

CROP CANOPY MEASUREMENTS USING A LOW-COST LASER
FOR BIOMASS ESTIMATION

By

Roberto Mario Buelvas Gómez

Department of Bioresource Engineering
Macdonald Campus of McGill University
Montreal, Canada
April, 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the
degree of
Master of Science

© Roberto Buelvas, 2018

ABSTRACT

In-season sensing of crop architectonics is important for identifying and preventing potential stresses as well as optimizing crop management logistics. Measurements of chlorophyll content, size, density, and/or temperature of the canopy have been used as primary *in situ* diagnostic tools. The goal of this study was to develop a prototype sensor system that would integrate laser proximity measurements as the basis for a low-cost instrumented system for green vegetable production. The system involves circular scanning of crop canopies to identify fresh biomass under different soil and management conditions.

The first experiment was conducted in a greenhouse with lettuce and kale. Biomass was estimated from the sensor system's measurements resulting in R^2 values between 0.74 and 0.93, percentage error between 25% and 55%, and root mean squared error (RMSE) between 0.295 ln(g) and 0.441 ln(g). These values include both dry and fresh biomass for lettuce and kale. The second experiment in a spinach field on a commercial farm produced similar results. Two approaches for processing the laser-based height profiles are discussed: regression of profile-representative features and inference of a canopy density function. Depending on the processing method, the R^2 was either 0.78 or 0.94, and the RMSE was 4.18 t/ha and 2.16 t/ha for each case. The mean absolute percentage error (MAPE) was around 30% for both methods. The second experiment featured improved hardware in terms of ease of operation.

Proper use of this technology will allow farm managers to improve harvest plans and shipment schedules under variable plant growth dynamics.

RÉSUMÉ

La détection saisonnière de l'architecture des plantes cultivées est importante pour discerner et prévenir les contraintes potentielles ainsi que pour optimiser la logistique de la gestion des cultures. Des mesures de la teneur en chlorophylle, de la taille, de la densité ou de la température de la canopée ont été utilisées comme principaux outils de diagnostic in situ. La présente étude avait pour objectif de faire le point sur le développement d'un prototype de système de capteurs intégrant les mesures de proximité au laser comme base d'un système muni d'équipement à faible coût pour la production de légumes verts. Le système nécessite le balayage circulaire des canopées pour identifier la biomasse fraîche de sols différents et de conditions de gestion variées.

La première expérience a été réalisée dans une serre avec de la laitue et du chou frisé. Dans cette étude, la biomasse a été estimée à partir des mesures du système de capteurs. Les valeurs R^2 obtenues étaient entre 0.74 et 0.93, le pourcentage d'erreur oscillait entre 25% et 55%, et l'erreur quadratique moyenne (RMSE) était entre $0.295 \ln(g)$ et $0.441 \ln(g)$. Ces valeurs incluent la biomasse fraîche et sèche pour la laitue et le chou frisé. Ensuite, la deuxième étude a été menée, cette fois-ci dans un champ d'épinards d'une ferme commerciale. Les résultats obtenus ont été similaires à ceux de l'étude précédente. Deux approches différentes pour traiter les profils de hauteur basés sur le laser sont discutées: la régression des caractéristiques représentatives du profil et l'inférence d'une fonction de densité de la canopée. Les mesures de R^2 obtenues étaient de 0.78 ou de 0.94, selon la méthode de traitement utilisée, et le RMSE était de 4.18 t/ha, et de 2.16 t/ha. L'erreur moyenne absolue en pourcentage (MAPE) était d'environ 30% pour les deux méthodes. Cette deuxième expérience faisait également appel à un matériel amélioré.

L'utilisation appropriée de cette technologie permettra aux gestionnaires des fermes d'améliorer leurs plans de récolte et les calendriers d'expédition dans des conditions de croissance variables.

