Laser Triangulation for Crop Canopy Measurements

Roberto M Buelvas, Viacheslav I Adamchuk

McGill University. Department of Bioresource Engineering. Macdonald Campus. 21 111
Lakeshore, Ste-Anne-de-Bellevue, QC, Canada

**A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada**

Abstract. *From a Precision Agriculture perspective, it is important to detect field areas where variabilities in the soil are significant or where there are different levels of crop yield or biomass. Information describing the behavior of the crop at any specific point in the growing season typically leads to improvements in the manner the local variabilities are addressed. The proper use of dense, in-season sensor data allows farm managers to optimize harvest plans and shipment schedules under variable plant growth dynamics, which may originate from soil spatial variability and management conditions. Sensing of crop architectonics has been used as a diagnostic tool in this context. Moving from the subjective visual estimation of farm workers to automated sensing technologies allows for improved repeatability and savings in cost, time, and labor. The goal of this paper is to report on the evaluation of a prototype sensor system embedded in a portable, low-cost instrument for green vegetable production. The prototype system is currently in its second iteration, featuring improvements for issues found in a previous experiment. The system involves circular scanning of crop canopies to identify crop biomass yield using laser triangulation. The results of these scans are height profiles along an angular position from 0° to 360°, which are the input for the biomass estimation. Two approaches for processing the laser-based height profiles are discussed: regression of profile-representative features and inference of a canopy density function. An experiment was conducted in a spinach field of a commercial farm in Sherrington, Quebec, Canada. The coefficient of determination (R^2) for regression between measured and predicted biomass was 0.78 and 0.94. The root mean square error (RMSE) was in turn 4.18 and 2.16 t/ha. The results indicate that the developed sensor system would be a suitable tool for rapid assessment of fresh biomass in the field. Its application would be beneficial in the process of optimizing crop management logistics, comparing the performance of different varieties of crops, and detecting potential stresses in a field environment.*

Keywords: *biomass, laser, phenotyping, crop sensing*

Introduction

Crop biomass is used as an indicator of plant growth in plant phenotyping and as an estimator of yield in agriculture (Golzarian et al., 2011; Van Henten, 1994). Crop biomass refers to the mass of the crop composed of live cells. In the present document, when biomass is mentioned, it refers to aboveground biomass. From the perspective of precision agriculture (PA), it would be useful to identify zones of the field where there are variabilities in the soil or where the crop produces different levels of biomass. This information could provide an improvement in the way in which local heterogeneities are addressed.

According to Catchpol & Wheeler (1992), aboveground biomass is usually measured by destructive methods. Several plants are wasted using this procedure to provide sufficient data to determine plant growth in the field. Non-destructive methods for measuring biomass are desirable, especially with a sensor-to-plant concept, as stated by Golzarian et al. (2011). Such a procedure would make high-throughput data collection possible, where the final users could take advantage of a practical implementation.

Most of the current non-destructive alternatives exploit the relationship between crop aboveground biomass and canopy properties like plant height, total volume or Leaf Area Index (LAI), all of which can be reliably measured without harming the crop (Tumbo et al., 2001; Biskup et al., 2007; Freeman et al., 2007; Ehlert et al., 2008; Reusch, 2009; Rosell et al., 2009; Keightleya and Bawdenb, 2010; Moorthya et al., 2010; Eitel et al., 2014; Kjaer and Ottosen, 2015; Tilly et al., 2015; Schaefer and Lamb, 2016; Su, 2017). With this approach, high accuracy can be achieved, depending on the technology used and the number of degrees of freedom involved in the measurement. The main limiting factor for this scenario is cost, driven by the increasing complexity of the developed systems.

The present work is relevant for considering the integration of lasers with other sensors and for a comparison with ultrasonic measurements; furthermore, its use is justified as a possible way to avoid the increased costs of 3D laser devices. Also, the application of this technique in lettuce and other horticultural crops is seldom found in the literature. It is direct continuation of Buelvas & Adamchuk (2017).

The objective of this study was to develop and evaluate a laser-based sensor system for the indirect measurement of aboveground crop biomass suitable for in situ deployment. The completion of this objective would be a step closer to the ultimate goal of equipping farmers with a tool for the rapid, non-destructive, reliable, and affordable assessment of their crops. Specific objectives were: (1) to develop the system prototype, (2) to evaluate the performance of the system in field growth environments, and (3) to study the most significant methods for retrieving the biomass estimate from the laser measurements.

Materials and methods

Design and construction of the device

The device combined laser, ultrasonic, thermal infra-red (IR), and ambient moisture measurements from the following commercially available sensors: IL-600 (Keyence Corporation, Itasca, IL, USA), ToughSonic14 (Senix Corporation, Hinesburg, VT, USA), SSS-LT (Process Sensors Corporation, Milford, MA, USA), and DHT22 (Adafruit Industries, New York, NY, USA) respectively[1]. The idea was to position these sensors above the crop at a determined height in nadir view and move them in a circular path parallel to the ground plane. Because of this, and to maintain the general setup of a hand-held device, a tripod was selected as the main frame for the

---

[1] Mention of a trade name, proprietary product, or company name is for presentation clarity and does not imply endorsement by the authors or McGill University, nor does it imply exclusion of other products that may also be suitable.

entire system (AX620B100 62-Inch Proline, Dolica, Rancho Cucamonga, CA, USA).

