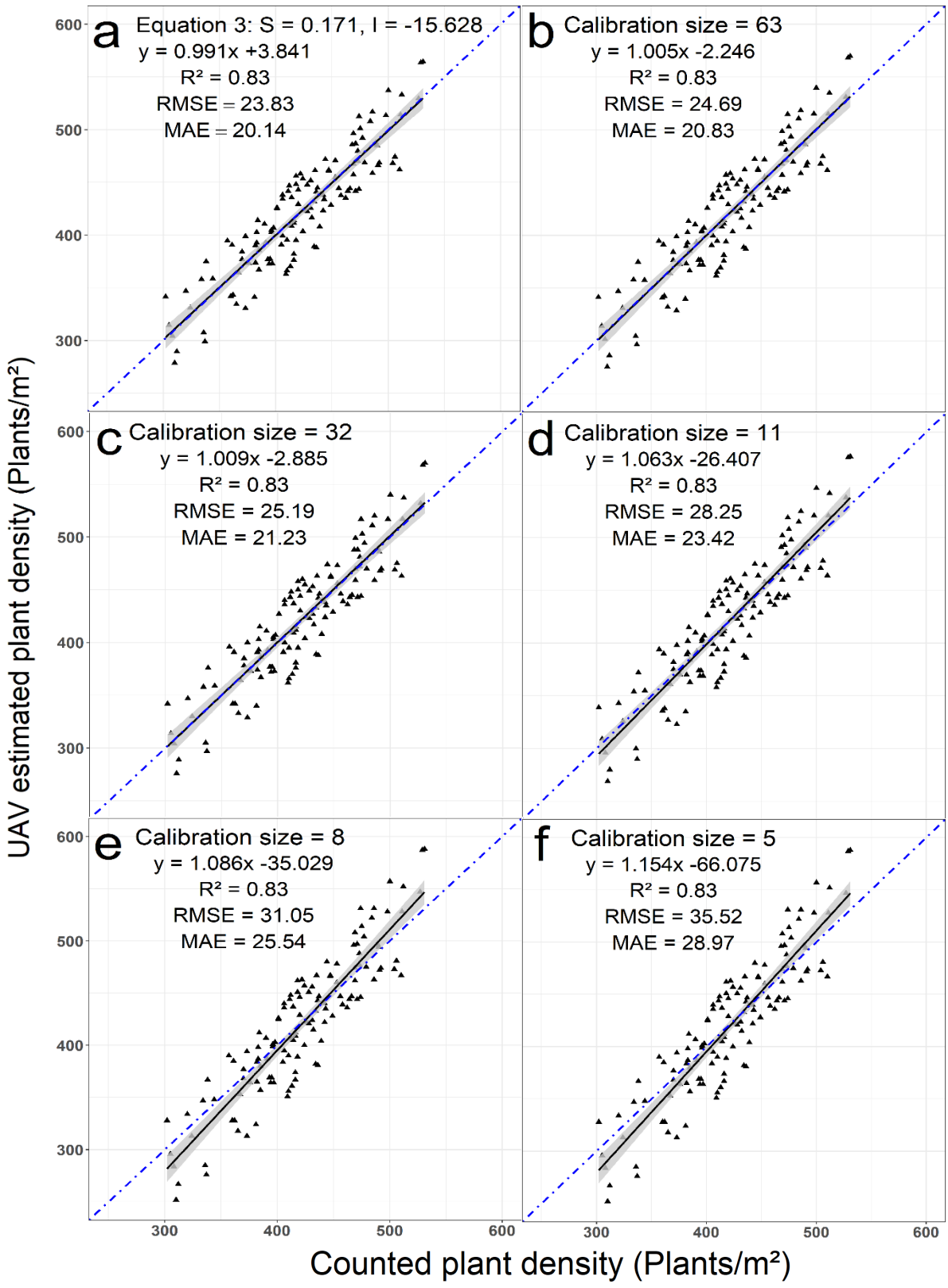


Fig. 5. Linear regression between manually counted wheat plant densities and those estimated from UAV data at BBCH stage 12 for experimental field 3 with two investigations (n = 126) **a** The empirical regression model from experimental field 2 was transferred to the data of experimental field 3 **b-f** Investigation of the necessary number of field measurements for a user-based calibration of an empirical regression model based on stepwise reduction of calibration size to **b** 63 samples; **c** 32 samples; **d** 11 samples; **e** 8 samples and; **f** 5 samples. The UAV estimated plant density represents an averaged number of plants computed by the randomly divided models. Black lines represent regression lines with 95% confidence intervals; blue lines represent 1:1 lines. S = slope; I = intercept.

