

RESEARCH

Open Access



# Field-measured canopy height may not be as accurate and heritable as believed: evidence from advanced 3D sensing

Jingrong Zang<sup>1</sup>, Shichao Jin<sup>1\*</sup>, Songyin Zhang<sup>1</sup>, Qing Li<sup>1</sup>, Yue Mu<sup>1</sup>, Ziyu Li<sup>1</sup>, Shaochen Li<sup>1</sup>, Xiao Wang<sup>1</sup>, Yanjun Su<sup>2</sup> and Dong Jiang<sup>1</sup>

## Abstract

Canopy height (CH) is an important trait for crop breeding and production. The rapid development of 3D sensing technologies shed new light on high-throughput height measurement. However, a systematic comparison of the accuracy and heritability of different 3D sensing technologies is seriously lacking. Moreover, it is questionable whether the field-measured height is as reliable as believed. This study uncovered these issues by comparing traditional height measurement with four advanced 3D sensing technologies, including terrestrial laser scanning (TLS), backpack laser scanning (BLS), gantry laser scanning (GLS), and digital aerial photogrammetry (DAP). A total of 1920 plots covering 120 varieties were selected for comparison. Cross-comparisons of different data sources were performed to evaluate their performances in CH estimation concerning different CH, leaf area index (LAI), and growth stage (GS) groups. Results showed that 1) All 3D sensing data sources had high correlations with field measurement ( $r > 0.82$ ), while the correlations between different 3D sensing data sources were even better ( $r > 0.87$ ). 2) The prediction accuracy between different data sources decreased in subgroups of CH, LAI, and GS. 3) Canopy height showed high heritability from all datasets, and 3D sensing datasets had even higher heritability ( $H^2 = 0.79-0.89$ ) than FM (field measurement) ( $H^2 = 0.77$ ). Finally, outliers of different datasets are analyzed. The results provide novel insights into different methods for canopy height measurement that may ensure the high-quality application of this important trait.

## Highlights

- The effects of canopy height, leaf area index, and growth stage on the accurate monitoring of canopy height with different 3D sensors were systematically evaluated.
- Field-measured canopy height may not be as accurate as believed, especially in the plots with higher canopy height and at later growth stages.
- 3D sensing methods achieved higher heritable canopy height estimation than field measurement.

**Keywords** Canopy height, Comparison, Field measurement, Terrestrial laser scanning, Backpack laser scanning, Gantry laser scanning, Digital aerial photogrammetry

\*Correspondence:

Shichao Jin

jschaon@njau.edu.cn; jinshichao1993@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Canopy height (CH) is an important and heritable agronomic trait for breeding and field management [34]. Breeders have paid much effort to selecting the ideal plant height to maximize light interception, increase yield [43], enhance lodging resistance [57, 77], and facilitate mechanical harvesting. Agronomists often use CH to indicate the growth of other complicated and difficultly accessible traits, such as phenology [72], leaf area index (LAI) [11], and biomass [50]. Therefore, high-throughput and accurate evaluation (e.g., ensuring high heritability) of CH are critical for accelerating crop breeding and production.

Traditional CH estimation methods mainly use rulers by selecting a few representative positions within a canopy. Manual measurement is time-consuming, labor-intensive, tedious, and error-prone due to subjective selection and visual observation. However, it is still the most widely adopted way due to its visibility and reliability during the past decades. Recently, many studies have demonstrated CH can be efficiently acquired from advanced three-dimensional (3D) sensing techniques [28, 30, 43, 59, 68]. It brings us naturally to a fundamental and essential question: are 3D sensing techniques as accurate as field measurement?

Recent studies have explored the applicability of some mainstream 3D sensing techniques for CH measurements in agriculture, including LiDAR (light detection and management) and multi-view images [25]. LiDAR is an active sensing technology that records 3D structure information of objects by measuring the distance with the laser [8, 30]. LiDAR has many advantages, including (1) strong penetration ability that can characterize the inner structure of the canopy, (2) real and direct 3D characterization of an object without a complicated reconstruction process, and (3) insensitive to illumination. According to different mounting platforms, LiDAR systems used for crop height measurement mainly include terrestrial laser scanning (TLS) [14, 64], backpack laser scanning (BLS) [78], gantry laser scanning (GLS) [36, 61], and unmanned-aerial-vehicle laser scanning (ULS) [40, 54, 77]. In contrast to the active LiDAR sensing technologies, passive sensing-based methods (e.g., Multi-view images) can also measure 3D structure through methods like structure from motion (SfM) [3, 18, 45, 69]. Among the passive sensing-based techniques, digital aerial photogrammetry (DAP) is one of the most popular ways for field CH estimation due to its low cost, high efficiency, and high accuracy comparable to ULS [17, 21, 75, 76]. These 3D sensing techniques have been successfully applied to CH measurement, including the adoption of TLS for accurate height measurement of maize ( $R^2=0.93$ ) [64], cotton ( $R^2=0.97$ ) [60], rice ( $R^2=0.91$ )

[63], barley ( $R^2=0.95$ ), pea ( $R^2=0.93$ ), and bean ( $R^2=0.91$ ) [9], the use of BLS for efficient height measurement of large-scale wheat [78] and forest [22, 32, 58]; the exploration of ULS for estimating CH of sugar beet ( $R^2=0.70$ ), wheat ( $R^2=0.78$ ), and potato ( $R^2=0.50$ ) [24], and DAP for measuring corn CH ( $R^2=0.78$ ) [57]. In all, current studies demonstrated that TLS and BLS usually performed better than ULS and DAP due to their close range of sensing, and the accuracy of DAP was comparable to ULS (Additional file 1: Table S1).

In addition to the exploration of high estimation accuracy, more and more studies are attempting to explore the genetic bases (e.g., heritability) of high-throughput phenotype [56, 62, 73]. CH is a high heritability trait, as effective as yield [53]. Higher heritability indicates that the environment has less influence on the trait, and further describes the value of breeding [5, 48]. Several studies have already verified the potential of CH from many 3D sensing platforms, including the use of LiDAR [33, 68] and UAV imagery [67]. Interestingly, recent studies declared 3D sensing-derived CH showed better heritability than field measurement. For example, Madec et al. [43] proved high heritability values ( $H^2>0.90$ ) of CH derived from both LiDAR and DAP; Volpato et al. [67] compared the height heritability from UAV imagery ( $H^2=0.71-0.97$ ) and field measurement ( $H^2=0.62-0.96$ ) across four different growth stages (GS), which showed the UAV imagery had better heritability. These novel studies inspire us to rethink a questionable and challenging question: is field-measured CH as accurate and heritable as believed?

Some critical discussions about the accuracy of field-measured CH have been raised in recent years. On the one hand, field measurements are believed as accurate benchmarks. For example, Wang et al. [70] found the heights measured by the LiDAR-Lite v2, the Kinect v2 camera, ultrasonic, and the imaging array sensors had high correlations ( $r\geq 0.90$ ) with manual measurements. They believed that the errors among sensors and field measurements come from the sensor's error. On the other hand, more and more studies emphasized that there may be systematic errors in the ground truth values. For example, Maesano et al. [44] pointed out that LiDAR can detect more precise height differences than field measurement by comparing the accuracy of grass CH derived from ULS and field measurement. The inaccuracy of field-measured CH may be attributed to the variations of CH [68] and canopy structure [77]. Similarly, the heritability between 3D sensing and field measurement is also worth exploring.

This study aims to compare CH extraction accuracy and heritability from field measurements and four different proximal 3D sensing technologies, including TLS,

BLS, GLS, and DAP in a wheat field of different varieties across different growth stages. Unlike previous studies, we make the following contributions: 1) systematically evaluating the accuracy of different data sources (TLS, BLS, GLS, DAP, and FM/field measurement) in estimating CH, 2) exploring the variations of height measurement accuracy concerning different CH, LAI, and the GS groups, 3) deciphering the error sources of CH measurement among different data sources, and 4) exploring the heritability of 3D sensing data sources in estimating CH.

## Materials and data collection

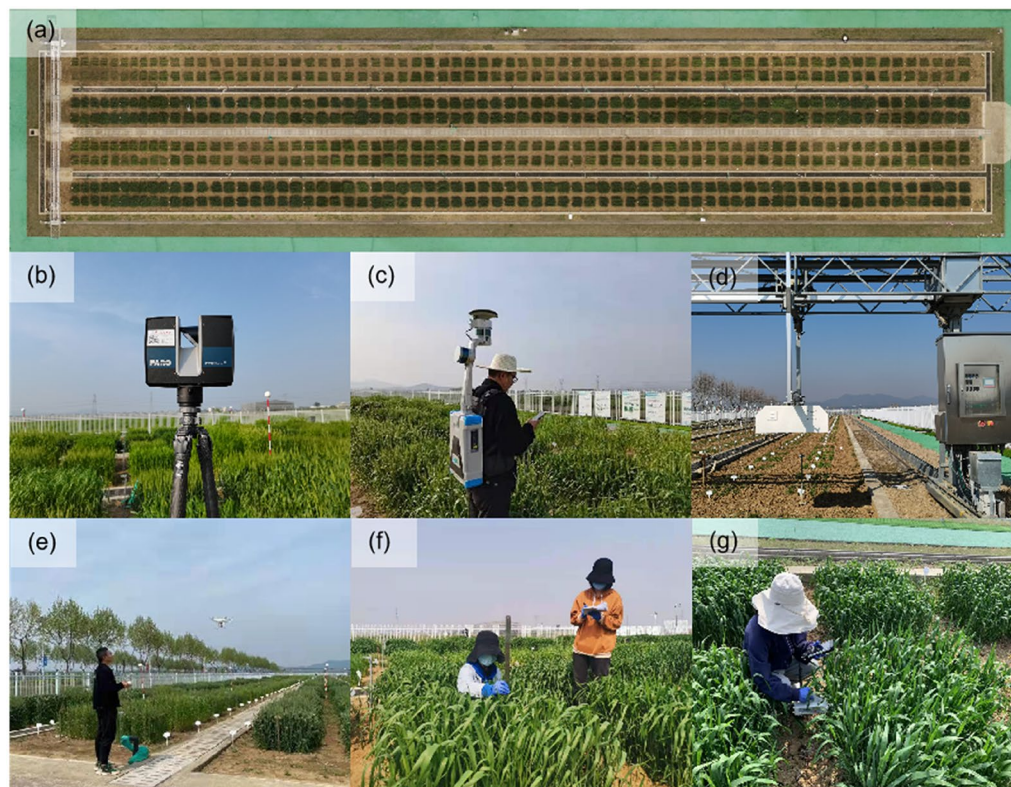
### Study area and experimental design

The study area was located at the Baima Experimental Station (119°18'71"E, 31°62'00"N) of Nanjing Agricultural University, China. A total of 480 plots were cultivated with 120 wheat varieties, two treatments of nitrogen fertilization (0 and 240 kg/ha), and two replications. The plot size is 1 m × 1 m with a plot spacing of 0.5 m, row spacing of 0.25 m, and sowing density of 300 seeds/m<sup>2</sup> (Fig. 1a). Different varieties, nitrogen treatments, and growth stages provided diverse canopy

structure for further comparison of CH from different data sources.

### Data collection

To make a systematic comparison of different height measurement methods, TLS, BLS, GLS, and DAP were selected to collect 3D data at four key growth stages that were jointing (134 days after seeding/DAS), heading (151 DAS), flowering (174 DAS), and maturity stages (188 DAS). These data were collected around noon (10:00–14:00) on sunny days, when the light and wind conditions are stable and preferred for optical image collection (e.g., DAP), although LiDAR sensors are insensitive to light conditions. Some important technical specifications used by the four 3D sensing systems are presented in Table 1. Meanwhile, field-measured CH and LAI were implemented with a ruler and the Sunscan Canopy Analyzer (*Delta-T Devices Ltd, U.K.*). Finally, different data sources were collected within one day at each growth stage to ensure cross-comparability.



**Fig. 1** **a** Study area and data collection by **b** terrestrial laser scanning (TLS), **c** backpack laser scanning (BLS), **d** gantry laser scanning (GLS), and **e** digital aerial photogrammetry (DAP); Manual measurement of **f** canopy height (CH) with a ruler and **g** leaf area index (LAI) with the SunScan Canopy Analyzer

**Table 1** Technical specifications of TLS, BLS, GLS, and DAP systems

	TLS	BLS	GLS	DAP
System	FARO Focus <sup>3D</sup> S70	LiBackpack D50	PlantEye F500	DJI Phantom4
Laser wavelength, nm	1550	905	940	RGB image
Field of view, °	H: 360° V: 300°	H: 360° V: - 90°~ + 90°	~ 53°	94°
Detection range, m	0.60–70 @10% ref	100 @ 20% ref	0.40–1.50	Flight altitude: 10–6000
Data resolution	0.30 mm @ 10 m@ 90% ref	30 mm	H: ~ 0.59 mm V: ~ 1.62 mm	4000 × 3000 pixels
Weight, kg	4.20	8	8.30	0.138
Size, mm	240 × 200 × 100	960 × 300 × 318	440 × 210 × 99	Wheelbase: 350
Battery capacity, h	4.50	2	unlimited	0.50

**TLS data**

The TLS data was collected using the FARO Focus<sup>3D</sup> S70 scanner (FARO Technology Inc, FL, USA). The sensor weight is 4.2 kg with a size of 240 mm × 200 mm × 100 mm. The field of view is 360° × 300°. The sensor emits lasers at a wavelength of 1550 nm and a pulse emitting rate of 244 kHz. The detection range is 0.6–70 m with upright incidence to a 10% reflective surface. The scanning accuracy is 0.3 mm @ 10 m @ 90% reflectance (Table 1).

The LiDAR sensor was mounted on a tripod (around 1.8 m above the ground) that was placed uniformly in the study area (Fig. 1b). The north–south and east–west distances between the two scanning locations were around 4 m and 7.5 m, respectively. The operating mode of the sensor was set as “Outdoor within 10 m Scanning Profile” without color information, which is suitable for acquiring detailed information with high efficiency (~5 min/scan) within a short distance (< 10 m) [26]. A total of 65 scans were implemented over the entire wheat field (Fig. 2b), taking around 6 h.

**BLS data**

The BLS data was acquired using the LiBackpack D50 system (Green Valley International Ltd., Beijing, China) that was equipped with two Velodyne VLP-16E sensors (Velodyne Lidar Inc., San Jose, CA, USA). The system weight is about 8 kg with a size of 960 mm × 300 mm × 318 mm. The field of view is 360° × 180° (- 90° ~ + 90°). The sensor emits lasers at a wavelength of 905 nm and a pulse emitting rate of 30 kHz. The detection range is 100 m with upright incidence to a 20% reflective surface. The scanning accuracy is ± 3 cm (Table 1).

BLS was carried on the shoulder (Fig. 1c), enabling efficient and flexible mobile acquisition. Because BLS uses the SLAM (simultaneous location and mapping) algorithm for data acquisition, the moving trajectory was

designed like a series of closed “triangles” (Fig. 2c). The collection time was around 20 min for the whole field.