A stepper motor (T-NM17C04, Zaber Technologies, Vancouver, BC, Canada) was located below the top of the tripod. The sensors were mounted in a 3D-printed holder connected from an edge to the shaft of the controlled stepper motor. The connector featured a fixed angle of 20° with respect to the vertical axis, as shown in Figure 1. In this configuration, the sensors were arranged to spin in a circular path with a specific radius in the range from 60 mm to 150 mm. The motor itself had another 3D-printed holder that was attached to the tripod. Both holders were designed in Inventor 2017 (Autodesk Inc., San Rafael, CA, USA) and their strength was validated with a Finite Element Analysis simulation under the expected loads. The testing field laps were taken with alternating clockwise and counter clockwise directions to avoid stress on the wiring.

Fig. 1. Picture illustrating the 3D-printed sensor holder

The included laser, ultrasonic, and thermal IR sensors had analog output, received by two Analog-to-Digital Converters (ADC) ADS1115 (Adafruit Industries, New York, NY, USA). These ADCs offer each two differential 16-bit channels that communicate with a Raspberry Pi 2 Model B (Raspberry Pi Foundation, Cambridge, UK) via I2C protocol. A Python script was used to read and log the sensors' measurements along with the corresponding angular position and send commands to the motor. The moisture sensor had digital output that was read directly by the Raspberry Pi. This sensor was not mounted in the spinning sensor holder, because only one measurement was desired per location, contrary to the other three, where a profile of measurements was required.

A Bluetooth serial communication protocol was enabled to allow the user to send commands to the Raspberry Pi with a smartphone, as well as to receive information from it. Any of the free serial monitor apps available for smartphones can be paired with the Raspberry Pi. In particular, Serial Bluetooth Terminal v1.12 (Kai Morich, Hockenheim, Germany) was used on an Android 6.0.1 smartphone during the production field experiment. On the Raspberry Pi side, a HC-05 Bluetooth module (Dilson Enterprises, Maharashtra, India) was added for this purpose. More recent versions of the Raspberry Pi come with a built-in Bluetooth transceiver. The block diagram is shown in Figure 2. The sensors and their accompanying circuitry, as well as the motor, were powered by a 6-cell LiPo battery with voltage converters. A box was adapted to contain the circuitry. Figure 3 shows the complete arrangement.
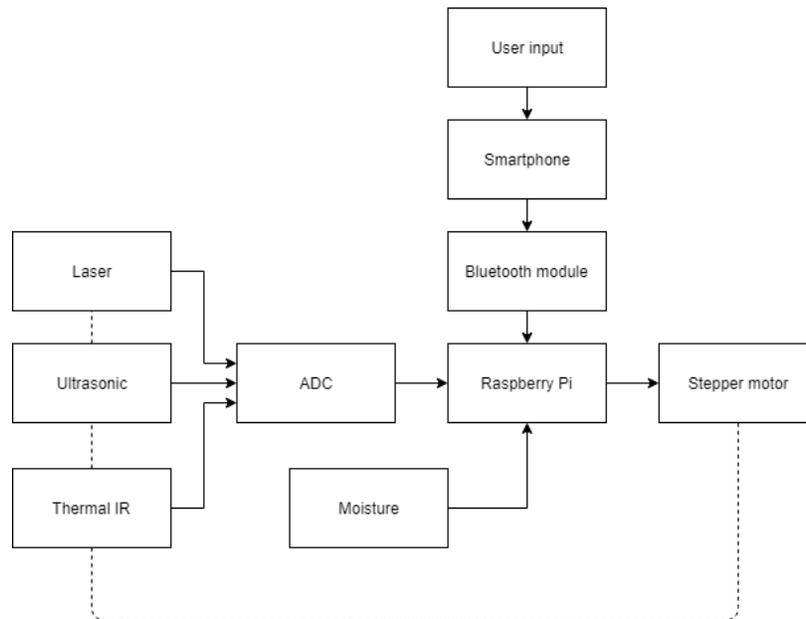
Fig. 2. Block diagram of the developed device

Fig. 3. Picture illustrating built device in the field

Data collection

Test data were collected in a field of VegPro International near Sherrington, Quebec, Canada on October 4th, 10th, and 11th, 2017. The field featured raised beds in organic soil with Stanton spinach (*Spinacia oleracea Hyb. Stanton*) at different growth dates. The presence of several growth dates allowed for the data collection over a short time span. A 500 mm x 500 mm frame was used to denote each zone for sampling. A total of 30 of these zones were measured manually and with the device over the three days. The manual measurements consisted of cutting all the plants inside the square frame and weighing them using an electronic scale MXX-612 (Denver Instrument Inc., Bohemia, NY, USA). This provided the fresh biomass. By dividing by the area of the square frame, the fresh biomass yield is obtained.

The laser, ultrasonic, and thermal IR sensors included in the device provided a measurement every 0.8°, creating a profile of 450 points per lap, along a circumference with a radius of 150 mm. The motor was set with a constant speed of 20°/s. For each location, about 6 laps were recorded, around 3 different centers inside the frame. The moisture sensor provided one measurement for each location.

Analysis methods

The estimation of biomass M is usually based on canopy characteristics, e.g. plant height H, as presented in Equation (1). A regression, either linear or nonlinear, can be used to test how well a certain type of function fits the data. Some error ε is always present, but by testing several functions from a set of candidate functions, the best representative can be found by selecting the one which minimizes the error.

$$M = f_1(H) + \varepsilon_1 \tag{1}$$

The idea behind the device's design is to define a relationship from several height values, i.e. the height profile $\boldsymbol{h}$, and improve the ability to accurately predict biomass. This is expressed in the following equation.

$$M = f_2\big(\boldsymbol{h}^{laser}, \boldsymbol{h}^{ultrasonic}\big) + \varepsilon_2 \tag{2}$$

Furthermore, it may be the case that the biomass model described in Equation (2) can be improved by including other sensor measurements x, like the thermal IR or moisture, as stated by the following equation.