4. Discussion

In the study presented here, UAV data were evaluated for high-throughput field phenotyping of plant density. The results demonstrated that plant density could be predicted with potential uncertainties of less than 10%. In the following, the introductory questions are discussed in detail.

BBCH stage 13 was identified as the plant developmental stage best suited for UAV data acquisition (Additional file 1: Table S1-S4). At this growth stage, the MAEs of the validation models were between 25 and 34 plants m^{-2} and therefore similar compared to previous studies focusing on wheat [21, 42, 43]. In addition, we observed a high prediction accuracy at BBCH stage 12 when plant density prediction was based on the data recorded from the higher GSD (Ø MAE < 30 plants m^{-2}) (Table 2, 4). In summary, the plants or the resolution needs to be large enough to detect the fractional cover accurately. Even with four unfolded leaves per plant (BBCH 14) predictions of plant density were possible with MAEs ranging from 28 to 46 plants m^{-2}. However, the model accuracy partly declined at BBCH stage 14 and 15 compared to the previous stages. The leaves grow in size and the plants develop additional leaves (Fig 2b), resulting in a higher overlap between neighboring plants.
In other studies, the investigated methodology was limited to a specific growth stage [21, 22, 42–44]. The approach presented here, however, allowed plant density to be estimated from UAV data acquired during BBCH stages 12 to 14. This makes the proposed method more practical, since UAV data acquisition is possible in a longer time window.

The models constructed using the higher GSD data enabled a higher prediction accuracy of plant density. However, especially at the later stages of plant development (BBCH 13 onwards), the differences in the predictions between the models for the GSDs were not great. At BBCH stage 13, the averaged MAE was lower than 31 plants m^{-2} for the low GSD image data (ExGR-30 m, NDVI-20 m) and had an almost similar prediction accuracy compared to the high GSD (ExGR-10 m, NDVI-10 m, Ø MAE < 28 plants m^{-2}). In principle, the prediction accuracy was influenced less by the spatial resolution using the NDVI for plant density assessment (Table 2, Table 4). We conclude that multispectral data recorded with a lower GSD than 1.38 cm may predict a similar accuracy to increase the throughput of the approach.
In comparison to other studies [21, 22, 42, 43], the present methodology is not limited to spatially very highly resolved image data (0.02 cm) in order to obtain precise estimates of plant density. The results showed that for ideal conditions RGB data with 0.6 cm and multispectral data with 1.4 cm spatial resolution were sufficient to predict the plant density of numerous plots, greatly enhance the throughput of the approach. In addition, the automatic image acquisition of the UAV and the use of orthomosaics instead of individual images [22, 42, 43] are ideally for large scales and decision-making process in precision farming and breeding.

Cereals such as wheat and barley have comparable plant structures, especially at the early leaf developmental stages. However, some individual characteristics such as the leaf arrangement can be different. The results stressed that it was essential to distinguish between planophile (wheat) and erectophile (barley) leaves in the process of developing empirical regression models having the capability to estimate plant density precisely. In a nadir UAV image, a leaf with horizontal orientation (planophile) covers a larger soil area and consists of a higher number of green pixels in comparison to a leaf with vertical orientation (erectophile) (Fig. 1 c,d). This results in higher fractional cover in the case of the observation of planophile cereals. For that reason, the wheat calibration models had steeper slopes compared to the barley calibration models. The impact of the leaf arrangement on the methodology was not investigated in previous studies focusing on the assessment of plant density [21, 22, 42–44, 60].

The comparison of the sensors revealed that the multispectral camera had a higher prediction accuracy compared to the RGB sensor (Additional file 1: Table S2, Table S4). Especially for the prediction of wheat plant density, the vegetation index was the most important factor (Additional file 1: Table S3). The predominantly horizontally oriented leaves of wheat caused more mixed pixels. The additional near- infrared spectral band used for the calculation of the NDVI is advantageous for distinguishing plant from soil pixels. Furthermore, the radiometric correction of the multispectral data based on the solar irradiance sensor and reference panel enabled to convert the raw pixel information

into absolute spectral reflectance. This made the method more robust to varying illumination conditions among data sets and during data acquisition (Table 1). Based on these findings, multispectral data are recommended in particular for genotypes with planophile leaves.