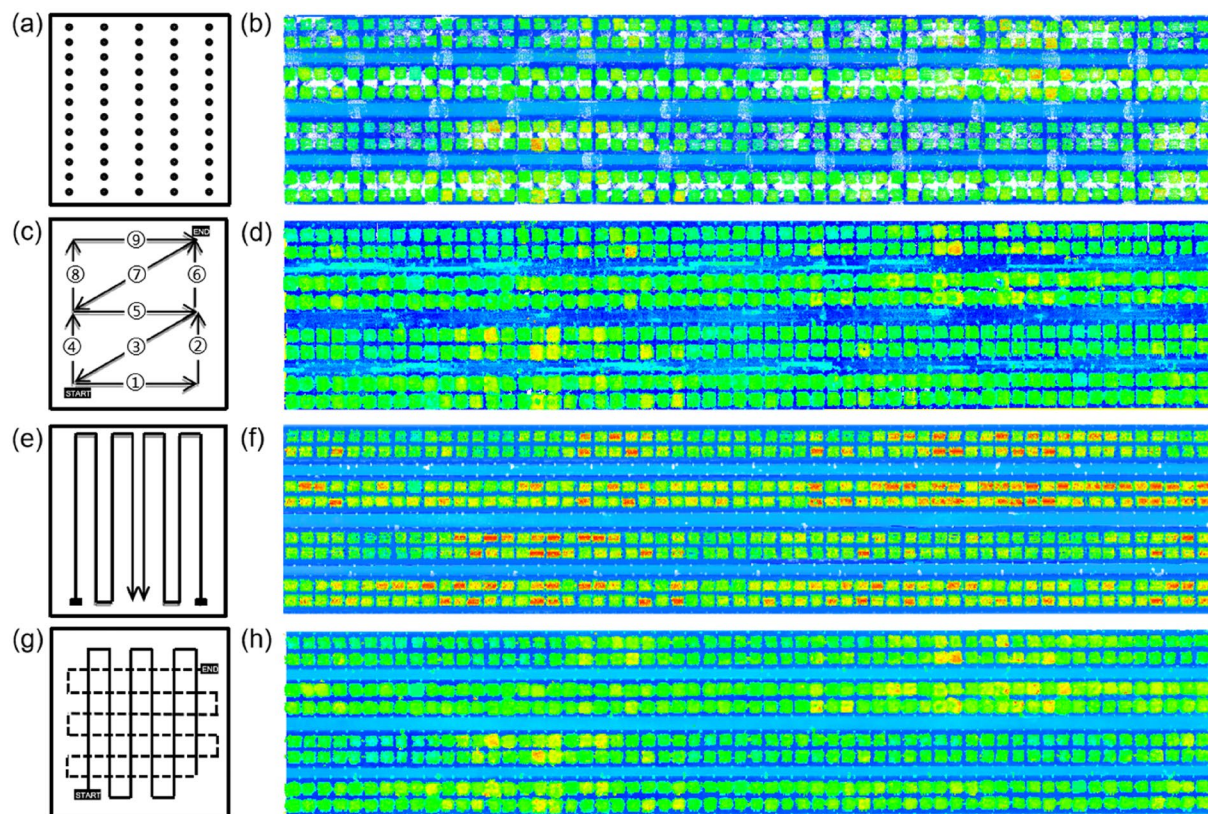
**GLS data**

The GLS data were acquired by using the FieldScan Phenotyping Platform (Fig. 1c), which is equipped with four high-resolution 3D laser scanners, PlantEye F500 (Phenospex Inc, Heerlen, The Netherlands) (Fig. 1d). The sensor weight is around 8.3 kg with a size of 440 mm × 210 mm × 99 mm, and the field of view is around 53°. The sensor emits lasers with a wavelength of 940 nm and a pulse emitting rate of 50 XZ-profiles/s. The ranging distance is between 0.4–1.5 m. The sensors scanning accuracies will decrease with the increase of distance along the vertical height range. The average horizontal and vertical resolutions are around 0.59 mm and 1.62 mm, respectively (Table 1).

The sensor system was carried by a gantry at a height of 1.5 m, and the maximum scan range is 1.1 m. The GLS system traveled automatically in the field with a defined regular trajectory (Fig. 2e). The system repeatedly collected data day and night for the whole field. Each round of collection took around 4.5 h, and then the system slept 1.5 h before the next round of collection. Notably, the integrated software system will remove ground points (i.e., filtering) by setting a height threshold of the lowest 0.28 m in this study, so the maximum detected canopy height of the GLS system is 0.82 m.

**DAP data**

The DAP data was collected using the DJI Phantom4 drone (SZ DJI Technology Co., Shenzhen, China) by carrying an RGB camera (Fig. 1e). The camera has a resolution of 4000 pixels × 3000 pixels. The field of view is 94°. Flight missions were planned using the Pix4D Capture software (PIX4D S.A., Lausanne, Switzerland). To balance the problem of acquisition accuracy and efficiency [31], we carried out comparisons at different flight altitudes,



**Fig. 2** Data acquisition schemes and point clouds collected by different 3D sensing techniques. **a** scanner positions and **b** point cloud of TLS; **c** trajectory and **d** point cloud of BLS; **e** trajectory and **f** point cloud of GLS; and **g** trajectory and **h** point cloud of DAP

including 10 m, 20 m, 30 m, and 40 m. The 20 m was selected because its accuracy is comparable to 10 m and higher than 30 m and 40 m (Additional file 1: Fig. S1). Oblique imageries were collected to ensure substantial overlap and reduced systematic errors [23]. Meanwhile, the cross flight was set up, covering an east–west and a north–south flight trajectories, to improve 3D reconstruction accuracy from images (Fig. 2g). Specifically, the forward and side overlaps were both set as 80%. The camera angle during the flight was set to 80° by referring to Rosnell and Honkavaara [51]. Seven ground control points were set up for image quality control in the field. A total of 216 images were collected during a 20 min flight.

#### Field measurements

In this study, the field CH is defined as the vertical distance from the ground to the highest point of a canopy in the natural growth state. In each plot, CHs were measured with a ruler of mm precision at three locations that look uniform and representative. The three replicated measurements were averaged as the reference CH (Fig. 1f) [68]. LAI was defined as half the total intercepting leaf area pre-unit ground area [6]. LAI was measured

with a SunScan Canopy Analyzer (*Delta-T Devices Ltd, Cambridge, U.K.*) that has a 1-m light-sensitive probe with 64 equally spaced photodiodes. The SunScan Canopy Analyzer estimates LAI by measuring the gap fraction [49]. In each plot, the probe was inserted into the bottom of the canopy and parallel to the row direction [47, 55]. Three replicated measurements were implemented and averaged as the reference LAI (Fig. 1g).

## Methods

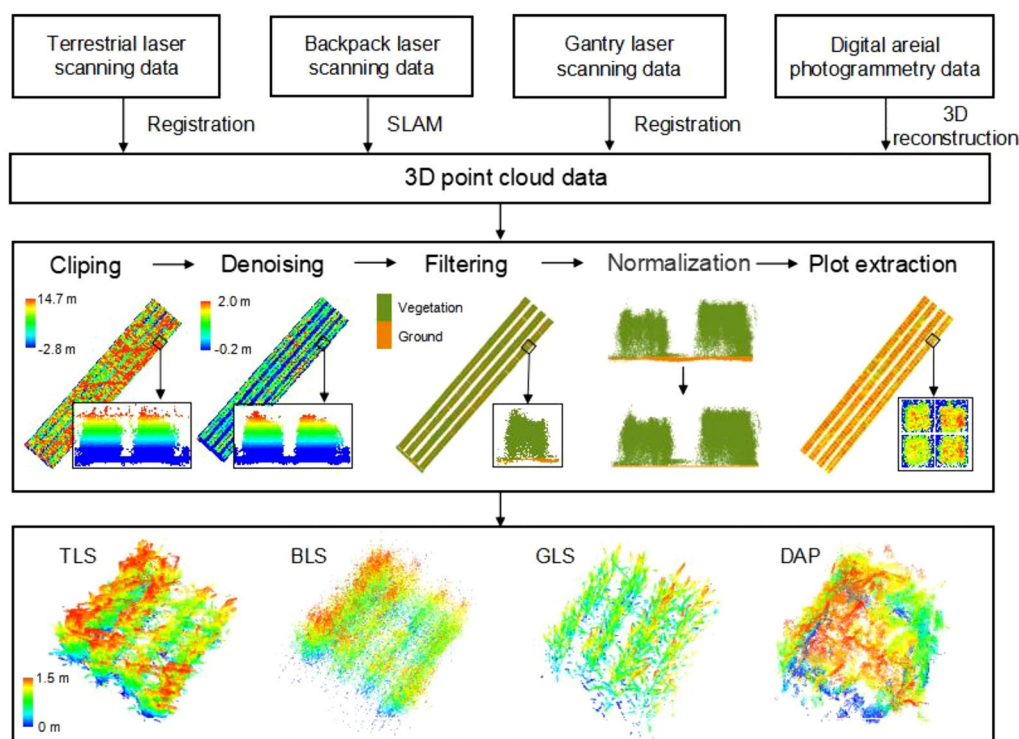
### Data preprocessing

Different 3D sensing data need to be first processed into point clouds with different methods before sharing similar point processing methods. TLS data at different scanning locations were automatically registered to generate a point cloud using SCENE software (*FARO Technology Inc, FL, USA*). BLS was registered during data collection because the system used the SLAM algorithm [58]. GLS data registration was implemented according to the relative position of sensors and the point features using the commercial HortControl software (*Phenospex Inc, Heerlen, The Netherlands*). DAP images were used to reconstruct the 3D point cloud using the PiX4D mapper

software (*Pix4D*, Lausanne, Switzerland). Once the 3D point clouds were generated, the following data processing processes were similar (Fig. 3).

The generated 3D point cloud data were further processed with a standard pipeline using the *LiDAR360* software (*Green Valley International Ltd.*, Beijing, China), including clipping, denoising, filtering, and normalization (Fig. 3). Clipping and denoising were manually implemented to ensure better accuracy, especially avoiding the loss of points in the sparse DAP and BLS point cloud. Filtering was first implemented using an integrated algorithm (i.e., improved progressive triangulated irregular network densification filtering algorithm), and the automatic results were carefully checked and revised

to decrease process errors. Normalization was achieved by subtracting the height of each point from the height of its nearest ground point in the horizontal direction. Specifically, GLS data was filtered with a given height threshold of 0.28 m and normalized during data collection. The normalized 3D point clouds of TLS, BLS, GLS, and DAP were shown in Fig. 2b, d, f, h. Taking pre-processed data at the heading stage as an example, the point density of TLS data is the highest (929,021.12 pts/m<sup>2</sup>), followed by GLS (697,092.18 pts/m<sup>2</sup>), DAP (40,051.30 pts/m<sup>2</sup>), and BLS (17,761.30 pts/m<sup>2</sup>). Meanwhile, the final point resolution, denoted by the average adjacent point distance, from fine to coarse was GLS (1.07 mm), TLS (2.46 mm), DAP (12.73 mm), and BLS (15.02 mm) (Table 2).



**Fig. 3** Processing of TLS, BLS, GLS, and DAP data. The processing pipeline was demonstrated using GLS data at the heading stage. SLAM means simultaneous location and mapping

**Table 2** Key information about the data quality of the preprocessed point clouds (taking data at the heading stage as an example) and the roughly estimated platform cost and data cost

Data sources	Point density, pts/m <sup>2</sup>	Point resolution, mm	Data volume, GB	Platform cost, \$	Data Cost, h		
					Collection	Preprocessing	Total
TLS	929,021.12	2.46	11.40	46,010.00	5.00	79.20	84.20
BLS	17,761.30	15.02	0.22	70,515.00	0.30	2.30	2.60
GLS	697,092.18	1.07	8.68	1,567,000.00	8.00	2.00	6.50
DAP	40,051.30	12.73	0.48	1253.60	0.50	15.70	16.20

Plot extraction is the prerequisite for CH extraction of each plot. Because different sources of point clouds have their sensor coordinate systems, this study manually aligned these data into the same coordinate origin and north–south directions in LiDAR360 software. After that, 480 plots of different source data at each growth stage can be extracted using a shared plot bounding box map defined manually (Fig. 3).

### Canopy height extraction

CH can be extracted from the normalized point cloud using different statistical metrics. In this study, Hmax, the maximum  $z$  value of all normalized points, was extracted. Meanwhile, difference height quantiles from 99% quantile height (i.e., H99) to 80% quantile height (i.e., H80) with an interval of 1% were also extracted [27]. These different height representations are compared and the optimal one was selected for comparing different sensing technologies.

### Cross-comparisons of canopy height estimates from field measurement and 3D sensing

The accuracies of the CH measured by different 3D sensing data were compared with the field measurement, and the cross-comparisons of different 3D sensing performances were also evaluated. Specifically, the comparisons between sensor data with field measurement include TLS vs.FM, BLS vs.FM, GLS vs.FM, DAP vs.FM, and the cross-comparisons include TLS vs. BLS, BLS vs. DAP, DAP vs. TLS, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS.

This study further evaluated the accuracy of different methods with respect to different field-measured CH groups, LAI groups, and GS groups, which are important indicators of canopy structure [41, 42] and affect the accuracy of CH monitoring. Four CH groups were considered, including 0.3–0.6 m (CH1), 0.6–0.8 m (CH2), 0.8–1 m (CH3), and 1–1.4 m (CH4). Each height group contains 360, 918, 501, and 141 plots, respectively. Four LAI groups were separated at 0–2 m<sup>2</sup>/m<sup>2</sup> (LAI1), 2–4 m<sup>2</sup>/m<sup>2</sup> (LAI2), 4–6 m<sup>2</sup>/m<sup>2</sup> (LAI3), and 6–8 m<sup>2</sup>/m<sup>2</sup> (LAI4). Each group contains 874, 641, 340, and 65 plots, respectively. Four compared growth stages were jointing stages, heading stages, flowering stages, and maturity stages.

Specifically, considering the scanning range and height threshold setting in filtering, the effective maximum height of the GLS system is 0.82 m. Therefore, only the plots that have a maximum measured height lower than 0.82 m were selected for comparison with GLS (1365 plots) in this study. Because there are a few plots belonging to the CH3 group and no plots belonging to the CH4 group, we only evaluated the GLS accuracies of CH1 and CH2 (360 and 918 plots, respectively).

The accuracy between the two compared groups was evaluated by Pearson's correlation coefficient ( $r$ ), root mean square error ( $RMSE$ ), relative  $RMSE$  ( $RMSE\%$ ), Bias, and relative Bias ( $Bias\%$ ).

$$r = \sqrt{1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}} \quad (1)$$

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

$$RMSE\% = \left( \frac{RMSE}{\bar{y}_i} \right) \times 100 \quad (3)$$

$$Bias = \sum_{i=1}^n (y_i - \hat{y}_i) / n \quad (4)$$

$$Bias\% = \left( \frac{Bias}{\bar{y}_i} \right) \times 100 \quad (5)$$

where  $i$  represents a sample index,  $n$  represents the number of samples,  $y_i$  represents reference measurements (e.g., FM),  $\hat{y}_i$  represents predicted CH from different 3D sensing datasets, and  $\bar{y}_i$  is the mean of  $y_i$ .

Moreover, the CHs of different data sources were compared in terms of broad-sense heritability ( $H^2$ ). Broad-sense heritability was defined as the proportion of heritability variance, which was computed as the ratio between the genotypic to the total variance [65, 66]. In this study, the interaction effect of different varieties and  $N$  treatments was considered, i.e., G by E.