$$M = f_3\big(\boldsymbol{h}^{laser}, \boldsymbol{h}^{ultrasonic}, x_1, \dots, x_n\big) + \varepsilon_3 \tag{3}$$

The goal is then to find some $f_2 \colon \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}$ or $f_3 \colon \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \times \dots \times \mathbb{R} \mapsto \mathbb{R}$ such that $\sigma_1^2 > \sigma_2^2 > \sigma_3^2$, under the assumption that the errors are normally distributed $\varepsilon_1 \sim N(0, \sigma_1^2)$, $\varepsilon_2 \sim N(0, \sigma_2^2)$, and $\varepsilon_3 \sim N(0, \sigma_3^2)$; where m is the number of points recorded in a lap by either laser or ultrasonic sensors. One disadvantage of this approach is that the existence of a physical interpretation for the chosen f cannot be guaranteed.

The data were imported into MATLAB R2017a (MathWorks Inc., Natick, MA, USA). A pre-treatment was performed on the height profiles. First, points considered outliers were removed. Second, the replicates of the same location were averaged to produce a more significant profile. All the measurements taken around the same center point produced one height profile.

Two approaches were taken to analyze the data from the production field experiment. In the first approach (regression of profile-representative features), the functions were built based on the features listed in Table 1, which resulted in scalar values that could be transformed by linear, exponential or polynomial transformations, leaving certain constant parameters to be determined later by regression, typically non-linear. The last two features are analogous to the processing methods by Su (2017). Cases where only one of the features was used at a time were considered, as well as combinations of several of them to evaluate what produced the best fitting.

Table 1. Definition of features used in first analysis approach

Short name or abbreviation	Brief description
Integral	Compute numeric definite integral
Average recorded crop height	Compute arithmetic mean
Max	Find the maximum value
Energy	Compute numeric definite integral of the square
Variance	Compute sample variance
Average trimmed crop height	Compute arithmetic mean only in-between 5 to 95-percentiles
Trimmed variance	Compute sample variance only in-between 5 to 95-percentiles
Energy wavelet	Compute Energy after applying a wavelet filter
Lower envelope	Find lower envelope and compute arithmetic mean
Lower envelope integral	Find lower envelope and compute Integral
Lower envelope energy	Find lower envelope and compute Energy
Upper envelope	Find upper envelope and compute arithmetic mean
Upper envelope integral	Find upper envelope and compute Integral
Upper envelope energy	Find upper envelope and compute Energy
Derivative	Estimate derivative and compute arithmetic mean
Derivative variance	Estimate derivative and compute sample variance
Derivative energy	Estimate derivative and compute Energy
Count	Find the number of points where abrupt changes happen
Peak count	Find the number of local maxima
Frequency	Estimate the median normalized frequency of the power spectrum
Bandwidth	Estimate bandwidth of the power spectrum
SFDR	Estimate ratio between fundamental frequency and first spurious peak in power spectrum
SNR	Estimate Signal-to-Noise Ratio
THD	Compute Total Harmonic Distortion
Mean of max	Find the maxima between sets of 10 points, then average them
Pseudo-max	Find the value of 95-percentile

The second approach (inference of a canopy density function) relied on the formulation indicated in Equation (4), as follows. A cylindrical coordinate system was used because it relates directly with the circular paths followed by the device. In this case, the z axis was perpendicular to the ground plane and aligned with the center point of the circular path, while the angular position θ corresponds to the placement of the sensors by the stepper motor and the radius r to the distance between any point in the ground plane to the center of the circular path. The key assumptions were that the height profile measured along the circular path was representative of the entire frame and that the density of the crop only varies on a noticeable scale with the height. About the former, one way that the assumption could be included was by considering concentric circles where the height profile was repeated, so that the height of the ith plant (labeled H_i) depended only on the angular position, and not on the radius.

$$M_i = \int \rho \, dV = \int_0^{R_l} \int_0^{2\pi} \int_0^{H_i(\theta,r)} r \, \rho(z,\theta,r) \, dz \, d\theta \, dr = \int_0^{R_l} \int_0^{2\pi} \int_0^{H_i(\theta)} r \, \rho(z) \, dz \, d\theta \, dr + \varepsilon_i$$

$$= \int_0^{R_l} r \, dr * \int_0^{2\pi} \int_0^{H_i(\theta)} \rho(z) \, dz \, d\theta + \varepsilon_i = \frac{R_l^2}{2} * \int_0^{2\pi} \int_0^{H_i(\theta)} \rho(z) \, dz \, d\theta + \varepsilon_i \qquad (4)$$

where M_i denotes the mass of the ith plant, R_l is the maximum radius used for the lap measurements, ρ is the density, and ε_i the error produced from the assumptions in the ith plant. It is worth noting that Equation (4) can be rewritten as Equation (5) based on the Fundamental theorem of calculus.

$$M_i = \frac{R_l{}^2}{2} * \int\limits_0^{2\pi} \int\limits_0^{H_i(\theta)} \rho(z)\,dz\,d\theta + \varepsilon_i = \frac{R_l{}^2}{2} * \int\limits_0^{2\pi} f\big(H_i(\theta)\big)\,d\theta + \varepsilon_i \tag{5}$$

for some function $f\colon \mathbb{R} \mapsto \mathbb{R}$ such that $f'(x) = \rho(x)$. This step allows for a single integration instead of a double integration. Different candidates for this function f were tested. To compare with the manually measured biomass, which was sampled within the square frame, the biomass yield can be found by dividing over the area, which for the case of each measurement was that of the circle with the largest radius. Thus, Equation (5) becomes Equation (6).