ACKNOWLEDGEMENTS

This research was supported in part by funds provided through the National Science and Engineering Research of Canada (NSERC) Discovery Grant and Programa Crédito-Beca from Colfuturo.

Thanks to Peter Tikasz and Dr. Mark Lefsrud for providing the greenhouse setup for the controlled environment experiment. Thanks to Jarek Holoszkiewicz and the team at VegPro International, Inc. (Sherrington, QC, Canada) for providing an opportunity to field test the developed system.

I would like to thank my colleagues and lab mates who supported me through the whole process and helped me with the data collection.

I would also like to express my most grateful appreciation to my parents, my brother, and my entire family, as well as my girlfriend for their unconditional support. Finally, I would like to thank Dr. Viacheslav Adamchuk, my supervisor, for his patient guidance and constant motivation.

CONTRIBUTIONS OF AUTHORS

The research work of this thesis was presented at an early stage in the 2017 ASABE Annual International Meeting held in Spokane, Washington, USA (Paper No. 1701002). The author of this thesis was responsible for building the prototype, building the code to collect and process the data, designing and carrying out the experiments, and preparing the manuscript based on this thesis. Dr. Viacheslav Adamchuk, an Associate Professor in the Department of Bioresource Engineering of McGill University, is the thesis supervisor. He created the idea for this research and offered scientific advice and technical guidance throughout the study. He is also responsible for editing and reviewing the prepared manuscript.

Publication related to the thesis

Buelvas, R., & Adamchuk, V. 2017. Crop canopy measurement using laser and ultrasonic sensing integration. 2017 ASABE Annual International Meeting Paper No. 1701002. ASABE: St. Joseph, Michigan, USA.

TABLE OF CONTENTS

ABSTRACT.....	2
RÉSUMÉ.....	3
ACKNOWLEDGEMENTS.....	4
CONTRIBUTIONS OF AUTHORS.....	5
TABLE OF CONTENTS.....	6
LIST OF FIGURES.....	7
LIST OF TABLES.....	9
LIST OF ABBREVIATIONS AND SYMBOLS.....	10
1. INTRODUCTION.....	11
1.1. OBJECTIVES.....	11
2. LITERATURE REVIEW.....	12
3. MATERIAL AND METHODS.....	16
3.1. DESIGN AND CONSTRUCTION OF THE DEVICE.....	16
3.2. CONTROLLED ENVIRONMENT EXPERIMENT DESIGN.....	18
3.3. CONTROLLED ENVIRONMENT ANALYSIS.....	21
3.4. MODIFICATIONS TO THE DEVICE.....	23
3.5. PRODUCTION FIELD EXPERIMENT DESIGN.....	26
3.6. PRODUCTION FIELD ANALYSIS.....	27
4. RESULTS AND DISCUSSION.....	31
4.1. CONTROLLED ENVIRONMENT EXPERIMENT.....	31
4.2. PRODUCTION FIELD EXPERIMENT.....	43
5. SUMMARY AND CONCLUSION.....	52
6. REFERENCES.....	53
APPENDICES.....	57