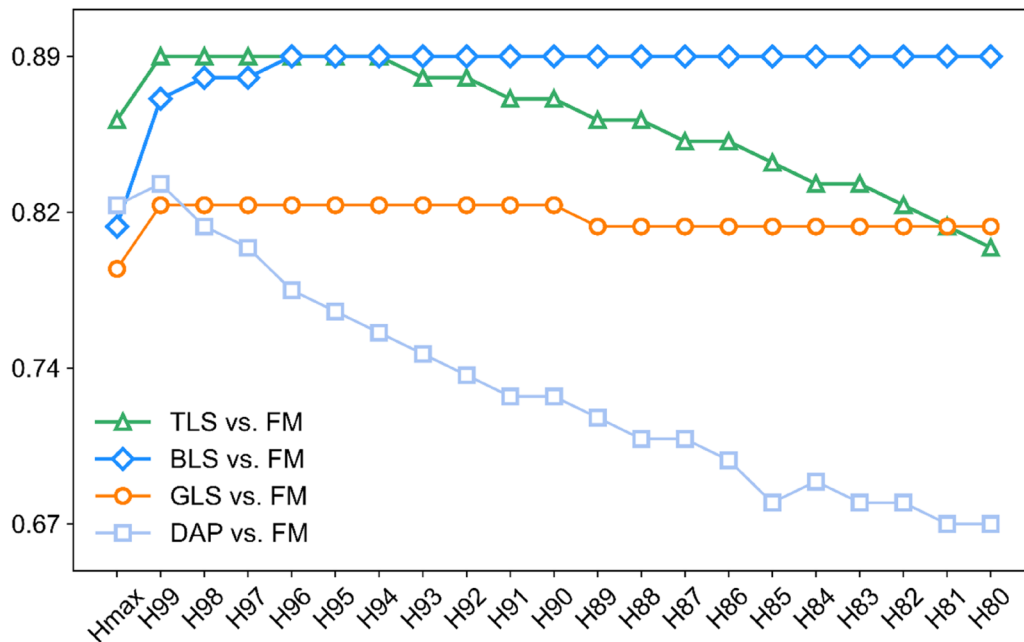
$$\sigma_G^2 = \frac{1}{g-1} \sum_i G_i^2 \quad (6)$$

$$\sigma_E^2 = \frac{1}{e-1} \sum_j E_j^2 \quad (7)$$

$$\sigma_{GE}^2 = \frac{1}{(g-1)(e-1)} \sum_{ij} GE_{ij}^2 \quad (8)$$

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{GE}^2}{e} + \frac{\sigma_E^2}{re}} \quad (9)$$

where  $H^2$  is broad-sense heritability,  $\sigma_G^2$ ,  $\sigma_E^2$ , and  $\sigma_{GE}^2$  are genotypic variance, environmental variances, and genotype-by-environment interaction variance, respectively.  $g$  is the number of genotypes,  $i$  is the index of genotype;  $e$



**Fig. 4** Correlation values between FM and height quantiles (e.g., Hmax and H99) derived from the different 3D sensing data, including TLS, BLS, GLS, and DAP. The green triangle, blue diamond, dark orange circle, and baby blue square represent correlations of TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively

is the number of N treatments, *j* is the index of N treatments, and *r* is the number of replications per genotype.

### Error source analysis

As we know, CHs measured by different methods will not be exactly the same. This study analyzed which data source the error comes from by referring to the method of Wang et al. [71]. First, we calculate the relative residual between the 3D sensing estimated CHs and FM (Eq. 10). Then, screening out the plots where the above calculated relative residuals greater than 20% as the suspicious cases (*S*) (Eq. 11). The intersections of *S*<sub>TLS</sub>, *S*<sub>BLS</sub>, *S*<sub>GLS</sub>, and *S*<sub>DAP</sub> were defined as the errors due to FM (Error\_FM) (Eq. 12). Based on Error\_FM, the intersection of *S*<sub>TLS</sub>, *S*<sub>BLS</sub>, *S*<sub>GLS</sub>, *S*<sub>DAP</sub>, and non-Error\_FM was defined as the errors due to TLS (Error\_TLS), BLS (Error\_BLS), GLS (Error\_GLS), and DAP (Error\_DAP), respectively (Eq. 13–16). Notably, when regarding TLS or any other 3D sensing datasets as the errors, it is not mean the other three 3D sensing datasets do not contain outliers because the conditions for Error\_FM are very strict.

$$\Delta^i_{(a,field)} = \left| \frac{H^i_a - H^i_{field}}{H^i_{field}} \right| \tag{10}$$

$$S_a = \left\{ P^i \mid \Delta^i_{(a,field)} \geq 0.2 \right\} \tag{11}$$

$$Error\_FM = \left\{ P^i \mid S_{TLS} \cap S_{BLS} \cap S_{GLS} \cap S_{DAP} \right\} \tag{12}$$

$$Error\_TLS = \left\{ P^i \mid S_{TLS} \cap (!Error\_field) \right\} \tag{13}$$

$$Error\_BLS = \left\{ P^i \mid S_{BLS} \cap (!Error\_field) \right\} \tag{14}$$

$$Error\_GLS = \left\{ P^i \mid S_{GLS} \cap (!Error\_field) \right\} \tag{15}$$

$$Error\_DAP = \left\{ P^i \mid S_{DAP} \cap (!Error\_field) \right\} \tag{16}$$

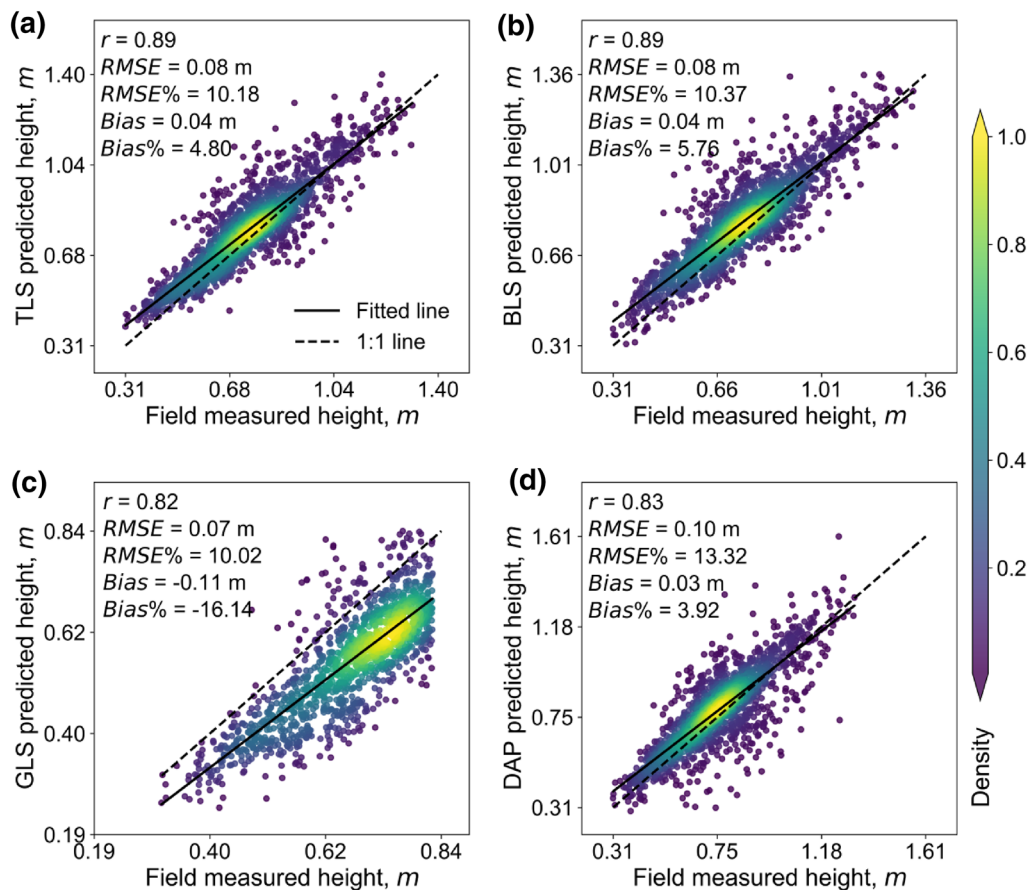
where *i* is the sample index and *P*<sup>*i*</sup> represents sampled data (i.e., a plot).  $\Delta^i_{(a,field)}$  is the relative residual between 3D sensing-derived CH and FM, *H*<sup>*i*</sup><sub>*a*</sub> represent predicted CH, where *a* can be TLS, BLS, GLS, and DAP. Meanwhile, the exclamation mark (!) is the “NOT” in logic operations

## Results

### Canopy height from different 3D sensing datasets

To fairly compare different 3D sensing datasets for CH estimation, it is important to first explore which height representation metric is optimal according to their correlations with FM. In this study, the influences of different





**Fig. 5** Correlations between FM and predicted heights by different 3D sensing technologies. **a–c**, and **d** represent TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line. The color bar shows the kernel density value of the point distribution, and the green to yellow represents the increase in kernel density

height quantiles for CH extraction were evaluated using point clouds of all stages. The results showed that the evaluation accuracy of TLS, BLS, and GLS were all high and stable when using different height quantiles (Fig. 4). By contrast, height estimation accuracy from DAP data was lower and more sensitive to the selection of height quantiles. According to the highest correlation (Fig. 4) and the lowest error metrics (Additional file 1: Fig. S2), H99 was selected as the best representation of CH for TLS, GLS, and DAP, while H96 was the best for BLS. These best height quantiles (H99 or H96) for each data source was used for all subsequent analysis.

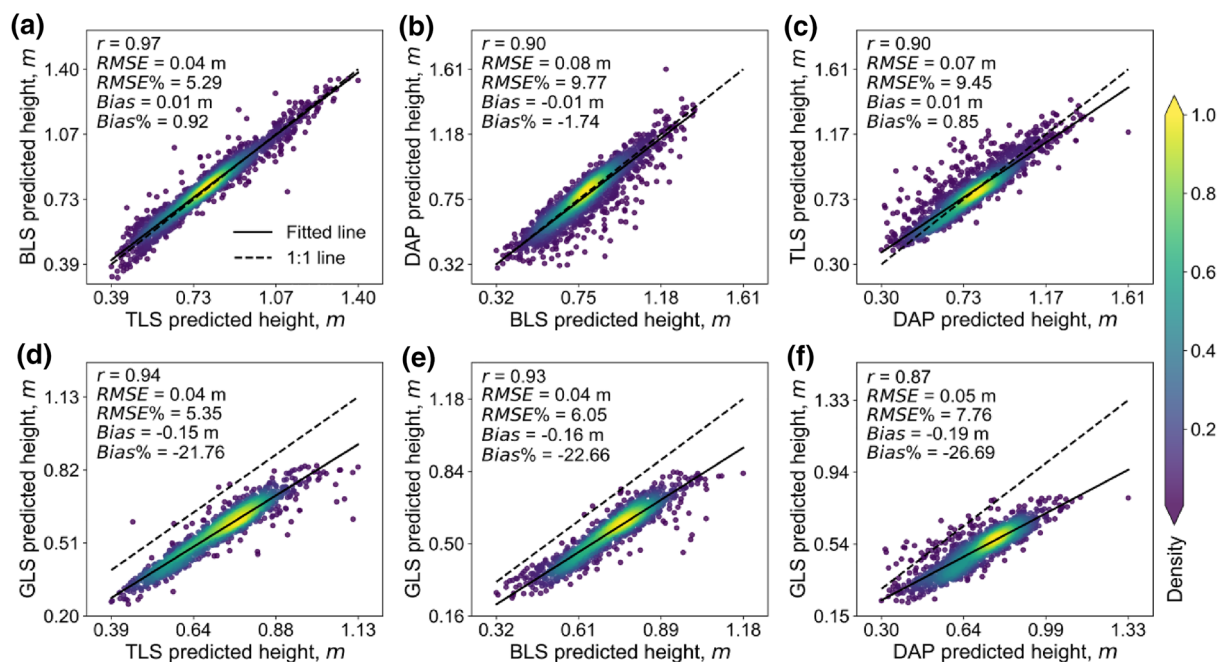
The best correlations of TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM were 0.89, 0.89, 0.82, and 0.83, respectively (Fig. 5). The fitted lines of TLS, BLS, and DAP were close to the reference lines (1:1) except a little overestimation when CH was small (Fig. 5). In contrast, GLS showed an overall underestimation (Fig. 5c).

Cross-comparisons among different sensor datasets showed higher correlations ( $r$ ) ranging from 0.87 to 0.97,

which was much higher than the above comparisons with FM (0.82–0.89). The highest correlation value is 0.97 between TLS and BLS (Fig. 6a), followed by TLS vs. GLS ( $r=0.94$ ) (Fig. 6d), BLS vs. GLS ( $r=0.93$ ) (Fig. 6e), DAP vs. TLS ( $r=0.90$ ) (Fig. 6c), BLS vs. DAP ( $r=0.90$ ) (Fig. 6b), and DAP vs. GLS ( $r=0.87$ ) (Fig. 6f). Among them, DAP had a relative larger RMSE with other sensing datasets ( $RMSE > 0.05$  m, Fig. 6b, c, f), especially the comparison with BLS ( $RMSE=0.08$  m, Fig. 6b). Moreover, the fitting *Bias* are all very small (0.01 m) except for comparisons with GLS (Fig. 6d, e, f). Although GLS showed an overall underestimation, it still keeps a low RMSE (0.04 m–0.05 m) with other 3D sensing datasets.

#### Comparing canopy height measurement of different methods among different canopy height groups

The correlation coefficients of CHs derived from 3D sensing and FM decreased obviously when evaluated with respect to different subgroups of CH ( $r < 0.71$ ). Similarly,



**Fig. 6** Canopy height correlations between different 3D sensing estimates. **a–e**, and **f** represent TLS vs. BLS, BLS vs. DAP, DAP vs. TLS, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line. The color bar shows the kernel density value of the point distribution, and the green to yellow represents the increase in kernel density

the correlation coefficients of cross-comparisons of different 3D sensing also decreased, although the largest  $r$  was up to 0.93 (Table 3).

As for comparing 3D sensing with FM, GLS was the best according to the highest mean  $r$  (0.60), followed by BLS (mean  $r=0.59$ ), TLS (mean  $r=0.58$ ), and DAP (mean  $r=0.52$ ) (Table 3). From the prospect of subgroup comparisons, the best methods for estimating CH1, CH2, CH3, and CH4 were DAP (mean  $r=0.71$ ), TLS (mean  $r=0.56$ ), BLS (mean  $r=0.52$ ), and BLS (mean  $r=0.64$ ), respectively (Table 3). The fitting lines of TLS, BLS, and DAP were very close to the reference lines in CH3 and CH4 groups, while slight overestimation appeared in CH1 and CH2 groups (Fig. 7). Consistently, GLS showed underestimation in both CH1 and CH2 groups (Fig. 7c).

The cross-comparisons of different 3D methods showed much higher correlation values. Among them, TLS vs. GLS showed the highest correlation (mean  $r=0.90$ ), followed by TLS vs. BLS (mean  $r=0.89$ ), BLS vs. GLS (mean  $r=0.85$ ), DAP vs. GLS (mean  $r=0.76$ ), DAP vs. TLS (mean  $r=0.74$ ), and BLS vs. DAP (mean  $r=0.74$ ). From the perspective of subgroup comparisons, the most consistent method for estimating CH1 was TLS vs. GLS, and the most consistent methods for estimating CH2, CH3, and CH4 were always TLS vs. BLS (Table 3).

The fitted lines for TLS vs. BLS and BLS vs. DAP were both close to 1:1 for different CH groups (Additional file 1: Fig. S3 a, b). DAP vs. TLS showed overestimation at low heights and underestimation at high heights for each height group (Additional file 1: Fig. S3 c). Underestimations also almost existed in comparisons between GLS and other 3D sensing datasets at every CH group (Additional file 1: Fig. S3 d–f). The fitted line of GLS vs. TLS was nearly parallel to the reference line, while underestimations to other 3D data become more obvious with height growth.