$$D_i = \frac{M_i}{A} = \frac{\frac{R_l{}^2}{2} * \int_0^{2\pi} f\big(H_i(\theta)\big)\,d\theta + \varepsilon_i}{\pi R_l^2} = \frac{1}{2\pi} \int\limits_0^{2\pi} f\big(H_i(\theta)\big)\,d\theta + \varepsilon_i^{(6)} \tag{6}$$

Finally, it is worth mentioning that this approach can be analogous to the first approach taken, depending on the selection of the function f. For example, assuming $\rho(z) = \rho \to D_i = \frac{\rho}{2\pi} \int_0^{2\pi} H_i(\theta)\,d\theta = a * Integral + b$, where *Integral* refers to one of the processing methods from the first approach, which was used in a linear regression model to relate to D_i, with some parameters a and b, taking $a = \frac{\rho}{2\pi}$ and $b = 0$. Another example, taking $f(x) = x^2 \to D_i = \frac{1}{2\pi} \int_0^{2\pi} [H_i(\theta)]^2\,d\theta = a * Energy + b$, where the same situation is found to occur with a different processing method. The advantage of this second approach was that it guarantees the existence of a physical interpretation for the model once an appropriate function f has been selected.

For both approaches, the coefficient of determination (R^2) and root mean squared error (RMSE) were calculated as indicators of the system performance related to the ability to predict fresh biomass and used to compare methods. For cases where the number of parameters varied, the adjusted R^2 was used as the decision criteria. Finally, when a working model was achieved, the mean absolute percentage error (MAPE) was computed to indicate the precision of the system using the following equation. This was not used as a criterion to choose between different models.

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{actual_i - estimate_i}{actual_i} \right| \tag{7}$$

Results and discussion

Figure 4 presents the fresh biomass yield manually measured within the square frame for all dates. This yield is simply the biomass divided by the area of the square frame (0.25 m²) and transformed to t/ha units. This data has a mean value of 10.77 t/ha and a standard deviation of 7.93 t/ha. The difference between the maximum and minimum values is 32.9 t/ha, proving that a varied range of biomass yield has been recorded. It seems as if it is not normally distributed, but rather similar to a log-normal. Whatever the case, the exact distribution is not relevant for the analysis. A drawback of the experiment is the absence of samples in the range of 20 t/ha to 30 t/ha.
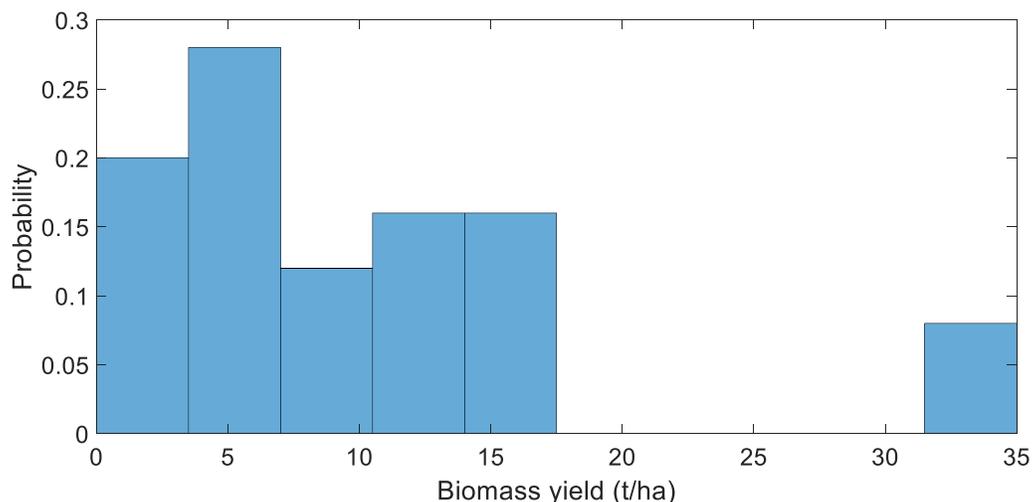
Fig. 4. Normalized histogram of fresh biomass yield

Taking the first approach with the laser measurements, Table 2 was built with the features described in Table 1. The RMSE values are in t/ha. The best performing method was the Lower envelope integral with linear regression. Figure 5 illustrates the best performing regression at this stage.