In order to test the repeatability and robustness of the approach, in particular with regard to genotypic varieties, the calibrated empirical regression model constructed for spring wheat was transferred to an independent breeding experiment with 42 different winter wheat genotypes. Visual field observation revealed only little variability among genotypes in the early leaf developmental stages. This observation was similar to the findings of Jin et al. [43]. The transferred regression model, therefore, provided good prediction accuracy across the genotypes ($R^2 = 0.83$, MAE < 20 plants m^{-2}). The large variation in the training dataset resulted in a robust model, which avoids the need for further genotypic model adaptation. For this reason, a specific re-calibration, which is usually required for machine-learning techniques [42, 43, 60], was not obligatory using this approach and comparable conditions in this study.

Estimating the required number of reference measurements in the field to calibrate a user-based empirical regression model is needed in the case of different seed row spacing, illumination conditions and different spatial resolution of the UAV data. For this user-based calibration, 11 randomly selected field measurements were already sufficient in this study to train a robust model (MAE < 23 plants m^{-2}). No significant differences in prediction accuracy were observed by using a small training subset of 11 samples instead of a larger one consisting of 63 samples (Table 5, Fig. 5). Nevertheless, attention has to be paid to outliers when such a small calibration sample size is used. Moreover, some variation of plant density in the field, which covers a range of possible values, is needed to adequately train a user-based empirical regression model. The required time to manually count the plants in an area of one square meter with a sowing density of 460 seeds m^{-2} is approximately 5 minutes.

In summary, the prediction accuracy of the transferred regression model (Fig. 5a) was closely related to the user-based calibrated model (Fig. 5b-e), although the experiments differed by season, soil color, genotypes and sowing density. Therefore, we assumed that user-based calibration is not necessary as long as the conditions are comparable, such as seed row spacing, the spatial resolution of the UAV data and the illumination conditions during data acquisition. In this respect, it is important to note that the empirical regression model was calibrated and transferred with data acquired under overcast conditions. Hence, the transferability of the model was not affected by the position of the sun. It is important to note that a calibrated model should be transferred to data with comparable illumination conditions. This is an important prerequisite for the transferability of empirical regression models [61]. Therefore, we recommend that data should be primarily acquired under overcast conditions or as an alternative timely close to solar noon to enable high transferability of the calibrated model. Under direct sunlight, data acquisition around solar noon conditions has the benefit of minimizing shadow effects and enabling comparable illumination conditions between datasets.

The investigation of repeatability focused on wheat with predominantly planophile leaves, higher regression error metrics and high sowing density of up to 500 plants m^{-2} to demonstrate the transferability and robustness of the calibrated regression model. The prediction capability in this study was not negatively affected at early leaf developmental stages by high sowing densities of up to 500 plants with a higher probability of overlapping leaves [21, 42, 43]. However, a weed-free crop stand is required to achieve a sufficient prediction accuracy, since weeds have a negative impact on the assessment of fractional cover. Furthermore, local soil heterogeneities or sowing failures need to be identified and subsequently considered to enable a successful differentiation between different varieties based on sowing density.

5. Conclusions

The study presents a novel way to assess the plant density of cereal crops with high-throughput UAV image data. The determination of fractional cover for plant density assessment takes advantage of the fact that the average plant size is relatively stable and the number of overlapping plants is relatively low at early leaf developmental stages. BBCH stage 13 was identified as the ideal growth stage when the plants were large enough to detect the fractional cover accurately even with lower GSDs in

this study. Nevertheless, with appropriate parameters such as a high GSD for BBCH stage 12, the specific growth stage was not critical for UAV based plant density assessment.
The calibrated models proved to be robust with respect to unknown sites with comparable conditions such as seed row spacing, spatial resolution and illumination conditions. Across a data set of 42 genotypes, it was shown that genotypic model adaptation is not necessary. It is only essential to distinguish between planophile and erectophile cereals in the application and process of empirical regression models. Better results were observed when multispectral data were used for plant density quantification, in particular for planophile cereals. However, also RGB data provided comparable model performance and was sufficient for plant density assessment.
The results stressed the transferability, practicability and repeatability of the developed methodology. With conditions comparable to the presented study, the operator can choose whether to apply the trained regression model without re-calibration or whether to perform a user-based approach with a small number of field measurements. Furthermore, also a broader application of the method is conceivable, e.g. for other cereals with heterogeneous germination and overlapping leaves such as oats or triticale. In contrast to machine-learning approaches, expert guidance in feature extraction and large training datasets are not necessary.
The developed methodology facilitates high-throughput plant density assessment of cereal crops for decision-making in precision farming and breeding. Breeding trials are particularly difficult to monitor within a reasonable time, resulting in an increasing need for faster selection of superior lines. In addition, considering the plant density variation caused, for example, by different soil or nutrition conditions, can increase farmer's crop production and yield estimations.