#### Comparing canopy height measurement of different methods among different LAI groups

The correlation coefficients of CHs derived from 3D sensing and FM only decreased slightly (mean  $r=0.79$  to 0.87) with respect to different LAI groups. Likewise, the correlation coefficients of cross-comparisons of different 3D sensing also decreased slightly (mean  $r=0.84$  to 0.96), with little change for TLS vs. BLS (Table 4).

As for comparing 3D sensing with FM, TLS was the best according to the highest mean  $r$  (0.87), followed by BLS (mean  $r=0.85$ ), DAP (mean  $r=0.81$ ), and GLS (mean  $r=0.79$ ). From the presence of subgroup comparisons, the best method for estimating the height of the LAI1 group were BLS (mean  $r=0.93$ ) and TLS (mean  $r=0.93$ ), while the best method for LAI2, LAI3, and

**Table 3** Detailed statistics on comparing canopy height measurement methods

CH group	RMSE, m (RMSE%)	Bias, m (Bias%)	r	RMSE, m (RMSE%)	Bias, m (Bias%)	r
	TLS vs. FM			BLS vs. FM		
CH1	0.06 (10.92)	0.05 (10.24)	0.66	0.07 (13.65)	0.06 (12.60)	0.65
CH2	0.07 (10.48)	0.05 (6.79)	<u>0.56</u>	0.08 (10.96)	0.06 (7.85)	0.54
CH3	0.08 (9.57)	0.01 (0.95)	0.49	0.08 (8.75)	0.01 (1.48)	<u>0.52</u>
CH4	0.09 (7.98)	0.01 (0.86)	0.59	0.08 (7.58)	0.01 (0.99)	<u>0.64</u>
Mean	0.08 (9.74)	0.03 (4.71)	0.58	0.08 (10.23)	0.04 (5.73)	0.59
	GLS vs. FM			DAP vs. FM		
CH1	0.06 (11.39)	- 0.09 (- 17.96)	0.64	0.07 (12.91)	0.04 (8.15)	<u>0.71</u>
CH2	0.07 (9.82)	- 0.11 (- 15.49)	0.56	0.09 (13.10)	0.05 (6.76)	0.51
CH3	-	-	-	0.11 (12.79)	0.00 (- 0.17)	0.36
CH4	-	-	-	0.14 (12.25)	0.02 (- 1.49)	0.50
Mean	0.06 (10.6)	- 0.10 (- 16.72)	0.60	0.10 (12.76)	0.02 (3.31)	0.52
	TLS vs. BLS			TLS vs. GLS		
CH1	0.05 (8.89)	0.01 (2.15)	0.84	0.03 (5.32)	0.14 (- 25.57)	<u>0.92</u>
CH2	0.04 (4.98)	0.01 (0.99)	<u>0.91</u>	0.04 (5.21)	0.16 (- 20.86)	0.88
CH3	0.04 (4.30)	0.00 (0.52)	<u>0.91</u>	-	-	-
CH4	0.04 (3.71)	0.00 (0.13)	<u>0.93</u>	-	-	-
Mean	0.04 (5.47)	0.01 (0.95)	0.89	0.03 (5.27)	- 0.15 (- 23.22)	0.90
	BLS vs. DAP			BLS vs. GLS		
CH1	0.06 (10.80)	- 0.02 (- 3.95)	0.75	0.04 (7.47)	0.16 (- 27.14)	0.82
CH2	0.07 (8.98)	- 0.01 (- 1.00)	0.77	0.04 (5.31)	0.17 (- 21.64)	0.88
CH3	0.09 (9.80)	- 0.01 (- 1.62)	0.69	-	-	-
CH4	0.11 (9.46)	- 0.03 (- 2.45)	0.74	-	-	-
Mean	0.08 (9.76)	- 0.02 (- 2.26)	0.74	0.04 (6.39)	- 0.16 (- 24.39)	0.85
	DAP vs. TLS			DAP vs. GLS		
CH1	0.04 (7.58)	0.01 (1.93)	0.83	0.06 (13.67)	0.13 (31.82)	0.79
CH2	0.06 (7.83)	0.00 (0.03)	0.75	0.07 (12.38)	0.16 (26.33)	0.73
CH3	0.07 (8.36)	0.01 (0.01)	0.65	-	-	-
CH4	0.07 (6.66)	0.03 (0.03)	0.75	-	-	-
Mean	0.06 (7.61)	0.01 (1.37)	0.74	0.05 (7.95)	- 0.15 (- 22.49)	0.76

The top side of the table showed the evaluation results of 3D sensing datasets with FM; the bottom side of the table showed the results of 3D sensing datasets cross-comparisons. RMSE and RMSE%, Bias and Bias%, and correlation coefficient (r) were given for distinct canopy height (CH) groups. The underlined values were the best result for each CH group among different comparisons

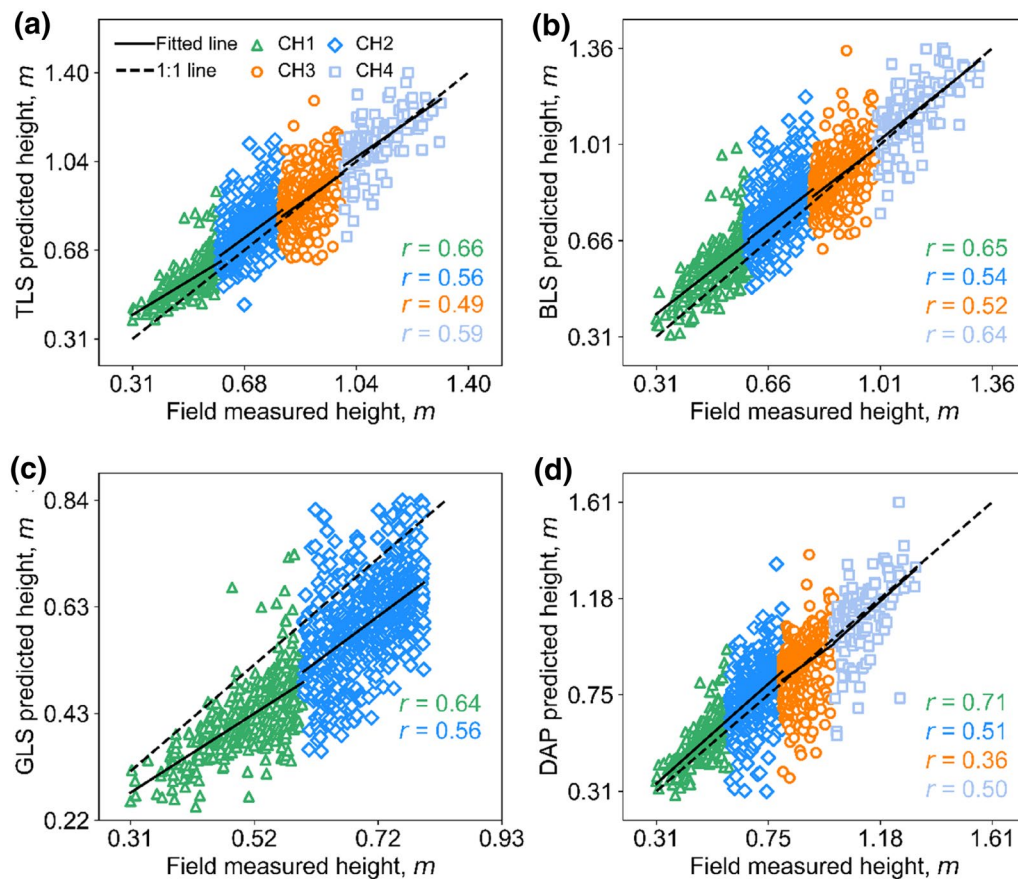
- represents the comparison was not available due to the limited ranging ability of the GLS system

LAI4 was always TLS (mean  $r > 0.83$ ) (Table 4). The fitting lines of TLS, BLS, and DAP were very close to the reference lines in all LAI groups, while GLS showed underestimation in all LAI groups (Fig. 8).

As for cross-comparison of different 3D methods, TLS vs. BLS showed the highest correlation (mean  $r = 0.96$ ), followed by TLS vs. GLS (mean  $r = 0.91$ ), BLS vs. GLS (mean  $r = 0.87$ ), DAP vs. TLS (mean  $r = 0.87$ ), BLS vs. DAP (mean  $r = 0.86$ ), and DAP vs. GLS (mean  $r = 0.84$ ). From the perspective of subgroup comparisons, the most consistent methods for estimating LAI1 were TLS vs. BLS and TLS vs. GLS (mean  $r = 0.97$ ). Besides, the most

consistent methods for estimating LAI3 and LAI4 were still TLS vs. BLS (mean  $r = 0.94$  and  $0.95$ ) (Table 4).

The fitted line for TLS vs. BLS almost coincided with the reference line (Additional file 1: Fig. S4 a). The fitted lines for BLS vs. DAP and DAP vs. TLS were also relatively close to the reference line, but they became worse when LAI increased (Additional file 1: Fig. S4b, c). Underestimations also existed in comparisons between GLS and other 3D sensing datasets at all LAI groups, and the correlations decreased when LAI increased (Additional file 1: Fig. S4d-f).



**Fig. 7** Correlations between FM heights and predicted heights by different 3D sensing technologies under four canopy height (CH) groups. **a–d** represent TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively. The green triangle, blue diamond, dark orange circle, and baby blue square represent the CH1, CH2, CH3, and CH4 groups, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line

### Comparing canopy height measurement of different methods among different GS groups

The correlation coefficients of CHs derived from 3D sensing and FM were less accurate (mean  $r=0.65$  to  $0.83$ ) with regard to different GS groups, especially for GLS vs. FM. By contrast, the correlation coefficients of cross-comparisons of different 3D sensing data decreased slightly (mean  $r=0.80$  to  $0.94$ ) (Table 5).

As for comparing 3D sensing with FM, TLS was the best according to the highest mean  $r$  ( $0.83$ ), followed by BLS (mean  $r=0.81$ ), DAP (mean  $r=0.76$ ), and GLS (mean  $r=0.65$ ) (Table 5). From the perspective of subgroup comparisons, DAP was the best method for estimating CH at the jointing stage (mean  $r=0.89$ ). Moreover, TLS was also the best method for the heading, flowering, and maturity stages (Table 5). The fitting lines of TLS, BLS, and DAP were very close to the reference lines, especially at the heading stage ( $r=0.72$ – $0.92$ ).

However, GLS showed underestimation at all growth stages, which was more obvious at late stages (Fig. 9c).

As for cross-comparison of different 3D methods, TLS vs. BLS showed the highest correlation (mean  $r=0.94$ ), followed by TLS vs. GLS (mean  $r=0.88$ ), BLS vs. GLS (mean  $r=0.86$ ), DAP vs. TLS (mean  $r=0.86$ ), BLS vs. DAP (mean  $r=0.84$ ), and DAP vs. GLS (mean  $r=0.80$ ). From the perspective of subgroup comparisons, the most consistent method for estimating the jointing stage was TLS vs. GLS (mean  $r=0.93$ ), while the best methods for heading, flowering, and maturity stages were TLS vs. BLS (mean  $r=0.92$ – $0.98$ ) (Table 5).

The fitted lines of TLS vs. BLS and BLS vs. DAP were closer to the reference line than DAP vs. TLS (Additional file 1: Fig. S5 a, c, e). Underestimations also existed in comparisons between GLS and other 3D sensing datasets at every GS group, especially at the maturity stage (Additional file 1: Fig. S5 d-f).

**Table 4** Detailed statistics on comparing canopy height measurement methods

LAI group	RMSE, m (RMSE%)	Bias, m (Bias%)	<i>r</i>	RMSE, m (RMSE%)	Bias, m (Bias%)	<i>r</i>
	TLS vs. FM			BLS vs. FM		
LAI1	0.07 (9.11)	0.03 (4.71)	<u>0.93</u>	0.06 (8.78)	0.04 (5.13)	<u>0.93</u>
LAI2	0.09 (11.43)	0.04 (5.32)	<u>0.86</u>	0.09 (11.74)	0.05 (7.11)	0.85
LAI3	0.07 (9.79)	0.03 (3.88)	<u>0.83</u>	0.08 (10.64)	0.03 (4.68)	0.79
LAI4	0.07 (8.98)	0.04 (5.46)	<u>0.86</u>	0.07 (9.10)	0.05 (6.16)	0.84
Mean	0.07 (9.83)	0.04 (4.84)	0.87	0.07 (10.06)	0.04 (5.77)	0.85
	GLS vs. FM			DAP vs. FM		
LAI1	0.05 (7.66)	− 0.09 (− 14.98)	0.91	0.08 (11.51)	0.02 (2.43)	0.89
LAI2	0.08 (12.05)	− 0.10 (− 14.99)	0.73	0.13 (16.53)	0.04 (5.61)	0.77
LAI3	0.06 (8.86)	− 0.14 (− 19.58)	0.78	0.07 (9.89)	0.03 (4.09)	0.77
LAI4	0.06 (7.85)	− 0.13 (− 18.74)	0.73	0.06 (8.55)	0.04 (5.78)	0.82
Mean	0.06 (9.10)	− 0.12 (− 17.07)	0.79	0.09 (11.62)	0.03 (4.48)	0.81
	TLS vs. BLS			TLS vs. GLS		
LAI1	0.04 (5.05)	0.00 (0.40)	<u>0.97</u>	0.03 (4.46)	− 0.15 (− 21.93)	<u>0.97</u>
LAI2	0.04 (5.25)	0.01 (1.70)	<u>0.97</u>	0.04 (5.40)	− 0.15 (− 20.89)	0.94
LAI3	0.04 (5.68)	0.01 (0.77)	<u>0.94</u>	0.04 (6.02)	− 0.16 (− 22.76)	0.90
LAI4	0.04 (4.74)	0.01 (0.67)	<u>0.95</u>	0.05 (6.33)	− 0.17 (− 22.32)	0.82
Mean	0.04 (5.18)	0.01 (0.88)	0.96	0.04 (5.55)	− 0.16 (− 21.98)	0.91
	BLS vs. DAP			BLS vs. GLS		
LAI1	0.07 (8.64)	0.02 (− 2.56)	0.93	0.04 (5.48)	− 0.15 (− 22.54)	0.95
LAI2	0.09 (11.08)	0.01 (− 1.40)	0.89	0.04 (5.51)	− 0.17 (− 22.21)	0.94
LAI3	0.06 (8.52)	0.00 (− 0.56)	0.82	0.05 (7.00)	− 0.17 (− 23.52)	0.86
CH4	0.07 (8.33)	0.00 (− 0.36)	0.80	0.06 (7.38)	− 0.18 (− 23.24)	0.74
Mean	0.07 (9.14)	− 0.01 (− 1.22)	0.86	0.05 (6.34)	− 0.17 (− 22.88)	0.87
	DAP vs. TLS			DAP vs. GLS		
LAI1	0.06 (8.21)	0.02 (− 2.18)	0.94	0.04 (6.48)	− 0.13 (− 20.08)	0.93
LAI2	0.10 (12.06)	0.00 (0.28)	0.87	0.07 (9.79)	− 0.15 (− 21.10)	0.80
LAI3	0.06 (8.14)	0.00 (0.20)	0.84	0.06 (7.96)	− 0.17 (− 23.75)	0.81
LAI4	0.06 (8.12)	0.00 (0.31)	0.82	0.05 (6.33)	− 0.18 (− 23.65)	0.81
Mean	0.07 (9.32)	0.00 (0.36)	0.87	0.06 (7.64)	− 0.16 (− 22.14)	0.84