Table 2. Summary of results from regressions by processing method

| Name in MATLAB	Linear			Quadratic			Exponential		
	R^2	R^2adj	RMSE [t/ha]	R^2	R^2adj	RMSE [t/ha]	R^2	R^2adj	RMSE [t/ha]
Integral	0.54	0.52	5.94	0.56	0.52	5.96	0.56	0.54	5.76
Average recorded crop height	0.53	0.51	5.96	0.55	0.51	5.98	0.56	0.54	5.78
Max	0.01	<0	8.67	0.02	<0	8.82	0.01	<0	8.66
Energy	0.51	0.49	6.10	0.52	0.48	6.19	0.52	0.50	6.07
Variance	0.01	<0	8.69	0.01	<0	8.88	0.01	<0	8.69
Average trimmed crop height	0.53	0.51	5.99	0.55	0.51	5.98	0.56	0.54	5.76
Trimmed variance	0.00	<0	8.72	0.02	<0	8.82	0.00	<0	8.72
Energy wavelet	0.54	0.52	5.89	0.54	0.50	6.02	0.53	0.51	5.99
Lower envelope	0.62	0.60	5.39	0.62	0.59	5.48	0.61	0.59	5.45
Lower envelope integral	0.62	0.60	5.38	0.62	0.59	5.48	0.61	0.59	5.45
Lower envelope energy	0.61	0.59	5.46	0.61	0.57	5.56	0.59	0.57	5.58
Upper envelope	0.44	0.42	6.54	0.46	0.41	6.55	0.48	0.46	6.31
Upper envelope integral	0.44	0.42	6.54	0.46	0.41	6.56	0.48	0.46	6.32
Upper envelope energy	0.12	0.08	8.21	0.13	0.05	8.34	0.10	0.06	8.26
Derivative	0.00	<0	8.73	0.00	<0	8.92	0.00	<0	8.73
Derivative variance	0.00	<0	8.72	0.15	0.07	8.25	0.00	<0	8.73
Derivative energy	0.00	<0	8.73	0.14	0.06	8.29	0.00	<0	8.73
Count	0.02	<0	8.64	0.24	0.17	7.76	0.01	<0	8.67
Peak count	0.00	<0	8.72	0.03	<0	8.77	0.00	<0	8.72
Frequency	0.13	0.09	8.15	0.18	0.11	8.09	0.18	0.14	7.92
Bandwidth	0.15	0.11	8.04	0.15	0.07	8.22	0.14	0.10	8.08
SFDR	0.27	0.24	7.44	0.32	0.26	7.36	0.33	0.30	7.15
SNR	0.13	0.09	8.14	0.14	0.06	8.30	0.11	0.07	8.22
THD	0.04	0.00	8.54	0.06	<0	8.66	0.05	0.01	8.51
Mean of max	0.44	0.42	6.52	0.46	0.41	6.52	0.48	0.46	6.28
Pseudo-max	0.37	0.34	6.90	0.43	0.38	6.75	0.44	0.42	6.50
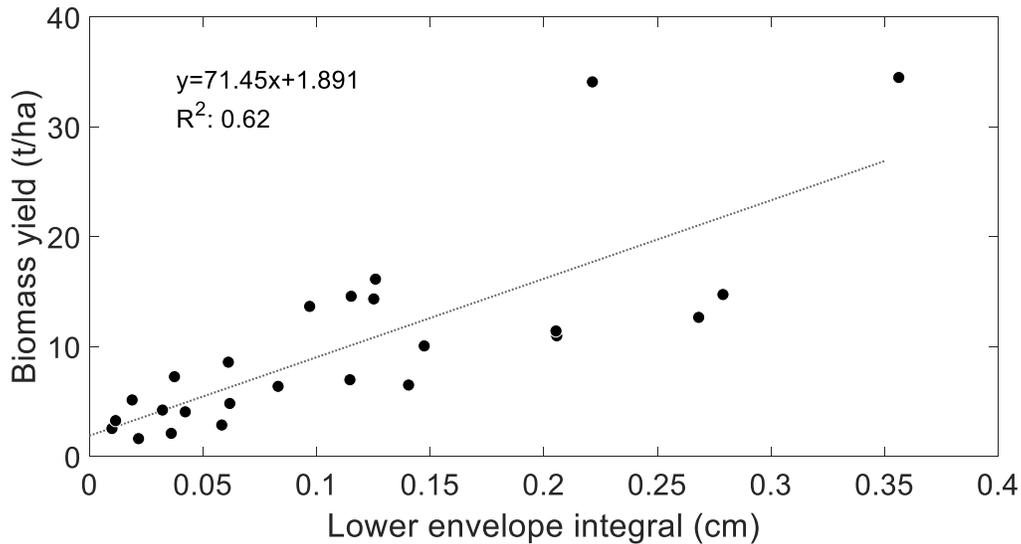
Fig. 5. Linear regression of biomass and lower envelope integral

The resulting RMSE of 5.38 t/ha with that processing method is still relatively high. Because of this, new factors that included combinations of the mentioned processing methods were considered. The combination of processing methods gave rise to an improved fitting, but to avoid overfitting, the number of parameters was limited to a maximum of 6. This constraint in the complexity of the model would prove useful in a validation experiment. Table 3, Figure 6, and the following equations summarize the most reliable solutions.

Table 3. Summary of results from regressions by equation

Equation number	R^2	R^2adj	RMSE [t/ha]	Number of parameters
(12)	0.68	0.65	5.06	3
(13)	0.83	0.81	3.72	4
(14)	0.92	0.90	2.71	5
(15)	0.93	0.91	2.55	6

$$f\left(\boldsymbol{h}_i^{laser}\right) = a + b\frac{[Lower\ envelope\ energy]}{[Upper\ envelope\ energy]}$$
$$+ c[Mean\ of\ max]\frac{[Lower\ envelope\ energy]}{[Upper\ envelope\ energy]} \tag{8}$$

$$f\left(\boldsymbol{h}_i^{laser}\right) = a + b[SFDR] + c[Lower\ envelope\ energy]^2 + d[SFDR]^2 \tag{9}$$

$$f\left(\boldsymbol{h}_i^{laser}\right) = a\frac{[Integral]^2}{[Max]^2} + b\frac{[Average\ recorded\ crop\ height]^2}{[Variance]^2}$$
$$+ c\frac{[Integral]^2[Average\ recorded\ crop\ height]}{[Max]^2[Variance]}$$
$$+ d\frac{[Integral][Average\ recorded\ crop\ height]^2}{[Max][Variance]^2}$$
$$+ e\frac{[Average\ recorded\ crop\ height]^3}{[Variance]^3} \tag{10}$$