Authors' contributions

All authors have made significant contributions to this research. NW and RP conceived and designed the experiments; NW and VK performed the validation; NW performed the data acquisition; OM and UR acquired funding, NW, BS, JAP processed and analyzed the data; BS, OM, JAP, UR supervision; NW, BS, JAP and VK wrote and edited the paper. All authors read and approved the final manuscript.

Author details

[1] Institute of Bio- and Geosciences, Plant Sciences (IBG-2), Forschungszentrum Jülich GmbH, 52428 Jülich, Germany
[2] Field Lab Campus Klein, Altendorf, University of Bonn, 53359 Rheinbach, Germany
[3] Institute for Crop Science and Resource Conservation, Renewable Resources, University of Bonn, 53115 Bonn, Germany

Acknowledgements

We thank Dr. Beate Doerffel for her statistical advice. Part of this work was supported by the 'Strukturwandel-Projekt BioökonomieREVIER' (project identification number 031B0918A). Additionally, this work has partially been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2070 – 390732324.

Declaration of Competing Interest

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

Funding

This study was performed within the German Plant Phenotyping Network (DPPN), which is funded by the German Federal Ministry for Education and Research (BMBF), project identification number BMBF 031A053. The work was additional supported by SPECTORS, which is funded by INTERREG V A-Programm Deutschland-Nederland, project identification number 143081.

References

1. Yang G, Liu J, Zhao C, Li Z, Huang Y, Yu H, et al. Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. Frontiers in Plant Science. 2017;8.
2. Zhang C, Kovacs JM. The application of small unmanned aerial systems for precision agriculture: A review. Precis Agric. 2012;13:693–712.
3. Vargas JQ, Bendig J, Arthur A Mac, Burkart A, Julitta T, Maseyk K, et al. Unmanned Aerial Systems (UAS)-Based Methods for Solar Induced Chlorophyll Fluorescence (SIF) Retrieval with Non-Imaging Spectrometers: State of the Art. Remote Sens. 2020.
4. Candiago S, Remondino F, Giglio M De, Dubbini M, Gattelli M. Evaluating Multispectral Images and Vegetation Indices for Precision Farming Applications from UAV Images. Remote Sens. 2015;7 Vi:4026–47.
5. Colomina I, Molina P. Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing. 2014.
6. Primicerio J, Di Gennaro SF, Fiorillo E, Genesio L, Lugato E, Matese A, et al. A flexible unmanned aerial vehicle for precision agriculture. Precis Agric. 2012.
7. Hunt ER, Hively WD, Daughtry CST, Mccarty GW, Fujikawa SJ, Ng TL, et al. Remote Sensing of Crop Leaf Area Index Using Unmanned Airborne Vehicles. In: Proceedings of the Pecora 17 Symposium. 2008.
8. Siegmann B, Jarmer T. Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. Int J Remote Sens. 2015.
9. Verger A, Vigneau N, Chéron C, Gilliot JM, Comar A, Baret F. Green area index from an unmanned aerial system over wheat and rapeseed crops. Remote Sens Environ. 2014.
10. Holman FH, Riche AB, Michalski A, Castle M, Wooster MJ, Hawkesford MJ. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. Remote Sens. 2016;8.
11. Madec S, Baret F, De Solan B, Thomas S, Dutartre D, Jezequel S, et al. High-throughput phenotyping of plant height: Comparing unmanned aerial vehicles and ground lidar estimates. Front Plant Sci. 2017.
12. Bendig J, Bolten A, Bareth G. UAV-based Imaging for Multi-Temporal, very high Resolution Crop Surface Models to monitor Crop Growth Variability. Photogramm - Fernerkundung - Geoinf. 2013;6 December:551–62. doi:10.1127/1432-8364/2013/0200.
13. Bendig J, Yu K, Aasen H, Bolten A, Bennertz S, Broscheit J, et al. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int J Appl Earth Obs Geoinf. 2015;39:79–87. doi:10.1016/j.jag.2015.02.012.
14. Hunt ER, Cavigelli M, Daughtry CST, McMurtrey JE, Walthall CL. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precis Agric. 2005.
15. Kross A, McNairn H, Lapen D, Sunohara M, Champagne C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. Int J Appl Earth Obs Geoinf. 2015.
16. Wilke N, Siegmann B, Klingbeil L, Burkart A, Kraska T, Muller O, et al. Quantifying lodging percentage and lodging severity using a UAV-based canopy height model combined with an objective threshold approach. Remote Sens. 2019.
17. Chu T, Starek MJ, Brewer MJ, Murray SC, Pruter LS. Assessing lodging severity over an