The top side of the table showed the evaluation results of 3D sensing datasets with FM; the bottom side of the table showed the results of 3D sensing datasets cross-comparisons. RMSE and RMSE%, Bias and Bias%, and correlation coefficient (*r*) were given for distinct leaf area index (LAI) groups. The underlined values were the best result for each LAI group among different comparisons

### Comparing the broad sense heritability of canopy height measurement from different methods

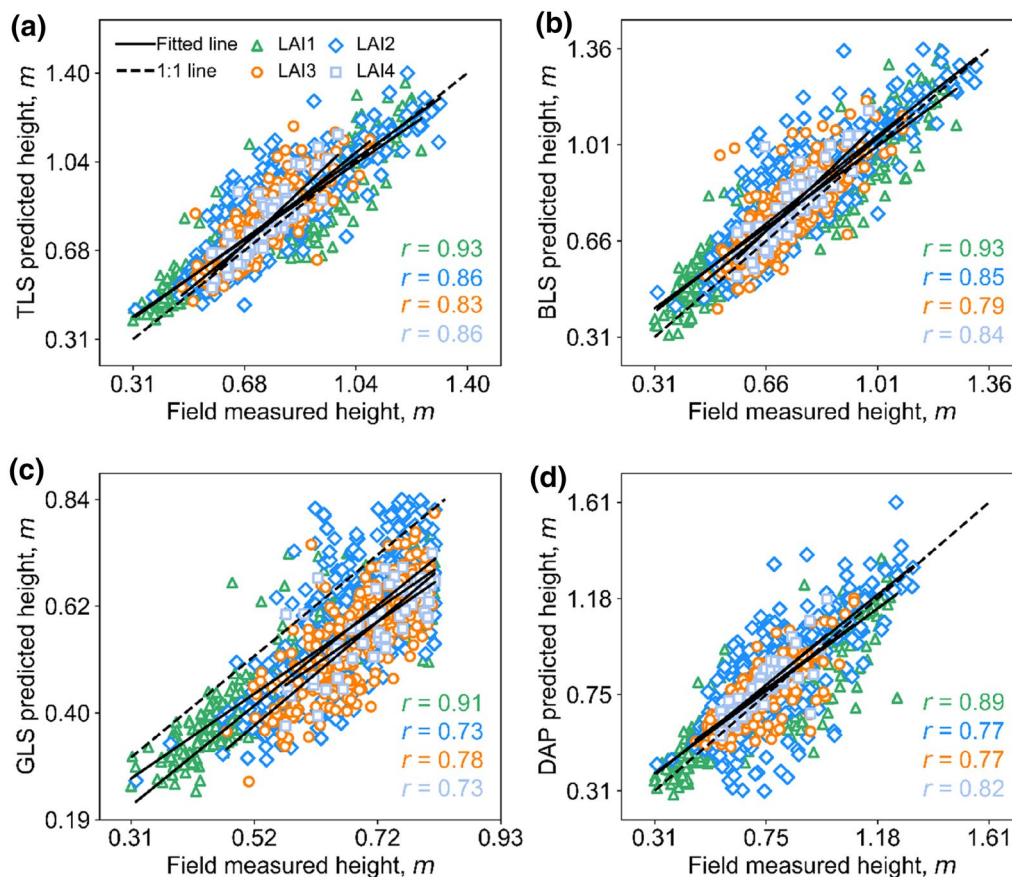
This study found the  $H^2$  of CH derived from 3D sensing datasets was overall higher than FM no matter analyzed with CH, LAI, or GS groups (Table 6). At different CH groups, TLS showed the highest  $H^2$  (mean  $H^2 = 0.73$ ), followed by BLS (mean  $H^2 = 0.70$ ), GLS (mean  $H^2 = 0.66$ ), DAP (mean  $H^2 = 0.66$ ), and FM (mean  $H^2 = 0.60$ ). The  $H^2$  of the lower CH group (CH1) derived from 3D sensing was much larger than the higher CH group (CH2). At different LAI groups, TLS also showed the highest  $H^2$  (mean  $H^2 = 0.90$ ), followed by GLS (mean  $H^2 = 0.86$ ), BLS (mean  $H^2 = 0.85$ ), DAP (mean  $H^2 = 0.84$ ), and FM (mean  $H^2 = 0.83$ ). At different GS groups, TLS also showed the

highest  $H^2$  (mean  $H^2 = 0.89$ ), followed by BLS (mean  $H^2 = 0.85$ ), GLS (mean  $H^2 = 0.81$ ), DAP (mean  $H^2 = 0.79$ ), and FM (mean  $H^2 = 0.77$ ). Overall,  $H^2$  of LiDAR-derived CH was larger than that derived from DAP, and  $H^2$  of all 3D sensing-derived CH was larger than FM. The overall heritability in the later growth period decreased, especially in the maturity stage.

### Discussions

#### Height quantities of 3D point cloud affect the best estimates of canopy height

Height quantities have been widely used for depicting CH due to their insensitivity to noisy points [19]. However, it has been found that different height quantiles may



**Fig. 8** Correlations between FM heights and predicted heights by different 3D sensing technologies under four leaf area index (LAI) groups. **a–d** represent TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively. The green triangle, blue diamond, dark orange circle, and baby blue square represent the LAI1, LAI2, LAI3, and LAI4 groups, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line

be suitable for different 3D data with regard to different crop types [38] and sensor types [43].

In this study, we explored the effects of height quantities on the accuracy of height estimation from four kinds of 3D sensing techniques by collecting 1920 wheat plots of various varieties and nitrogen treatments at four growth stages. Our results found that H99 was the best CH quantile of TLS, GLS, and DAP, while H96 was the best for BLS data (Fig. 4). These results are reasonable considering previous studies found the best height quantiles mainly located between H90 and H99, especially near H99, such as the best height quantile for maize was H99 ( $R^2=0.9$ ) [46] and H99.9 [39], for wheat was H99.5 ( $R^2=0.90$ ) [43], and for soybean was H99.9 ( $R^2>0.85$ ) [38].

Although the best height quantiles are similar, the influences of height quantile selection on height estimation are different. DAP was easy to lose small targets such as the leaf tips of the canopy [46]. Meanwhile,

DAP was difficult to capture the internal structure of the canopy [13], which leads to sparse point density (Fig. 10) and may illustrate why DAP-predicted CH accuracy was more sensitive to height quantiles (Fig. 4) and had a relative larger RMSE with other sensing datasets ( $RMSE>0.05$  m, Fig. 6b, c, f). By contrast, TLS, BLS, and GLS can generate high-density point clouds, enabling the characterization of inner canopy structure (Fig. 10). This may illustrate why GLS are less sensitive to the selection of height quantities, so are TLS and BLS (Fig. 4). Additionally, the GLS system used in this study may lose points near the ground due to the filtering method (Fig. 10b), which illustrated the overall underestimation and relative high bias of GLS-predicted data (Fig. 5c, Fig. 6d–f, Fig. 7c, Fig. 8c). However, it had a slight influence on the overall trend of CH assessment and RMSE ( $r=0.82$ , Fig. 5c). Notably, despite the high point resolution of GLS, its ranging extent is much closer, making it easier to be saturated

**Table 5** Detailed statistics on comparing height measurement methods

GS group	RMSE, m (RMSE%)	Bias, m (Bias%)	<i>r</i>	RMSE, m (RMSE%)	Bias, m (Bias%)	<i>r</i>
	TLS vs. FM			BLS vs. FM		
J	0.04 (6.69)	0.02 (3.37)	0.88	0.05 (9.58)	0.03 (5.32)	0.81
H	0.05 (6.64)	0.04 (5.42)	<u>0.92</u>	0.05 (6.49)	0.04 (4.89)	0.91
F	0.10 (11.82)	0.03 (4.10)	<u>0.76</u>	0.10 (12.03)	0.05 (6.41)	0.76
M	0.08 (9.85)	0.05 (5.99)	<u>0.76</u>	0.07 (9.16)	0.05 (6.26)	0.75
Mean	0.07 (8.75)	0.04 (4.72)	0.83	0.07 (9.32)	0.04 (5.72)	0.81
	GLS vs. FM			DAP vs. FM		
J	0.05 (8.96)	-0.13 (-22.39)	0.79	0.05 (8.94)	0.03 (5.25)	<u>0.89</u>
H	0.05 (7.00)	-0.10 (-13.84)	0.72	0.05 (6.56)	0.05 (5.75)	0.92
F	0.08 (10.96)	-0.08 (-10.99)	0.45	0.14 (16.58)	0.03 (4.13)	0.64
M	0.06 (7.94)	-0.11 (-14.74)	0.62	0.11 (14.18)	0.01 (0.85)	0.60
Mean	0.06 (8.71)	-0.10 (-15.49)	0.65	0.09 (11.57)	0.03 (3.99)	0.76
	TLS vs. BLS			TLS vs. GLS		
J	0.04 (7.61)	0.01 (1.89)	0.87	0.03 (5.15)	-0.15 (-24.93)	<u>0.93</u>
H	0.03 (3.33)	0.00 (-0.50)	<u>0.98</u>	0.03 (3.88)	-0.15 (-20.19)	0.91
F	0.04 (4.44)	0.02 (2.22)	<u>0.97</u>	0.03 (4.28)	-0.17 (-20.79)	0.92
M	0.04 (5.26)	0.00 (0.25)	<u>0.92</u>	0.05 (6.07)	-0.16 (-20.44)	0.76
Mean	0.04 (5.16)	0.01 (0.97)	0.94	0.04 (4.84)	-0.16 (-21.59)	0.88
	BLS vs. DAP			BLS vs. GLS		
J	0.06 (9.68)	0.00 (-0.07)	0.85	0.04 (6.58)	-0.16 (-26.33)	0.88
H	0.05 (5.81)	0.01 (0.82)	0.93	0.04 (4.78)	-0.15 (-20.42)	0.86
F	0.08 (9.24)	-0.02 (-2.15)	0.89	0.03 (3.49)	-0.18 (-22.29)	0.95
M	0.10 (11.93)	-0.04 (-5.09)	0.70	0.05 (6.65)	-0.16 (-20.92)	0.70
Mean	0.07 (9.17)	-0.01 (-1.62)	0.84	0.04 (5.51)	-0.16 (-22.52)	0.86
	DAP vs. TLS			DAP vs. GLS		
J	0.03 (5.10)	-0.01 (-1.78)	0.92	0.04 (7.05)	-0.16 (-26.27)	0.86
H	0.04 (5.27)	0.00 (-0.31)	0.95	0.03 (4.54)	-0.15 (-20.41)	0.87
F	0.07 (8.52)	0.00 (-0.02)	0.87	0.05 (5.87)	-0.17 (-21.14)	0.84
M	0.09 (10.98)	0.04 (5.10)	0.68	0.06 (7.76)	-0.13 (-17.09)	0.61
Mean	0.06 (7.47)	0.01 (0.74)	0.86	0.05 (6.30)	-0.15 (-21.23)	0.80

The top side of the table showed the evaluation results of 3D sensing datasets with FM; the bottom side of the table showed the results of 3D sensing datasets cross-comparisons. RMSE and RMSE%, Bias and Bias%, and correlation coefficient (*r*) were given for distinct growth stage (GS) groups. The underlined values were the best result for each GS group among different comparisons

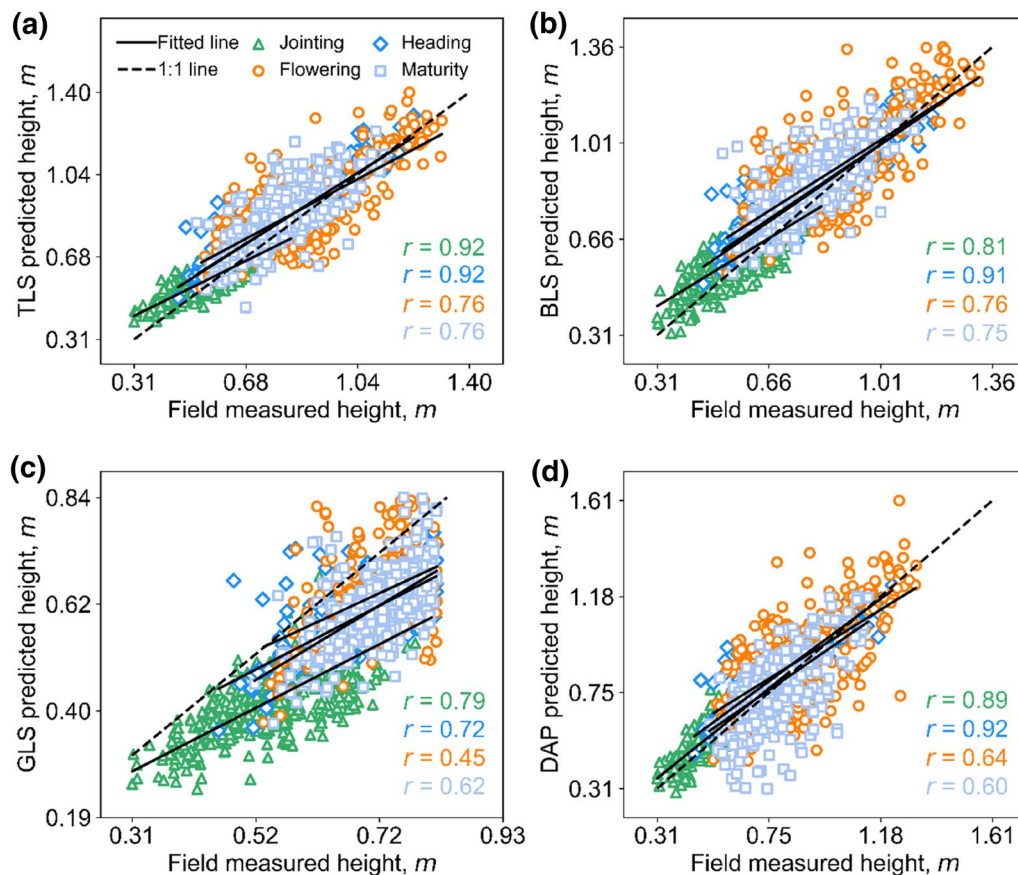
J, H, F, and M represent jointing, heading, flowering, and maturity stages

when predicting higher canopies, which can be seen if all the plots are used for height estimation in this study (Additional file 1: Fig. S6). This suggests that the choice of laser ranging extent is as important as the sensor resolution for high-precision crop phenotyping.

In conclusion, selecting the optimal height quantiles is critical in the evaluation of CH. Despite subtle differences, these best height metrics were very close in performance. Considering the more diverse datasets used in this study than in previous studies [38, 39, 43], the systematic evaluation of 3D sensing methods were unprecedented, which lays reliable foundations for the further cross-comparisons.

### CH estimation under various height groups, LAI groups, and GS groups

The CH estimate accuracies will obviously decrease when evaluated at CH subgroups (Fig. 7). This has been rarely reported in agriculture, but some similar findings have been drawn in forest CH estimation [31, 71]. The subgroup of lower CH plots (e.g., CH1) showed higher correlations (Table 3), which are consistent with previous studies that indicated the uncertainty of CH assessment by 3D sensing increased with height [59]. This may attribute to the increasing canopy complexity (e.g., crop canopy cover and plant density) with height [4, 15]. Meanwhile, canopy senescence and logging may



**Fig. 9** Correlations between FM heights and predicted heights by different 3D sensing technologies under four growth stages (GS) groups. **a–d** represent TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively. The green triangle, blue diamond, dark orange circle, and baby blue square represent the jointing stage, heading stage, flowering stage, and maturity stage respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line

also influence height estimation accuracy at high-height groups.