$$f\left(h_i^{laser}\right) = a[Lower\ envelope\ energy] + b[Lower\ envelope\ energy]^2$$
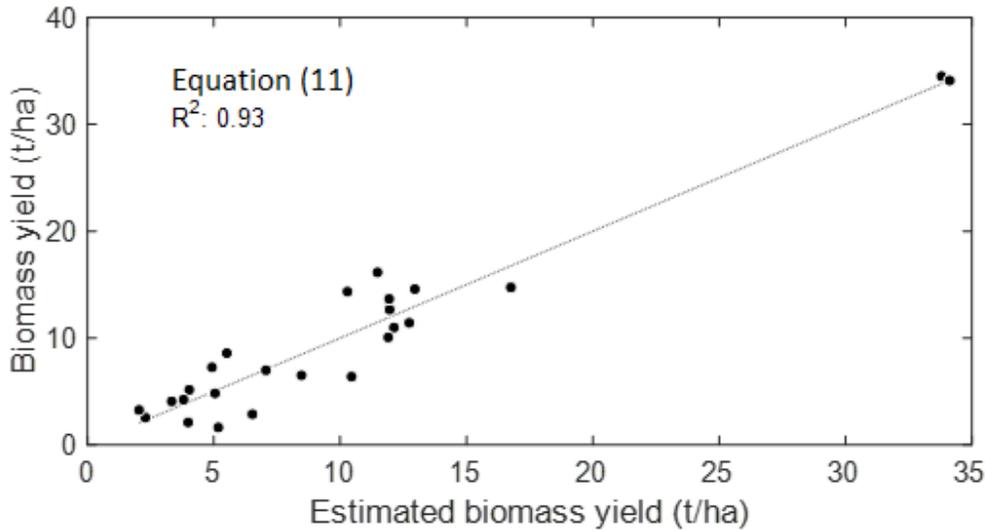$$+c[Lower\ envelope\ energy]^3 + d[SFDR] + e[SFDR]^2 + f[SFDR]^3 \qquad (11)$$

Fig. 6. Best fit achieved with the first approach

The best fitting regression, expressed in Equation (11) and used to build Figure 6, relies on 2 processing methods that are relatively hard to find physical meaning. Nonetheless, an intuition about what SFDR and Lower envelope energy mean can be provided in this context. SFDR is a ratio between the power at the fundamental frequency of a signal and the power at a significant higher frequency. This ratio is then a measure of how important the higher frequency components are, where typically the spurious peak is due to noise. On the other hand, the envelope of a signal provides information about the low frequency components. By considering low frequency and high frequency components separately, the processing method can produce a more reliable estimation. In the context of height profiles, low frequency components refer to large-scale sections of the canopy, rather than the sections of the profile with larger local variability.

With the second approach, inference of a canopy density function, the performance of different density functions is summarized in Table 4. To achieve results comparable to those of the first approach, a relatively large number of parameters is required. For example, the best fitting of a density function uses 9 parameters, as specified in Equation (12). Figures 7 and 8 illustrate this density function, while Figure 9 shows the fitting of the data using this approach. It was expected that the density at lower parts of the canopy was higher, because the stems are heavier than the leaves. The elements of this approach have a clear physical meaning.

Table 4. Summary of results from density functions

Density function	R^2	R^2adj	RMSE [t/ha]	Number of parameters
Constant	0.44	0.44	6.27	1
Linear	0.45	0.43	6.21	2
Quadratic	0.49	0.44	5.98	3
Cubic	0.50	0.43	5.94	4
Exponential	0.44	0.42	6.25	2
Gaussian	0.48	0.43	6.03	3
Sinusoidal	0.50	0.45	5.90	3
Rational	0.52	0.48	5.80	3
Rational	0.54	0.47	5.70	4
Logistic*	0.49	0.44	5.98	3
Generalized logistic*	0.49	0.42	5.97	4
Rational tanh*	0.63	0.58	5.10	4
Modified logistic*	0.74	0.61	4.24	9

*: referring to the cumulative density, of which the density function is the derivative

$$f\big(H_i(\theta)\big) = a * tanh\left(b * \frac{H_i(\theta) + c}{H_i(\theta) + d}\right)\frac{H_i(\theta) - e}{H_i(\theta) + f} + g * sin\big(p(H_i(\theta) + l)\big) \qquad (12)$$
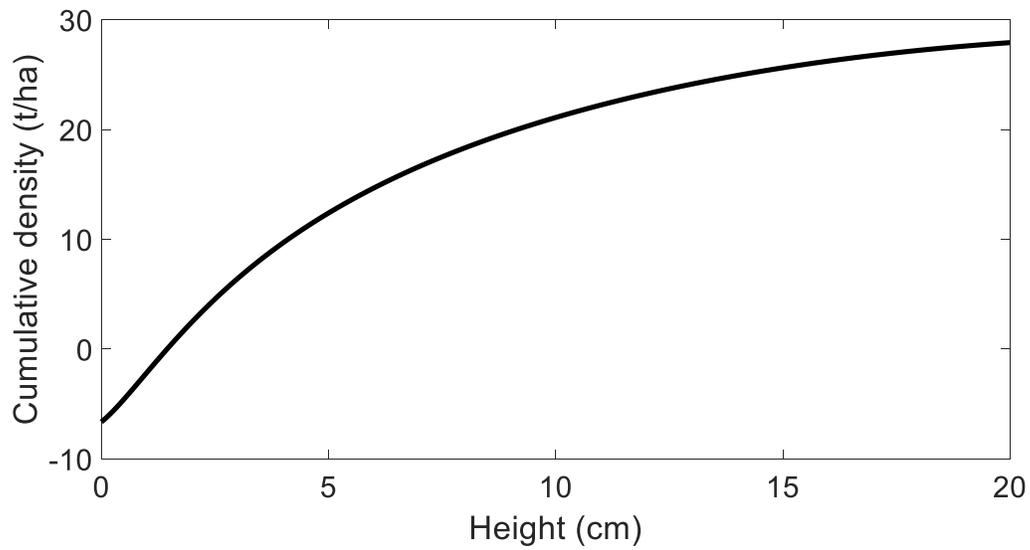