585 experimental maize (Zea mays L.) field using UAS images. Remote Sens. 2017;9:1–24.
586 18. Lawles K, Raun W, Desta K, Freeman K. Effect of delayed emergence on corn grain yields. J Plant
587 Nutr. 2012.
588 19. Olsen J, Kristensen L, Weiner J. Influence of sowing density and spatial pattern of spring wheat
589 (Triticum aestivum) on the suppression of different weed species. Weed Biol Manag. 2006.
590 20. Olsen J, Weiner J. The influence of Triticum aestivum density, sowing pattern and nitrogen
591 fertilization on leaf area index and its spatial variation. Basic Appl Ecol. 2007.
592 21. Liu S, Baret F, Allard D, Jin X, Andrieu B, Burger P, et al. A method to estimate plant density and
593 plant spacing heterogeneity: Application to wheat crops. Plant Methods. 2017;:1–11.
594 22. Aich S, Josuttes A, Ovsyannikov I, Strueby K, Ahmed I, Duddu HS, et al. DeepWheat: Estimating
595 Phenotypic Traits from Crop Images with Deep Learning. In: Proceedings - 2018 IEEE Winter
596 Conference on Applications of Computer Vision, WACV 2018. 2018.
597 23. Joseph KDSM, Alley MM, Brann DE, Gravelle WD. Row Spacing and Seeding Rate Effects on
598 Yield and Yield Components of Soft Red Winter Wheat1. Agron J. 1985.
599 24. Whaley JM, Sparkes DL, Foulkes MJ, Spink JH, Semere T, Scott RK. The physiological response of
600 winter wheat to reductions in plant density. Ann Appl Biol. 2000.
601 25. Jin X, Li Z, Yang G, Yang H, Feng H, Xu X, et al. Winter wheat yield estimation based on multi-
602 source medium resolution optical and radar imaging data and the AquaCrop model using the particle
603 swarm optimization algorithm. ISPRS J Photogramm Remote Sens. 2017.
604 26. Valério IP, De Carvalho FIF, De Oliveira AC, Benin G, De Souza VQ, De Almeida Machado A, et
605 al. Seeding density in wheat genotypes as a function of tillering potential. Sci Agric. 2009.
606 27. Postma JA, Hecht VL, Hikosaka K, Nord EA, Poorter PH. Dividing the pie : A quantitative review
607 on plant density responses. 2020;:0–2.
608 28. Kirby EJM. Effect of sowing depth on seedling emergence, growth and development in barley and
609 wheat. F Crop Res. 1993.
610 29. Lindstrom MJ, Papendick RI, Koehler FE. A Model to Predict Winter Wheat Emergence as
611 Affected by Soil Temperature, Water Potential, and Depth of Planting1. Agron J. 1976.
612 30. Winter SR, Musick JT, Porter KB. Evaluation of Screening Techniques for Breeding Drought-
613 Resistanct Winter Wheat. Crop Sci. 1988.
614 31. Murungu FS. Effects of seed priming and water potential on seed germination and emergence of
615 wheat (Triticum aestivum L.) varieties in laboratory assays and in the field. African J Biotechnol. 2011.
616 32. Steiner JJ, Grabe DF, Tulo M. Single and Multiple Vigor Tests for Predicting Seedling Emergence
617 of Wheat. Crop Sci. 1989.
618 33. Rajala A, Niskanen M, Isolahti M, Peltonen-Sainio P. Seed quality effects on seedling emergence,
619 plant stand establishment and grain yield in two-row barley. Agric Food Sci. 2011.
620 34. Al-Mulla YA, Huggins DR, Stöckle CO. Modeling the emergence of winter wheat in response to
621 soil temperature, water potential, and planting depth. Trans ASABE. 2014.
622 35. Araus JL, Cairns JE. Field high-throughput phenotyping: The new crop breeding frontier. Trends
623 in Plant Science. 2014.
624 36. Furbank RT, Tester M. Phenomics - technologies to relieve the phenotyping bottleneck. Trends in
625 Plant Science. 2011.
626 37. Tang L, Tian LF. Plant identification in mosaicked crop row images for automatic emerged corn
627 plant spacing measurement. Trans ASABE. 2008;51:2181–91.