This study found the TLS, BLS, and DAP showed overestimation in low CH groups (i.e., CH1 and CH2 groups) but are closer to field measurement in CH3 and CH4 groups (Figs. 7, 11). The possible reason is the canopy surface is not closed and looks uneven at the early stage. In this case, field measurement was hard to capture the highest CHs (observation) while the sensor measured height is the globally ranked height quantities (real max. height) of a plot. Although GLS had systematic underestimation due to its limited ranging extent, it had a better fitting effect with TLS and BLS (Additional file 1: Fig. S3), demonstrating the high reliability of ranging precision of 3D sensing technologies under different canopy structures. It is also the high precision of the GLS system (Table 1) that may illustrate why GLS keeps a low RMSE (0.04 m–0.05 m) with other 3D sensing datasets (Fig. 6d–f).

In addition, DAP-estimated height showed lower correlations with other 3D sensing datasets (Additional file 1: Fig. S3). This may be caused by the relatively lower data quality of the DAP point cloud. DAP point cloud was reconstructed from images, which are sensitive to environmental illumination, image quality, and reconstruction algorithms [2, 12, 16]. Some studies have demonstrated that the DAP has comparable accuracy with LiDAR in monitoring canopy height [10]. In this study, we further proved that DAP showed similar better results with LiDAR in field plots with lower CH (e.g., CH1), and found the accuracy would decrease at higher CH groups (Fig. 7). The decreasing accuracy may be caused by the large variations of estimated height at large CH groups where canopy structures are denser and complicated (Fig. 11).

By contrast, the CH estimate accuracies did not show an obvious decrease when evaluated at LAI or GS groups (Figs. 8 and 9). The possible reasons are the height range of data within each LAI or GS subgroup



**Table 6** The values of Broad-sense heritability ( $H^2$ ) from different 3D sensing datasets with regard to different canopy height (CH), leaf area index (LAI), and growth stage (GS) groups

	FM	TLS	BLS	GLS	DAP
CH group					
CH1	0.58	<u>0.81</u>	0.77	0.73	0.74
CH2	0.61	<u>0.65</u>	0.64	0.59	0.57
CH3	-	-	-	-	-
CH4	-	-	-	-	-
Mean	0.60	<b>0.73</b>	<b>0.70</b>	<b>0.66</b>	<b>0.66</b>
LAI group					
LAI1	0.83	<u>0.90</u>	0.85	0.86	0.84
LAI2	-	-	-	-	-
LAI3	-	-	-	-	-
LAI3	-	-	-	-	-
Mean	0.83	<b>0.90</b>	<b>0.85</b>	<b>0.86</b>	<b>0.84</b>
GS group					
Jointing	0.79	<u>0.90</u>	0.85	0.86	0.83
Heading	0.83	<u>0.90</u>	0.85	0.86	0.84
Flowering	0.77	<u>0.94</u>	0.86	0.79	0.85
Maturity	0.70	0.82	<u>0.85</u>	0.73	0.62
Mean	0.77	<b>0.89</b>	<b>0.85</b>	<b>0.81</b>	<b>0.79</b>

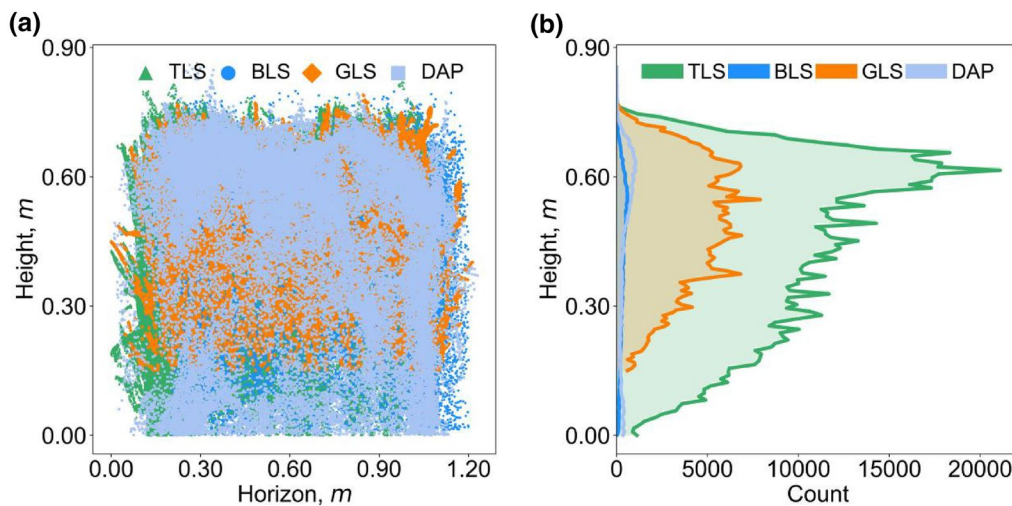
The calculation of  $H^2$  of a variety was based on the variety's four plot CHs, including two N treatments and two replicates. Because only CHs under 0.82 m of FM were used for comparison, the  $H^2$  value of varieties can be calculated only when all four plots' CH of a variety were below 0.82 m. - represents no varieties within the group that meet the above conditions. The underlined values represent the best results among different height measurement methods (each row) with regard to different subgroups of CH, LAI, and GS. The bold values were the mean values of different 3D sensing datasets (each column) across all subgroups of CH, LAI, and GS

was relatively large. However, the accuracy at high LAI or late GS was also relatively lower, which may attribute to the more complex canopy structure [18, 45].

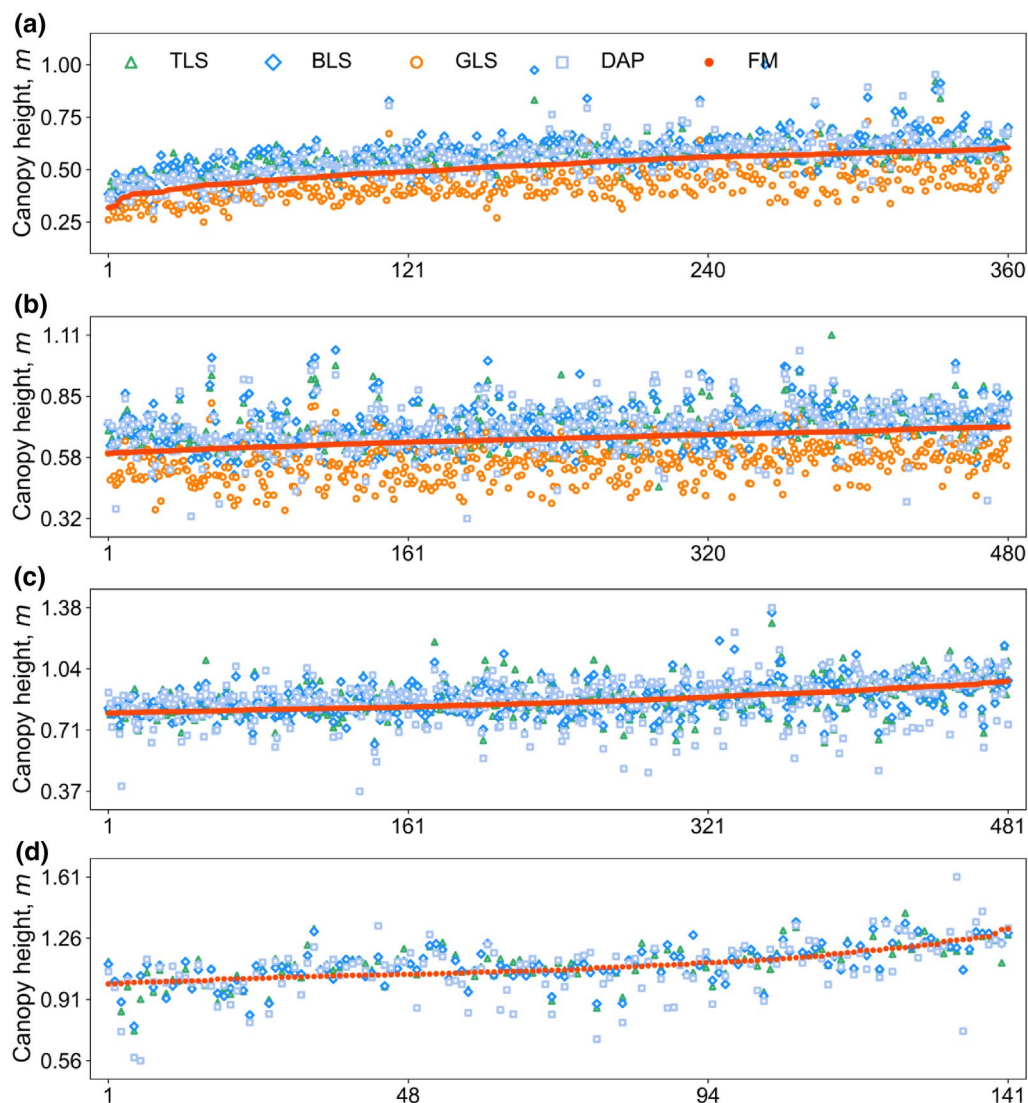
**Outlier analysis of different datasets**

Error source analysis revealed that 8 plots existed FM error according to our definitions in Sect. "Error source analysis" (Fig. 12a). In these plots, heights estimated from all 3D sensing methods were 20% greater than FM, and the heights between different 3D sensing methods were closer. This indicated that FM may be inaccurate. By contrast, there are more potential suspicious CH results estimated from GLS (451), DAP (253), BLS (224), and TLS (164) (Fig. 12). Reasons for why the number of suspicious FM is fewer than other sensors may attribute to the strict judging conditions in Eq. 12. A FM value is suspicious only when it is suspicious to TLS, BLS, GLS, and DAP at the same time. In other words, if a FM is suspected as long as there are more than two suspects in the four kinds of comparisons (FM v.s. TLS, FM v.s. BLS, FM v.s. GLS, FM v.s. DAP), then the number of suspicious FMs will be more (Fig. 12).

In fact, there should be more errors coming from FM. For example, Fig. 12b shows the error source case of TLS, but it can be easily found that most TLS measurements were very consistent with BLS and DAP. This may imply FM and GLS are both suspicious, instead of TLS. Similar more suspicious cases of FM can be found in Fig. 12c–e. Although the overall underestimation of GLS data brought challenges for the above outlier analysis, the



**Fig. 10** a Front view and b frequency distribution of points' height value of TLS, BLS, GLS, and DAP data in the same plot



**Fig. 11** Canopy height (CH) observations from TLS (in the green triangle), BLS (in the blue rhombus), GLS (in the orange circle), DAP (in the blue light square), and FM (in the orange-red point) for **a** CH1, **b** CH2, **c** CH3, and **d** CH4 group. The x-axis represents the sorting order of field plots, and the y-axis represents the value of canopy height

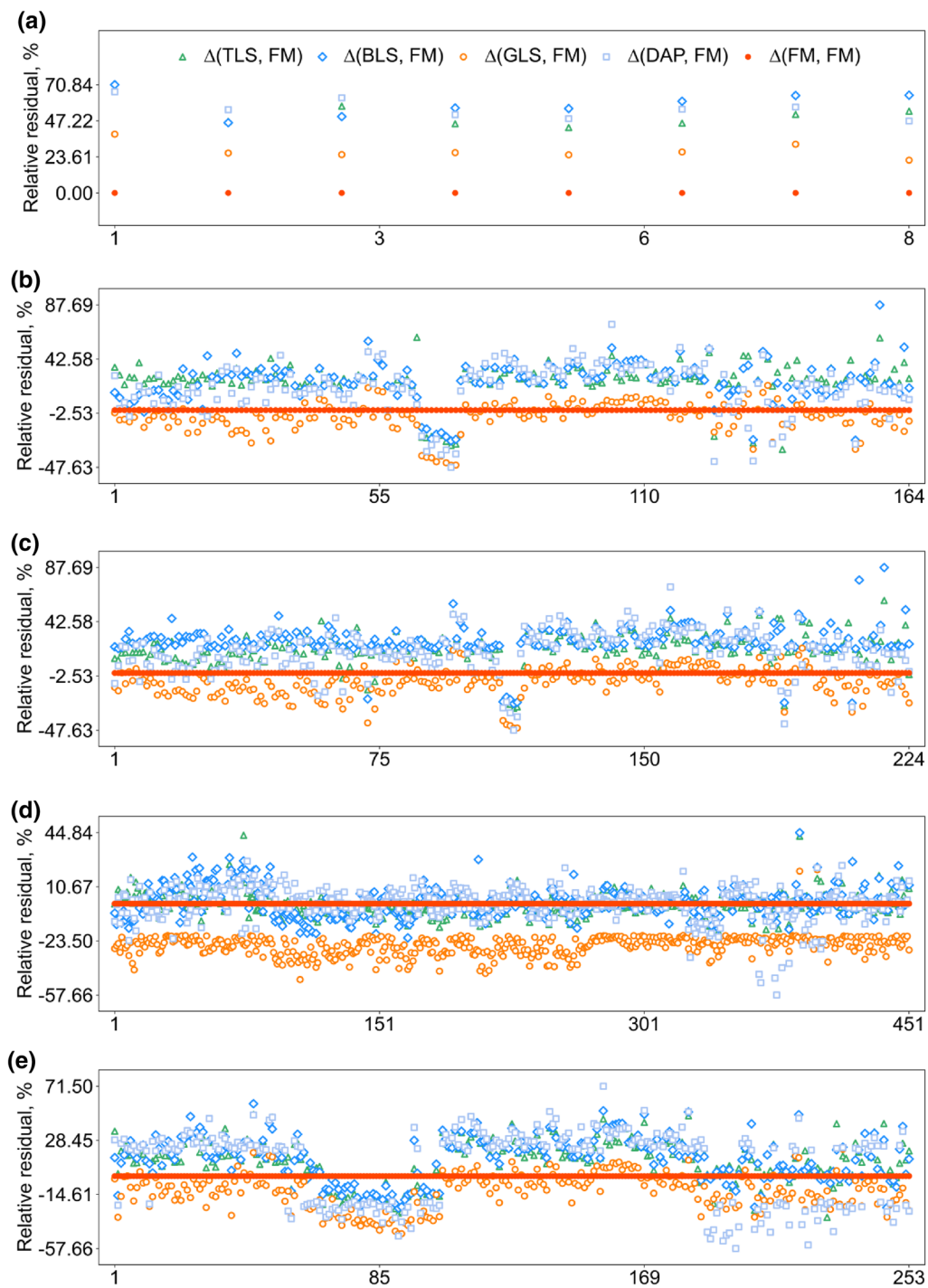
general trends still exist. As for the outlier estimations of GLS, most relative residuals were below  $-20\%$  (Fig. 12d), which was mainly caused by the lack of ground point (Fig. 10b), indicating the importance of ground filtering in CH estimation.