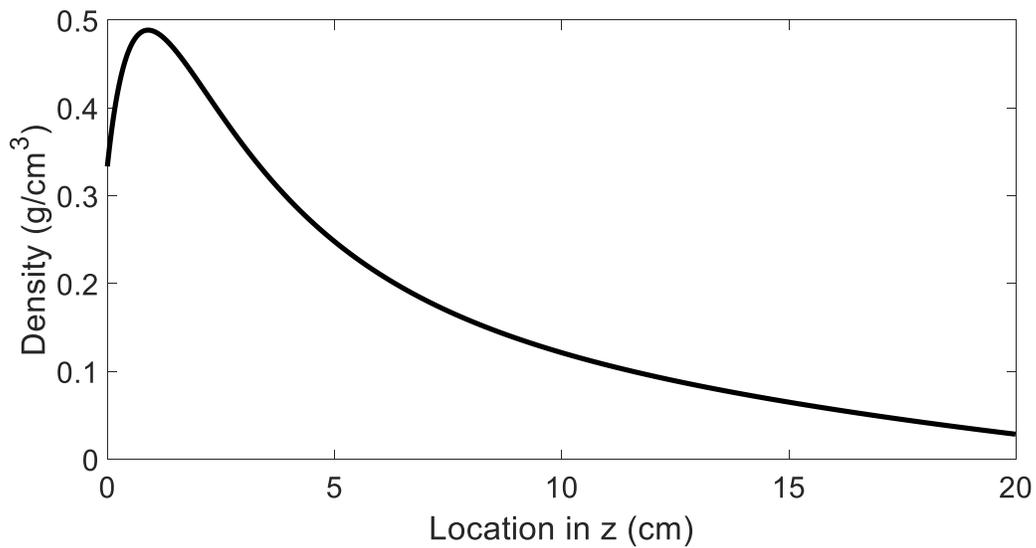**Fig. 7. Best fit for cumulative density function**
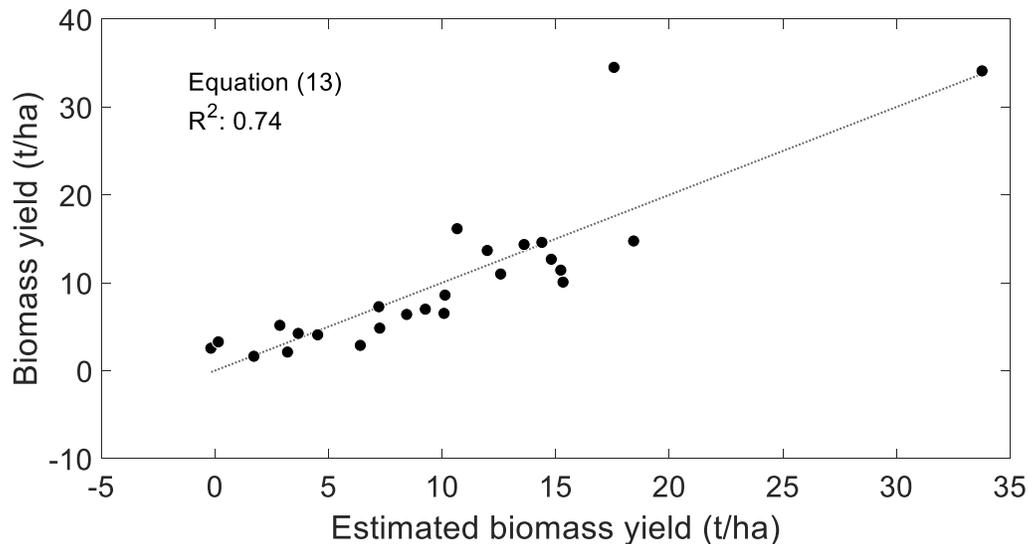

Fig. 8. Best fit for density function

Fig. 9. Best fit achieved with the second approach

At this point, it was considered the effect of the measurements where the biomass yield was above 30 t/ha, which could be regarded as outliers. It is important to consider that with the current results, removing these higher yield points from the computation of RMSE would turn the value of this indicator into 2.70 t/ha. By removing the higher yield points and using the second approach to retry the fitting of the density function, Table 5 was built. For all the considered cases, the fitting improved, resulting in a reduction of the gap among the best and worst density functions. These results seem to show that removing the higher yield points is convenient for simpler models, which could be explained by considering that the whole range of the function is significantly reduced. The case of having a logistic cumulative density function produces the best results in terms of adjusted R^2, and the RMSE is lower than the corresponding best case when the higher yield points were included. Nonetheless, in spite of the convenience of these results, this process proved itself inconclusive in the determination of the outlier status of the higher yield points. Thus, it will be assumed that the higher yield is produced by the natural variability of the crop, and will be retained in the final analysis.