LIST OF FIGURES

Fig. 1. Block diagram of the system.....	17
Fig. 2. Picture illustrating the complete device in the greenhouse setup.....	17
Fig. 3. Illustration of the Rockwool cubes' distribution in the hydroponic bed..	18
Fig. 4. Picture illustrating experimental setup in greenhouse.....	20
Fig. 5. Picture illustrating the new holder.....	24
Fig. 6. Block diagram of modified system.....	25
Fig. 7. Picture illustrating setup for the production field experiment.....	26
Fig. 8. Evolution of average lettuce dry biomass against bed.....	31
Fig. 9. Evolution of average kale dry biomass against bed.....	31
Fig. 10. Exponential regression of lettuce fresh biomass vs maximum height...	32
Fig. 11. Exponential regression of kale fresh biomass vs maximum height.....	33
Fig. 12. Picture illustrating circular path over the crops at middle location for (a) DAS 36 and (b) DAS 43.....	34
Fig. 13. Example of height profile as recorded by ultrasonic sensor and calculated processing methods.....	34
Fig. 14. Best model of exponential regression for (a) lettuce and (b) kale fresh biomass based on ultrasonic.....	35
Fig. 15. Best model of exponential regression for (a) lettuce and (b) kale dry biomass based on ultrasonic.....	36
Fig. 16. Example of height profiles as recorded by laser and ultrasonic sensors for different days of the experiment.....	36
Fig. 17. Best model of exponential regression for (a) lettuce and (b) kale fresh biomass based on laser.....	37
Fig. 18. Best model of exponential regression for (a) lettuce and (b) kale dry biomass based on laser.....	38
Fig. 19. Actual fresh biomass against fresh biomass predicted by the device for lettuce	42
Fig. 20. Actual fresh biomass against fresh biomass predicted by the device for kale	42
Fig. 21. Normalized histogram of fresh biomass yield.....	43
Fig. 22. Linear regression of biomass and lower envelope integral.....	45

Fig. 23. Best fit achieved with the first approach.....	46
Fig. 24. Best fit for cumulative density function.....	48
Fig. 25. Best fit for density function.....	48
Fig. 26. Best fit achieved with the second approach.....	49

LIST OF TABLES

Table 1. Thresholds used to simulate water stress.....	18
Table 2. Schedule for data collection and plant harvesting.....	19
Table 3. Definition of processing methods for production field experiment	23
Table 4. Definition of processing methods for production field experiment.....	28
Table 5. Coefficients of determination from exponential regression between the biomass and processed ultrasonic height measurements for each processing method.....	35
Table 6. Coefficients of determination from exponential regression between the biomass and processed laser height measurements for each processing method.....	37
Table 7. Processing methods found to produce the best results in exponential regression	38
Table 8. Comparison of fitting in exponential regression between laser, ultrasonic and both.....	39
Table 9. Pearson correlation values between plant biomass and other measured properties.....	39
Table 10. Coefficients of determination from exponential regression between the biomass and combination of laser and ultrasonic height measurements with other measured properties.....	40
Table 11. Summary of results from exponential regressions.....	41
Table 12. Summary of results from regressions by processing method.....	44
Table 13. Summary of results from regressions by equation.....	45
Table 14. Summary of results from density functions.....	48
Table 15. Summary of best results by approach.....	49

LIST OF ABBREVIATIONS AND SYMBOLS

ADC	Analog-to-Digital Converter
D	Biomass yield
DAS	Days After Seeding
H	Height as scalar or scalar function
h	Height profile as vector
IR	Infra-Red
LAI	Leaf Area Index
LiDAR	Light Detection And Ranging
LiPo	Lithium-Polymer
M	Fresh biomass
MAPE	Mean Absolute Percentage Error
NDRE	Normalized Difference Red Edge
NDVI	Normalized Difference Vegetation Index
NIR	Near Infra-Red
PA	Precision Agriculture
PE	Percentage Error
RMSE	Root Mean Squared Error
SFDR	Spurious-Free Dynamic Range
SNR	Signal-to-Noise Ratio
THD	Total Harmonic Distortion
UAV	Unmanned Aerial Vehicle
YM	Yield Monitor
ε	Error term

1. INTRODUCTION

Crop biomass is used as an indicator of plant growth in plant phenotyping and as a way to estimate 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 have 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.

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 products is seldom found in the literature.

1.1. OBJECTIVES

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, and reliable assessment of their crops. Specific objectives were: (1) to develop the system prototype, (2) to evaluate the performance of the system in controlled and in field growth environments, and (3) to study the most significant methods for retrieving the biomass estimate from the laser measurements.