628 38. Shrestha DS, Steward BL. Automatic corn plant population measurement using machine vision.
629 Trans Am Soc Agric Eng. 2003;46:559–65.
630 39. Li B, Xu X, Han J, Zhang L, Bian C, Jin L, et al. The estimation of crop emergence in potatoes by
631 UAV RGB imagery. Plant Methods. 2019;15:1–13.
632 40. Sankaran S, Quirós JJ, Knowles NR, Knowles LO. High-Resolution Aerial Imaging Based
633 Estimation of Crop Emergence in Potatoes. Am J Potato Res. 2017;94:658–63.
634 41. Lottes P, Khanna R, Pfeifer J, Siegwart R, Stachniss C. UAV-based crop and weed classification
635 for smart farming. In: Proceedings - IEEE International Conference on Robotics and Automation.
636 2017.
637 42. Liu S, Baret F, Andrieu B, Burger P, Hemmerlé M. Estimation of wheat plant density at early
638 stages using high resolution imagery. Front Plant Sci. 2017;8 May:1–10.
639 43. Jin X, Liu S, Baret F, Hemerlé M, Comar A. Estimates of plant density of wheat crops at emergence
640 from very low altitude UAV imagery. Remote Sens Environ. 2017;198:105–14.
641 44. Sankaran S, Khot LR, Carter AH. Field-based crop phenotyping: Multispectral aerial imaging for
642 evaluation of winter wheat emergence and spring stand. Comput Electron Agric. 2015;118:372–9.
643 doi:10.1016/j.compag.2015.09.001.
644 45. Meier U. Growth stages of mono- and dicotyledonous plants. 2001.
645 46. Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. "Structure-from-Motion"
646 photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology. 2012.
647 47. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis. 2004.
648 48. Agisoft. MicaSense RedEdge MX processing workflow (including Reflectance Calibration) in
649 Agisoft Metashape Professional 1.5 : Helpdesk Portal. Agisoft Helpdesk Portal. 2019.
650 49. Neto JC. a combined statistical-soft computing approach for classification and mapping weed
651 species in minimum-tillage systems. Cornell Hotel Restaur Adm Q. 2010.
652 50. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote
653 Sens Environ. 1979;8:127–50.
654 51. Otsu N. THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS. IEEE Trans
655 Syst Man Cybern. 1979.
656 52. López-Granados F, Torres-Sánchez J, Serrano-Pérez A, de Castro AI, Mesas-Carrascosa FJ, Peña
657 JM. Early season weed mapping in sunflower using UAV technology: variability of herbicide
658 treatment maps against weed thresholds. Precis Agric. 2016.
659 53. Torres-Sánchez J, López-Granados F, Peña JM. An automatic object-based method for optimal
660 thresholding in UAV images: Application for vegetation detection in herbaceous crops. Comput
661 Electron Agric. 2015;114 November 2016:43–52.
662 54. Bassine FZ, Errami A, Khaldoun M. Vegetation Recognition Based on UAV Image Color Index.
663 In: Proceedings - 2019 IEEE International Conference on Environment and Electrical Engineering and
664 2019 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2019. 2019.
665 55. Marcial-Pablo M de J, Gonzalez-Sanchez A, Jimenez-Jimenez SI, Ontiveros-Capurata RE, Ojeda-
666 Bustamante W. Estimation of vegetation fraction using RGB and multispectral images from UAV. Int
667 J Remote Sens. 2019.
668 56. Verrelst J, Camps-Valls G, Muñoz-Marí J, Rivera JP, Veroustraete F, Clevers JGPW, et al. Optical
669 remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review. ISPRS
670 Journal of Photogrammetry and Remote Sensing. 2015.

671 57. core Team R. R: A Language and Environment for Statistical Computing. R Found Stat Comput
672 Vienna, Austria. 2018.
673 58. Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments
674 against avoiding RMSE in the literature. Geosci Model Dev. 2014.
675 59. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean
676 square error (RMSE) in assessing average model performance. Clim Res. 2005.
677 60. Koh JCO, Hayden M, Daetwyler H, Kant S. Estimation of crop plant density at early mixed growth
678 stages using UAV imagery. Plant Methods. 2019.
679 61. Dorigo WA, Zurita-Milla R, de Wit AJW, Brazile J, Singh R, Schaepman ME. A review on reflective
680 remote sensing and data assimilation techniques for enhanced agroecosystem modeling. Int J Appl
681 Earth Obs Geoinf. 2007.
682