#### Field-measured canopy height may not be as accurate as believed

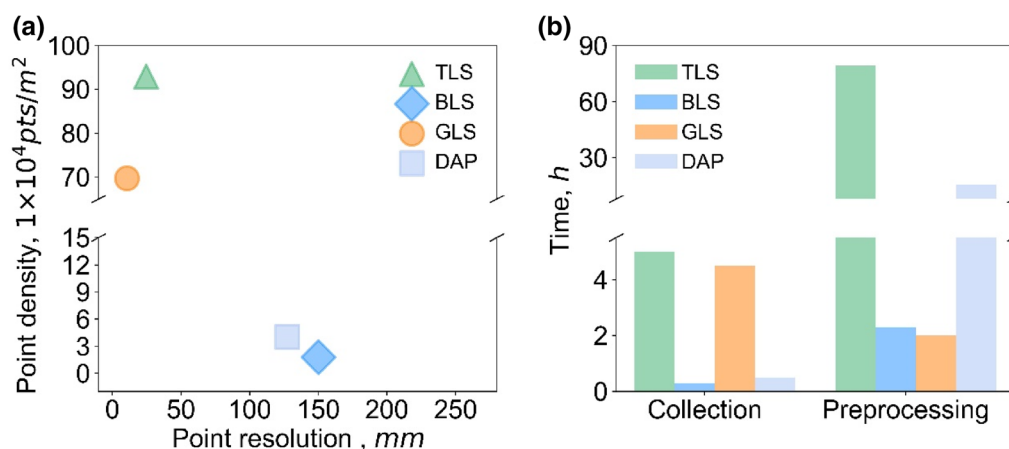
Our results showed that the height correlations between different 3D sensing ( $r=0.87-0.97$ ) are much better than the correlations between 3D sensing and FM ( $r=0.82-0.89$ ). The reasons may be two aspects. On the one hand, LiDAR and DAP are both accurate surveying and

mapping technologies, they have good repeatability and consistency despite a wide variety of sensors and platforms. The TLS, BLS, and GLS systems with centimeter and millimeter resolutions have been proven accurate for estimating not only height but also other 3D traits [68, 76]. On the other hand, FM may be suspicious because it is based on subjective samples and is easily influenced by the terrain and other factors [1]. Some studies have also indicated that LiDAR may be more accurate than manual inspection [44].

Heritability quantifies the repeatability of the canopy height trait estimation, which is another prospect to evaluate the reliability of phenotyping methods and their



**Fig. 12** Relative canopy height residuals between 3D sensing-derived canopy height (CH) and FM. Suspicious results existed in **a** FM, **b** TLS-derived CH, **c** BLS-derived CH, **d** GLS-derived CH, and **e** DAP-derived CH according to Eq. 12–16. The x-axis represents the ID of field plots, and the y-axis represents the value of relative residuals. The green triangle, blue rhombus, orange circle, light blue square, and orange-red point represent the relative residuals of  $\Delta(\text{TLS, FM})$ ,  $\Delta(\text{BLS, FM})$ ,  $\Delta(\text{GLS, FM})$ ,  $\Delta(\text{DAP, FM})$ , and  $\Delta(\text{FM, FM})$  in each subplot. Among them, the  $\Delta(\text{FM, FM})$  value is zero, which looks like a horizontal reference line in red



**Fig. 13** **a** Point accuracy and point density and **b** collection time and processing time from TLS (in the green triangle), BLS (in the blue rhombus), GLS (in the orange circle), and DAP (in the blue light square)

potential for the breeding program [52]. On the one hand, the differences in the heritability of different data reflect their ability to characterize the subtle differences of CH among different varieties, as mentioned by Volpato et al. [67]. Our results proved that  $H^2$  of 3D sensing, especially  $H^2$  of LiDAR-derived CH, was larger than that derived from FM (Table 6), which may be determined by the higher accuracy of LiDAR systems. Higher  $H^2$  of the advanced 3D sensing tools indicate that they will facilitate better trait extraction for breeding. On the other hand, the overall heritability in the later growth period decreased, which may attribute to the prominent environmental impact of nitrogen treatment in the later growth period (e.g., logging). The environmental effects on the heritability of LiDAR-derived plant height have been proved by Madec et al. [43]. The dynamic change of  $H^2$  would be interested by plant breeders for selecting the right time to study the genotypic and/or environmental influences on phenotype [67].

### Contributions and implications

This study systematically evaluated the accuracy of CH estimation from advanced 3D sensing systems (TLS, BLS, GLS, and DAP) and FM using wheat plots of different varieties, fertilization levels, and growth stages. To our knowledge, this is the first effort that uses multiple 3D sensing technologies to evaluate their reliability for estimating CH with regard to different CH, LAI, and GS groups. Moreover, we analyzed the heritability from 3D sensing datasets and FM, proving the potential advantages of 3D sensing technologies in crop breeding.

However, there are still some interesting and important directions that need to be explored in the future. First, it is meaningful to deeply analyze the effect of operating

modes of different 3D sensing technologies on CH monitoring. As for TLS, the scanning location settings (e.g., positions and total numbers) is important for acquiring a high-quality (higher density and less occlusion) point cloud [14, 68]. Although some pioneer studies have been conducted in forestry [74], it is still needed to have a scientific workflow of TLS in agriculture to ensure not only high accuracy but also improve efficiency. BLS is an economically friendly and easy-to-use platform. Designing the routine is critical and it has been discussed by Su et al. [58]. GLS is a kind of emerging phenotyping platform, which is mainly designed for crop phenotyping and has less been explored. This study highlights the necessity to integrate suitable sensors (e.g., longer-ranging ability) for different crop types, provide access to raw data, and enable more intelligent custom algorithms (e.g., filtering algorithm) for accurate phenotype extraction [29]. DAP is a low-cost system that has been widely used in phenotyping. However, the point cloud quality generated from DAP is affected by parameters such as sensor quality, camera shooting angle, routine overlap, and flight speed. This study determined the optimal flight height by a preliminary comparison experiment (Additional file 1: Fig. S1). More parameter comparison studies are worth exploring and can refer to Hu et al. [20]. Additionally, considering the UAV-LiDAR systems are more expensive than DAP and do not have obvious advantages in data quality [76], this study did not compare the UAV-LiDAR systems. However, we believe UAV-LiDAR systems are getting cheaper and the data quality is a good complement to DAP due to its higher penetration ability and robustness to light environments.

Secondly, the tradeoff between precision and efficiency is worth studying. Generally, data precision was depicted

by point density and resolution. High point density usually has a high resolution (Fig. 13a). The possible reason why TLS has a higher point density but a lower resolution is the multi-scan registration [37]. More importantly, this study highlights that higher precision always needs a longer collection time, but does not mean more processing time (e.g., GLS) (Fig. 13). Among them, TLS has the longest data acquisition and processing time, because the reference targets and scanner need to be laboriously laid out during the scanning, and multi-scan data registration is time-consuming during reprocessing [7]. BLS has the shortest time (collection plus preprocessing), implicating this type of mobile mapping technology is worth promoting in the future, especially as cost decreases and accuracy increases. GLS not only has the highest point resolution but also has the shortest preprocessing time and affordable collection time, which benefits from the automatic data collection system and processing software [36]. However, this kind of phenotyping platform is still too expensive (Table 2). DAP has high collection efficiency, but the data quality is relatively low. Besides, the processing time of DAP is long not only caused by 3D reconstruction but also attribute to the manual denoising process due to the low signal-to-noise ratio of the DAP point cloud. These preliminary explorations are of great significance for further in-depth and systematic analysis of cost and efficiency and the formulation of appropriate phenotypic working plans.

Finally, there is no standard for grouping CH and LAI. This study mainly divided 1920 plots into four different groups based on the value extent (maximum minus minimum) and frequency distribution. Although there are small differences in the spacing of the groupings and the number of groups is not exactly equal, the total sample sizes (i.e., 1920 plots) are unprecedented. The influence of CH, LAI, and GS on height measurement accuracy and heritability has been analyzed, but more quantitative evaluations are worth exploring, such as the specific CH and LAI thresholds for selecting the optimal measuring methods. Moreover, this study mainly studied the important CH trait in wheat, while more biologically meaningful and heritable traits in more crop types need further evaluation [35, 78].

## Conclusion

The study demonstrated novel insights into the accuracy and heritability of CH from 3D sensing and field measurement. Cross-comparisons among different sensor datasets showed higher correlations ( $r=0.87$  to  $0.97$ ) than comparisons with FM ( $r=0.82$  to  $0.89$ ). The correlation coefficients of CHs derived from 3D sensing and FM decreased obviously when evaluated with respect to

different subgroups (CH, LAI, and GS), especially different CH subgroups. TLS and BLS were more reliable in monitoring CH under different subgroups according to their cross-comparisons and comparisons with FM. The outlier analysis found cases where FM may be error-prone. Moreover, 3D sensing methods showed even higher heritability than FM. Further studies about the best configurations of sensors and working plans are needed, the tradeoff between data quality and efficiency is worth exploring, and more traits deserve future efforts. These novel findings may give insights into the selection of advanced 3D sensing platforms for crop monitoring and may shed new light on the high-quality development of crop sciences (e.g. providing higher heritable traits for breeding).

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13007-023-01012-2>.

**Additional file 1: Fig. S1.** Canopy height correlations between the field measurement and DAP estimates at different flight heights. (a), (b), (c), and (d) are the results of flight height at 10, 20, 30, and 40m, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 line. The color bar shows the kernel density value of the point distribution, and the green to yellow represents the increase in kernel density. **Fig. S2.** (a) RMSE, (b) RMSE%, (c) Bias, and (d) Bias% between field measured height (FM) and different height quantiles (Hmax and H99) derived from the different 3D point cloud, including TLS, BLS, GLS, and DAP. The green triangle, blue diamond, dark orange circle, and baby blue square represent TLS vs. FM, BLS vs. FM, GLS vs. FM, and DAP vs. FM, respectively. **Fig. S3.** Correlations of cross-comparisons between different 3D sensing data estimated canopy height (CH) at four CH subgroups. (a), (b), (c), (d), (e), and (f) are comparisons of TLS vs. BLS, BLS vs. DAP, DAP vs. TLS, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS, respectively. The green triangle, blue diamond, orange circle, and light blue square represent the CH1, CH2, CH3, and CH4 groups, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 reference. **Fig. S4.** Correlations of cross-comparisons between different 3D sensing data estimated canopy height (CH) at four leaf area index (LAI) groups. (a), (b), (c), (d), (e), and (f) are comparisons of TLS vs. BLS, BLS vs. DAP, DAP vs. TLS, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS, respectively. The green triangle, blue diamond, orange circle, and light blue square represent the LAI1, LAI2, LAI3, and LAI4 groups, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 reference line. **Fig. S5.** Correlations of cross-comparisons between different 3D sensing data estimated canopy height (CH) at four growth stages (GS) groups. (a), (b), (c), (d), (e), and (f) are comparisons of TLS vs. BLS, BLS vs. DAP, DAP vs. TLS, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS, respectively. The green triangle, blue diamond, orange circle, and light blue square represent the jointing, heading, flowering, and maturity stages, respectively. The solid line represents the fitted line, and the dashed line represents the 1:1 reference line. **Fig. S6.** Correlations between GLS predicted datasets (including the measured canopy height over 0.82m) and other datasets. (a), (b), (c), and (d) respectively represent GLS vs. FM, TLS vs. GLS, BLS vs. GLS, and DAP vs. GLS. The solid line represents the fitted line, and the dashed line represents the 1:1 line. The color bar shows the kernel density value of the point distribution, and the green to yellow represents the increase in kernel density. **Table S1.** The performance of canopy height measurement with TLS, BLS, GLS, and DAP systems.  $R^2$ ,  $r$ , and RMSE are the coefficient of determination, correlation coefficient, and root mean square error, respectively. – represents the performance metric that is not available.

**Acknowledgements**

Not applicable.

**Author contributions**

SJ designed this study; JZ, QL, and YM collected the data; JZ, SZ, and SL processed the data; JZ and SJ prepared the figures and tables; JZ and SJ wrote the manuscript; All authors helped to revise the manuscript. All authors read and approved the final manuscript.

**Funding**

This study was supported by the National Key Research and Development Program of China (2022YFD1900701), the National Natural Science Foundation of China (32201656), the Jiangsu Agricultural Science and Technology Independent Innovation Fund Project (CX (22) 1006), the JBGS Project of Seed Industry Revitalization in Jiangsu Province (JBGS [2021] 007), the High-Level Personnel Project of Jiangsu Province (JSSCBS20210271), and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24020202).

**Data availability**

The data that supports the findings of this study are available on request due to its large volume.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Competing interests**

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

**Author details**

<sup>1</sup>Plant Phenomics Research Centre, Academy for Advanced Interdisciplinary Studies, Collaborative Innovation Centre for Modern Crop Production Co-Sponsored By Province and Ministry, College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China. <sup>2</sup>State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.

Received: 31 December 2022 Accepted: 21 March 2023

Published online: 02 April 2023

**References**

- Aasen H, Burkart A, Bolten A, Bareth G. Generating 3d hyperspectral information with lightweight uav snapshot cameras for vegetation monitoring: from camera calibration to quality assurance. *ISPRS J Photogramm Remote Sens*. 2015;108:245–59.
- Ali-Sisto D, Gopalakrishnan R, Kukkonen M, Savolainen P, Packalen P. A method for vertical adjustment of digital aerial photogrammetry data by using a high-quality digital terrain model. *Int J Appl Earth Obs Geoinf*. 2020;84:101954–63.
- Bai D, Li D, Zhao C, Wang Z, Shao M, Guo B, Liu Y, Wang Q, Li J, Guo S, Wang R, Li YH, Qiu LJ, Jin X. Estimation of soybean yield parameters under lodging conditions using rgb information from unmanned aerial vehicles. *Front Plant Sci*. 2022;13:1012293–312.
- Busemeyer L, Mentrup D, Moller K, Wunder E, Alheit K, Hahn V, Maurer HP, Reif JC, Wurschum T, Muller J, Rahe F, Ruckelshausen A. Breedvision—a multi-sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors*. 2013;13(3):2830–47.
- Chandana BS, Mahto RK, Singh RK, Ford R, Vaghefi N, Gupta SK, Yadav HK, Manohar M, Kumar R. Epigenomics as potential tools for enhancing magnitude of breeding approaches for developing climate resilient chickpea. *Front Genet*. 2022;13:900253–78.
- Chen J, Black TA. Defining leaf area index for non-flat leaves. *Plant, Cell Environ*. 1992;15:421–9.
- Disney M. Terrestrial lidar: a three-dimensional revolution in how we look at trees. *New Phytol*. 2019;222(4):1736–41.
- Eitel JUH, Höfle B, Vierling LA, Abellán A, Asner GP, Deems JS, Glennie CL, Joerg PC, LeWinter AL, Magney TS, Mandlburger G, Morton DC, Müller J, Vierling KT. Beyond 3-d: the new spectrum of lidar applications for earth and ecological sciences. *Remote Sens Environ*. 2016;186:372–92.
- El-Naggar AG, Jolly B, Hedley CB, Horne D, Roudier P, Clothier BE. The use of terrestrial lidar to monitor crop growth and account for within-field variability of crop coefficients and water use. *Comput Electron Agric*. 2021;190(6):106416–32.
- Gao M, Yang F, Wei H, Liu X. Individual maize location and height estimation in field from uav-borne lidar and rgb images. *Remote Sensing*. 2022;14(10):2292–311.
- Gong Y, Yang K, Lin Z, Fang S, Wu X, Zhu R, Peng Y. Remote estimation of leaf area index (lai) with unmanned aerial vehicle (uav) imaging for different rice cultivars throughout the entire growing season. *Plant Methods*. 2021;17(1):88–104.
- Goodbody TRH, Coops NC, White JC. Digital aerial photogrammetry for updating area-based forest inventories: a review of opportunities, challenges, and future directions. *Current Forestry Reports*. 2019;5(2):55–75.
- Grenzdörffer GJ. Crop height determination with UAS point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. 2014;40(1):135–40.
- Guo T, Fang Y, Cheng T, Tian Y, Zhu Y, Chen Q, Qiu X, Yao X. Detection of wheat height using optimized multi-scan mode of lidar during the entire growth stages. *Comput Electron Agric*. 2019;165(6):104959–68.
- Han L, Yang G, Dai H, Xu B, Yang H, Feng H, Li Z, Yang X. Modeling maize above-ground biomass based on machine learning approaches using uav remote-sensing data. *Plant Methods*. 2019;15(1):10–29.
- Han X, Thomasson JA, Bagnall GC, Pugh NA, Horne DW, Rooney WL, Jung J, Chang A, Malambo L, Popescu SC, Gates IT, Cope DA. Measurement and calibration of plant-height from fixed-wing uav images. *Sensors (Basel)*. 2018;18(12):4092–113.
- Hartley RJL, Leonardo EM, Massam P, Watt MS, Estarija HJ, Wright L, Melia N, Pearse GD. An assessment of high-density uav point clouds for the measurement of young forestry trials. *Remote Sensing*. 2020;12(24):4039–59.
- Holman F, Riche A, Michalski A, Castle M, Wooster M, Hawkesford M. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using uav based remote sensing. *Remote Sensing*. 2016;8(12):1031–55.
- Hu P, Chapman SC, Wang X, Potgieter A, Duan T, Jordan D, Guo Y, Zheng B. Estimation of plant height using a high throughput phenotyping platform based on unmanned aerial vehicle and self-calibration: example for sorghum breeding. *Eur J Agron*. 2018;95:24–32.
- Hu T, Sun X, Su Y, Guan H, Sun Q, Kelly M, Guo Q. Development and performance evaluation of a very low-cost uav-lidar system for forestry applications. *Remote Sensing*. 2020;13(1):77–98.
- Hütt C, Bolten A, Hüging H, Bareth G. Uav lidar metrics for monitoring crop height, biomass and nitrogen uptake: a case study on a winter wheat field trial. *Photogramm Remote Sens Geoinform Sci*. 2022;22(6):2512–819.
- Hyyppä E, Yu X, Kaartinen H, Hakala T, Kukko A, Vastaranta M, Hyyppä J. Comparison of backpack, handheld, under-canopy uav, and above-canopy uav laser scanning for field reference data collection in boreal forests. *Remote Sens*. 2020;12(20):3327–58.
- James MR, Robson S. Mitigating systematic error in topographic models derived from uav and ground-based image networks. *Earth Surf Proc Land*. 2014;39(10):1413–20.
- Jelle ten Harkel J, Bartholomeus H, Kooistra L. Biomass and crop height estimation of different crops using uav-based lidar. *Remote Sens*. 2019;12(1):17–35.
- Jin S, Su Y, Gao S, Wu F, Hu T, Liu J, Li W, Wang D, Chen S, Jiang Y, Pang S, Guo Q. Deep learning: Individual maize segmentation from terrestrial lidar data using faster r-cnn and regional growth algorithms. *Front Plant Sci*. 2018;9:866–75.
- Jin S, Su Y, Song S, Xu K, Hu T, Yang Q, Wu F, Xu G, Ma Q, Guan H, Pang S, Li Y, Guo Q. Non-destructive estimation of field maize biomass using terrestrial lidar: an evaluation from plot level to individual leaf level. *Plant Methods*. 2020;16(3):69–89.
- Jin S, Su Y, Wu F, Pang S, Gao S, Hu T, Liu J, Guo Q. Stem–leaf segmentation and phenotypic trait extraction of individual maize using terrestrial lidar data. *IEEE Trans Geosci Remote Sens*. 2019;57(3):1336–46.