2. LITERATURE REVIEW

“Estimation of plant biomass is a central part of many ecological investigations. For instance, plant biomass is used to characterize biomes and ecosystems, to measure productivity, grazing pressure, and many other processes” (Jonasson, 1988). Traditionally, farmers have relied on manual measurements and visual estimations to keep track of the biomass changes across their fields. Although experienced farmers can conduct this estimation with relative accuracy, the advantages of having sensor systems performing these operations are evident. With sensor systems, the subjectivity of the estimation is removed, allowing for improved repeatability, as well as a reduction in labour-intensive and time-consuming sampling activities (Fricke, Richter, & Wachendorf, 2011).

One solution comes in the form of Yield Monitors (YM), which measure the production of the crop as it is harvested (Borgelt, 1993; Vansichen and DeBaerdemaeker, 1993; Plant, 2001; Fulton et al., 2009). Similar approaches have been extended to measurements to be taken before harvesting, but they all share the common factor of removing the whole plant or part of it per measurement, limiting the maximum number of measurements that can be taken in a field.

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 (Freeman et al., 2007; Eitel et al., 2014; Biskup et al., 2007; Rosell et al., 2009; Moorthya et al., 2010). 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.

Reusch (2009) used an ultrasonic sensor to estimate the dry mass density for four varieties of wheat. The sensor measured multiple echoes as they were reflected from different layers in the canopy while mounted on a vehicle. The collected information was used to determine which wheat variety was more convenient in terms of the variety which produced

the highest yield. Reusch (2009) found that the relationship between dry mass density and height changed with growth stage. He then proposed the separation of this behavior into three groups to fit linear models for each of the growth stages, as a way of avoiding nonlinear models.

Ehlert et al. (2008) took a similar approach by using a modified laser rangefinder mounted on a vehicle. The laser rangefinder worked using the principle of laser triangulation, producing a height profile across the path of the vehicle. The mean of the profile was taken as related to the biomass yield using linear and quadratic regressions. Fresh and dry biomass yield was estimated for oilseed rape and winter rye, while comparing the effect of the angle of incidence of the laser.

In an earlier work, Tumbo et al. (2001) made a comparison between laser and ultrasonic technologies for the estimation of canopy volume in citrus trees. Both types of sensors were mounted on moving platforms. The volume was computed as a numeric integration of the sensors' measurements for the ultrasonic technology. They found that the laser provided slightly better results, especially in defoliated trees.

Lasers with 3D capabilities were mounted on a tripod (Keightleya and Bawdenb, 2010; Eitel et al., 2014), rather than on a moving vehicle as was explained in the previously mentioned works. The former focused on retrieving a measurement of canopy volume for grapevine, while the latter estimated biomass based on a similar calculation of volume for wheat. The type of sensor used in both studies is referred to as Light Detection And Ranging (LiDAR), operating on time-of-flight principle.

Kjaer and Ottosen (2015) mounted a 3D laser on an automated boom, which allowed it to move along one axis, hence, covering a larger area of observation. The scan provided height, 3D leaf area, and projected leaf area. These measurements were used as predictors of properties usually measured destructively, including dry and fresh mass. The results were later used to identify the most convenient cultivars of rapeseed. Also, studying leaf movement and changes in canopy configuration was proposed as an application of the device.

The possibility of improving the estimation of biomass by adding multispectral sensors has been considered (Tilly et al., 2015; Schaefer and Lamb, 2016). The former study used a 3D laser scanner in a barley field, while the latter used a 2D LiDAR sensor mounted on a moving vehicle in a tall fescue field. After examination of NDVI and other vegetation indexes, it was determined that they produced better results. However, the improvements were small in both cases, casting doubt on the relevancy of this complementary sensor in relation to the added cost and complexity.