Table 5. Summary of results from density functions without higher yield points

Density function	R^2	R^2adj	RMSE [t/ha]	Number of parameters
Constant	0.65	0.65	2.72	1
Linear	0.67	0.65	2.62	2
Quadratic	0.69	0.66	2.57	3
Cubic	0.69	0.64	2.57	4
Exponential	0.66	0.64	2.65	2
Gaussian	0.69	0.66	2.54	3
Sinusoidal	0.69	0.66	2.55	3
Rational	0.58	0.54	2.97	3
Rational	0.58	0.51	2.97	4
Logistic*	0.70	0.67	2.52	3
Generalized logistic*	0.70	0.65	2.52	4
Rational tanh*	0.71	0.66	2.23	4
Modified logistic*	0.78	0.65	2.14	9

*: referring to the cumulative density, of which the density function is the derivative

Finally, Table 6 compares the best case of both approaches. The percentage error is considerably high for both cases, which is troubling, but is partially explained by the large range covered by the measured biomass. Whatever the case, the coefficient of determination is significant, showing that the laser measurements explain the biomass behavior. There is a possibility that the

uncertainty of the manual biomass measurement is also playing a role in this discrepancy. The moisture sensor information remains a good complement for the laser measurements. The moisture was linearly combined with the current estimations made with the first and second approaches by themselves, as stated by the following equation.

$$[final\ estimate] = a * [first\ or\ second\ approach] + b * [moisture] + c \qquad (13)$$

Table 6. Summary of best results by approach

Approach	R^2	R^2adj	RMSE [t/ha]	MAPE [%]
Regression of profile-representative features	0.93	0.91	2.55	34.08
Inference of a canopy density function	0.74	0.61	4.24	35.82
Regression of profile-representative features + Moisture	0.94	0.92	2.16	31.44
Inference of a canopy density function + Moisture	0.78	0.62	4.18	32.09

Conclusion

Two different approaches were considered, providing R^2 of 0.94 and 0.78, respectively. Though the first approach resulted in significantly better fitting, there might have been some overfitting, even when the complexity of the model was constrained. The second approach had the advantage of having a distinct physical meaning and could be considered more conservative. The large values of RMSE and MAPE could be due to the high variability intrinsic to the biomass measurement, but still raised concerns about the repeatability of the measurements prepared with the device. This issue could be addressed by taking repeated measurements inside the same frame.

For both approaches, the results were comparable to those of the reviewed literature, showing that the performance was not diminished by the use of the developed low-cost prototype in an open field environment. These results proved that the developed device is a viable tool for measuring biomass in an easy and fast way. This would result in faster and more frequent sampling becoming available for users. There were strong indicators that the device is not only more convenient than the traditional manual measurement of plant height, i.e. with ruler or measuring tape, but also more accurate in the estimation of biomass.

Acknowledgements

References

Biskup, B., Scharr, H., Schurr, U., & Rascher, U. (2007). A stereo imaging system for measuring structural parameters of plant canopies. Plant, Cell and Environment.

Buelvas, R., & Adamchuk, V. (2017). Crop canopy measurement using laser and ultrasonic sensing integration. 2017 ASABE Annual International Meeting. Spokane, WA.

Buelvas, R., & Adamchuk, V. (2018). Crop canopy measurements using a low-cost laser for biomass estimation. McGill University. MSc thesis.

Catchpol, W. R., & Wheeler, C. J. (1992). Estimating plant biomass: A review of techniques. Australian Journal of Ecology.

Ehlert, D., Horn, H.-J., & Adamek, R. (2008). Measuring crop biomass yield by laser triangulation. Computers and electronics in agriculture.

Eitel, J. U., Magney, T. S., Vierling, L. A., Brown, T. T., & Huggins, D. R. (2014). LiDAR based biomass and

crop nitrogen estimates for rapid,non-destructive assessment of wheat nitrogen status. Field Crops Research.

Freeman, K. W., Girma, K., Arnall, D. B., Mullen, R. W., Martin, K. L., Teal, R. K., & Raun, W. R. (2007). By-Plant Prediction of Corn Forage Biomass and Nitrogen Uptake at Various Growth Stages Using Remote Sensing and Plant Height. Agronomy Journal.

Golzarian, M. R., Frick, R. A., Rajendran, K., Berger, B., Roy, S., Tester, M., & Lun, D. S. (2011). Accurate inference of shoot biomass from high-throughput images of cereal plants. Plant Methods.

Keightleya, K. E., & Bawdenb, G. W. (2010). 3D volumetric modeling of grapevine biomass using Tripod LiDAR. Computers and electronics in agriculture.

Kjaer, K. H., & Ottosen, C. O. (2015). 3D Laser Triangulation for Plant Phenotyping in Challenging Environments . Sensors.

Moorthya, I., Miller, J. R., JimenezBerni, J. A., Zarco-Tejada, P., Hu, B., & Chen, J. (2010). Field characterization of olive (Olea europaea L.) tree crown architecture using terrestrial laser scanning data. Agricultural and Forest Meteorology.

Reusch, S. (2009). Use of ultrasonic transducers for on-line biomass estimation in winter wheat. 7th European Conference on Precision Agriculture (pp. 169-176). Wageningen: Wageningen academic publishers.

Rosell, J. R., Sanz, R., Llorens, J., Arnó, J., Escolà, A., Ribes-Dasi, M., . . . Palacín, J. (2009). A tractor-mounted scanning LIDAR for the non-destructive measurement of vegetative volume and surface area of tree-row plantations. Biosystems Engineering.

Schaefer, M. T., & Lamb, D. W. (2016). A Combination of Plant NDVI and LiDAR Measurements Improve the Estimation of Pasture Biomass. Remote Sensing.

Su, Y. (2017). Development of Proximal Sensing Systems for Crop Biomass Determination. McGill University.

Tilly, N., Aasen, H., & Bareth, G. (2015). Fusion of Plant Height and Vegetation Indices for the estimation of barley biomass. Remote Sensing.

Tumbo, S. D., Salyani, M., Whitney, J. D., Wheaton, T. A., & Miller, W. M. (2001). Investigation of laser and ultrasonic ranging sensor for measurements of citrus canopy volume. Applied Engineering in Agriculture.

Van Henten, E. J. (1994). Validation of a Dynamic Lettuce Growth Model for Greenhouse Climate Control. Agricultural Systems.