28. Jin S, Su Y, Zhang Y, Song S, Li Q, Liu Z, Ma Q, Ge Y, Liu L, Ding Y, Baret F, Guo Q. Exploring seasonal and circadian rhythms in structural traits of field maize from lidar time series. *Plant Phenomics*. 2021;2021(4):9895241–56.
29. Jin S, Su Y, Zhao X, Hu T, Guo Q. A point-based fully convolutional neural network for airborne lidar ground point filtering in forested environments. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2020;13:3958–74.
30. Jin S, Sun X, Wu F, Su Y, Li Y, Song S, Xu K, Ma Q, Baret F, Jiang D, Ding Y, Guo Q. Lidar sheds new light on plant phenomics for plant breeding and management: recent advances and future prospects. *ISPRS J Photogramm Remote Sens*. 2021;171(5):202–23.
31. Kawamura K, Asai H, Yasuda T, Khanthavong P, Soisouvanh P, Phongchanmixay S. Field phenotyping of plant height in an upland rice field in Laos using low-cost small unmanned aerial vehicles (uavs). *Plant Prod Sci*. 2020;23(4):452–65.
32. Ko C, Lee S, Yim J, Kim D, Kang J. Comparison of forest inventory methods at plot-level between a backpack personal laser scanning (bpls) and conventional equipment in Jeju Island, South Korea. *Forests*. 2021;12(3):308–21.
33. Kronenberg L, Yu K, Walter A, Hund A. Monitoring the dynamics of wheat stem elongation: genotypes differ at critical stages. *Euphytica*. 2017;213(7):157–70.
34. Li A, Hao C, Wang Z, Geng S, Jia M, Wang F, Han X, Kong X, Yin L, Tao S, Deng Z, Liao R, Sun G, Wang K, Ye X, Jiao C, Lu H, Zhou Y, Liu D, Fu X, Zhang X, Mao L. Wheat breeding history reveals synergistic selection of pleiotropic genomic sites for plant architecture and grain yield. *Mol Plant*. 2022;15(3):504–19.
35. Li D, Shi G, Li J, Chen Y, Zhang S, Xiang S, Jin S. Plantnet: a dual-function point cloud segmentation network for multiple plant species. *ISPRS J Photogramm Remote Sens*. 2022;184:243–63.
36. Li Q, Jin S, Zang J, Wang X, Sun Z, Li Z, Xu S, Ma Q, Su Y, Guo Q, Jiang D. Deciphering the contributions of spectral and structural data to wheat yield estimation from proximal sensing. *The Crop Journal*. 2022;11:2214–25.
37. Liang X, Kankare V, Hyypä J, Wang Y, Kukko A, Haggrén H, Yu X, Kaartinen H, Jaakkola A, Guan F, Holopainen M, Vastaranta M. Terrestrial laser scanning in forest inventories. *ISPRS J Photogramm Remote Sens*. 2016;115:63–77.
38. Liu F, Hu P, Zheng B, Duan T, Zhu B, Guo Y. A field-based high-throughput method for acquiring canopy architecture using unmanned aerial vehicle images. *Agric For Meteorol*. 2021;296:108231–42.
39. Lu J, Cheng D, Geng C, Zhang Z, Xiang Y, Hu T. Combining plant height, canopy coverage and vegetation index from UAV-based RGB images to estimate leaf nitrogen concentration of summer maize. *Biosys Eng*. 2021;202:42–54.
40. Luo S, Liu W, Zhang Y, Wang C, Xi X, Nie S, Ma D, Lin Y, Zhou G. Maize and soybean heights estimation from unmanned aerial vehicle (UAV) lidar data. *Comput Electron Agric*. 2021;182(9):106005–14.
41. Luo S, Wang C, Pan F, Xi X, Li G, Nie S, Xia S. Estimation of wetland vegetation height and leaf area index using airborne laser scanning data. *Ecol Ind*. 2015;48:550–9.
42. Ma H, Liang S. Development of the glass 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model. *Remote Sens Environ*. 2022;273:112985–3003.
43. Madec S, Baret F, de Solan B, Thomas S, Dutartre D, Jezequel S, Hemmerle M, Colombeau G, Comar A. High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground lidar estimates. *Front Plant Science*. 2017;8(1):2002–16.
44. Maesano M, Khoury S, Nakhle F, Firrieceli A, Gay A, Tauro F, Harfouche A. UAV-based lidar for high-throughput determination of plant height and above-ground biomass of the bioenergy grass *Arundo donax*. *Remote Sens*. 2020;12(20):3464–84.
45. Malambo L, Popescu SC, Murray SC, Putman E, Pugh NA, Horne DW, Richardson G, Sheridan R, Rooney WL, Avant R, Vidrine M, McCutchen B, Baltensperger D, Bishop M. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int J Appl Earth Obs Geoinf*. 2018;64:31–42.
46. Niu Y, Zhang L, Zhang H, Han W, Peng X. Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. *Remote Sensing*. 2019;11(11):1261–82.
47. Ogunbadewa EY. Tracking seasonal changes in vegetation phenology with a sunscan canopy analyzer in northwestern England. *For Sci Technol*. 2012;8(3):161–72.
48. Oumata S, Monneveux P, Zaharieva M, Mekliche-Hanifi L, David J. Variation of morphological traits among wheat (*Triticum aestivum* L.) landraces from two regions of the Algerian Sahara. Potential interest for wheat breeding. *Genet Resour Crop Evol*. 2022;1:429–45.
49. Potter, E., Wood, J. & Nicholl, C. (1996). Sunscan canopy analysis system: Users manual. In Delta-T Devices, Cambridge, UK. <https://www.manualslib.com/manual/37695/Delta-Ss1-Um-1-05.html#manual>
50. Ravi R, Lin Y-J, Shamseldin T, Elbahnasawy M, Masjedi A, Crawford M, Habib A. Wheel-based lidar data for plant height and canopy cover evaluation to aid biomass prediction. *IEEE Int Geosci Remote Sens Symp*. 2018;3(3):3242–5.
51. Rosnell T, Honkavaara E. Point cloud generation from aerial image data acquired by a quadcopter type micro unmanned aerial vehicle and a digital still camera. *Sensors*. 2012;12(1):453–80.
52. Schmidt P, Hartung J, Bennewitz J, Piepho HP. Heritability in plant breeding on a genotype-difference basis. *Genetics*. 2019;212(4):991–1008.
53. Singh V, Singh D, Singh N, Kumar S. Genetic analysis of wheat varieties for yield and its components. *Agric*. 2005;25(2):145–6.
54. Sofonia J, Shendryk Y, Phinn S, Roelfsema C, Kendoul F, Skocaj D. Monitoring sugarcane growth response to varying nitrogen application rates: a comparison of UAV SLAM lidar and photogrammetry. *Int J Appl Earth Obs Geoinf*. 2019;82:101878–93.
55. Sone C, Saito K, Futakuchi K. Comparison of three methods for estimating leaf area index of upland rice cultivars. *Crop Sci*. 2009;49(4):1438–43.
56. Song P, Wang J, Guo X, Yang W, Zhao C. High-throughput phenotyping: breaking through the bottleneck in future crop breeding. *The Crop Journal*. 2021;9(3):633–45.
57. Su W, Zhang M, Bian D, Liu Z, Huang J, Wang W, Wu J, Guo H. Phenotyping of corn plants using unmanned aerial vehicle (UAV) images. *Remote Sensing*. 2019;11(17):2021–140.
58. Su Y, Guo Q, Jin S, Guan H, Sun X, Ma Q, Hu T, Wang R, Li Y. The development and evaluation of a backpack lidar system for accurate and efficient forest inventory. *IEEE Geosci Remote Sens Lett*. 2021;18(9):1660–4.
59. Sun S, Li C, Paterson A. In-field high-throughput phenotyping of cotton plant height using lidar. *Remote Sensing*. 2017;9(4):377–98.
60. Sun S, Li C, Paterson AH, Jiang Y, Xu R, Robertson JS, Snider JL, Chee PW. In-field high throughput phenotyping and cotton plant growth analysis using lidar. *Front Plant Sci*. 2018;9(1):16–33.
61. Sun Z, Li Q, Jin S, Song Y, Xu S, Wang X, Cai J, Zhou Q, Ge Y, Zhang R, Zang J, Jiang D. Simultaneous prediction of wheat yield and grain protein content using multitask deep learning from time-series proximal sensing. *Plant Phenomics*. 2022;2022(3):1–13.
62. Tao H, Xu S, Tian Y, Li Z, Ge Y, Zhang J, Wang Y, Zhou G, Deng X, Zhang Z, Ding Y, Jiang D, Guo Q, Jin S. Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives. *Plant Communications*. 2022;3:100344–83.
63. Tilly N, Hoffmeister D, Cao Q, Huang S, Lenz-Wiedmann V, Miao Y, Bareth G. Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *J Appl Remote Sens*. 2014;8:083671–0836693.
64. Tilly, N., Hoffmeister, D., Schiedung, H., Hütt, C., Brands, J. & Bareth, G. Terrestrial laser scanning for plant height measurement and biomass estimation of maize. *Int Arch Photogramm, Remote Sens Spat Inform Sci*. <https://doi.org/10.5194/isprsarchives-XL-7-181-2014> (2014).
65. Tumbo SD, Salyani M, Whitney JD, Wheaton TA, Miller WM. Investigation of laser and ultrasonic ranging sensors for measurements of citrus canopy volume. *Appl Eng Agric*. 2002;18:367–72.
66. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet*. 2008;9(4):255–66.
67. Volpato L, Pinto F, Gonzalez-Perez L, Thompson IG, Borem A, Reynolds M, Gerard B, Molero G, Rodrigues FA Jr. High throughput field phenotyping for plant height using UAV-based RGB imagery in wheat breeding lines: feasibility and validation. *Front Plant Sci*. 2021;12: 591587.
68. Walter JDC, Edwards J, McDonald G, Kuchel H. Estimating biomass and canopy height with lidar for field crop breeding. *Front Plant Sci*. 2019;10:1145–61.

69. Wang D, Li R, Zhu B, Liu T, Sun C, Guo W. Estimation of wheat plant height and biomass by combining uav imagery and elevation data. *Agriculture*. 2022;13(1):9–30.
70. Wang X, Singh D, Marla S, Morris G, Poland J. Field-based high-throughput phenotyping of plant height in sorghum using different sensing technologies. *Plant Methods*. 2018;14:53–69.
71. Wang Y, Lehtomäki M, Liang X, Pyörälä J, Kukko A, Jaakkola A, Liu J, Feng Z, Chen R, Hyypä J. Is field-measured tree height as reliable as believed – a comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS J Photogramm Remote Sens*. 2019;147:132–45.
72. Wang Y, Yang X-D, Ali A, Lv G-H, Long Y-X, Wang Y-Y, Ma Y-G, Xu C-C. Flowering phenology shifts in response to functional traits, growth form, and phylogeny of woody species in a desert area. *Front Plant Sci*. 2020;11:536–47.
73. Xiao Q, Bai X, Zhang C, He Y. Advanced high-throughput plant phenotyping techniques for genome-wide association studies: a review. *J Adv Res*. 2022;35:215–30.
74. Yrttimaa T, Saarinen N, Kankare V, Liang X, Hyypä J, Holopainen M, Vastaranta M. Investigating the feasibility of multi-scan terrestrial laser scanning to characterize tree communities in southern boreal forests. *Remote Sensing*. 2019;11(12):1423–45.
75. Zhang C, Craine WA, McGee RJ, Vandemark GJ, Davis JB, Brown J, Hulbert SH, Sankaran S. High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. *Agron J*. 2021;113(4):3269–80.
76. Zhang F, Hassanzadeh A, Kikkert J, Pethybridge SJ, van Aardt J. Comparison of uas-based structure-from-motion and lidar for structural characterization of short broadacre crops. *Remote Sensing*. 2021;13(19):3975–96.
77. Zhou L, Gu X, Cheng S, Yang G, Shu M, Sun Q. Analysis of plant height changes of lodged maize using uav-lidar data. *Agriculture*. 2020;10(5):146–60.
78. Zhu Y, Sun G, Ding G, Zhou J, Wen M, Jin S, Zhao Q, Colmer J, Ding Y, Ober ES, Zhou J. Large-scale field phenotyping using backpack lidar and cropquant-3d to measure structural variation in wheat. *Plant Physiol*. 2021;187(2):716–38.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