Extending this idea to larger scales has been tested in (Zarco-Tejada, Diaz-Varela, Angileri, & Loudjani, 2014) and (Torres-Sánchez, López-Granados, Serrano, Arquero, & Peña, 2015), where the canopy properties are computed from camera images taken from a UAV. This is different from the use of reflectance-based sensors mounted in UAV, as in (Engström et al., 2009), where the canopy architectonics are not considered explicitly, yet the biomass or other plant properties can be estimated non-destructively. In (Tang & Shao, 2015) and (Elaksher, Bhandari, Carreon-Limones, & Lauf, 2017) canopy measurements are measured with a LiDAR system mounted on an UAV, but no other plant properties are further derived from the results of the direct measurement.

Biomass estimation is not the only application of this principle that has been presented. Further work has been done in using canopy properties as feedback within a control loop. For example, in (Zaman, Schumann, & Miller, 2005), prescription maps for fertilization were built based on the measured canopy volume. Similarly, in (Escolà et al., 2009), a foliar index was created and related to a cross-sectional area and height of tree canopies. The results were fed to a scheduler for fertilizer and pesticide applications. In tree orchards, the size of the tree canopy can be related to fruit yield, as was done in (Zaman, Schumann, & Miller, 2006) with citrus.

The work presented in this manuscript is a continuation of Chapter 3 of (Su, 2017), where an ultrasonic proximity sensor was used in a similar setup, characterized by the rotating sensor head, used for arugula and spinach, and its measurements were related with fresh and dry biomass. The base hypothesis is that the increased accuracy of using a laser

sensor would improve biomass estimation (Tumbo, Salyani, Whitney, Wheaton, & Miller, 2001). Additionally, the height profiles resulting from the rotating sensor head are expected to rank between simple height and volume-based estimations. In (Arnó et al., 2009), the intuition that volume-based estimates produce better results than height-based is confirmed. Nonetheless, given the results from (Su, 2017), height profiles may be a more practical option to handle the trade-off between accuracy and cost of the sensor system, given the restrictions that need to be followed for implementations meant to be operated in a farm environment. In (Moorthy, Miller, Hu, Chen, & Li, 2008), it is shown that canopy properties can be readily determined in a lab setting.

Other agricultural applications and characteristics of Light Detection And Ranging (LiDAR) are presented in (Omasa, Hosoi, & Konishi, 2006) and (Ehlert, Heisig, & Adamek, 2009). Among the aforementioned, the optimization of combined parameters during harvesting for current crop conditions and the improvement in guidance of agricultural machinery based on the detection of crop edges are some of the previously least explored topics.

3. MATERIALS AND METHODS

3.1. DESIGN AND CONSTRUCTION OF THE DEVICE

The device combined laser, ultrasonic and thermal infra-red (IR) measurements from the following commercially available sensors¹: IL-600 (Keyence Corporation, Itasca, IL, USA), ToughSonic14 (Senix Corporation, Hinesburg, VT, USA), and SSS-LT (Process Sensors Corporation, Milford, MA, USA), respectively. 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 keep the general setup of a hand-held device, a tripod was selected as the main frame for the 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 which allowed for changes in their position along the holder's central axis. The holder was connected from an edge to the shaft of the controlled stepper motor. In this configuration, the sensors were arranged to spin in a circular path with a specific radius. 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. The capability of the holder to change the position of the sensors allowed the radius to be modified between 7 cm and 15 cm.

The three included sensors had analog output. The laser and ultrasonic outputs were received by a data acquisition board (NI myDAQ, National Instruments Corporation, Austin, TX, USA), while the thermal IR measurements were read by an Arduino UNO board (Arduino AG, Chiasso, Switzerland). Both boards were working as Analog-to-Digital

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

Converters, transmitting the resulting signal to a laptop computer. LabView 2013 (National Instruments Corporation, Austin, TX, USA) was used to read and log the sensors' measurements along with a corresponding time stamp, send commands to the motor, and provide a user interface. The block diagram is shown in Figure 1. The sensors and their accompanying circuitry, as well as the motor, were powered by a 6-cell LiPo battery with voltage converters. A separate box was built with High-Density Polyethylene (HDPE) to contain the circuitry. Figure 2 shows the complete arrangement.

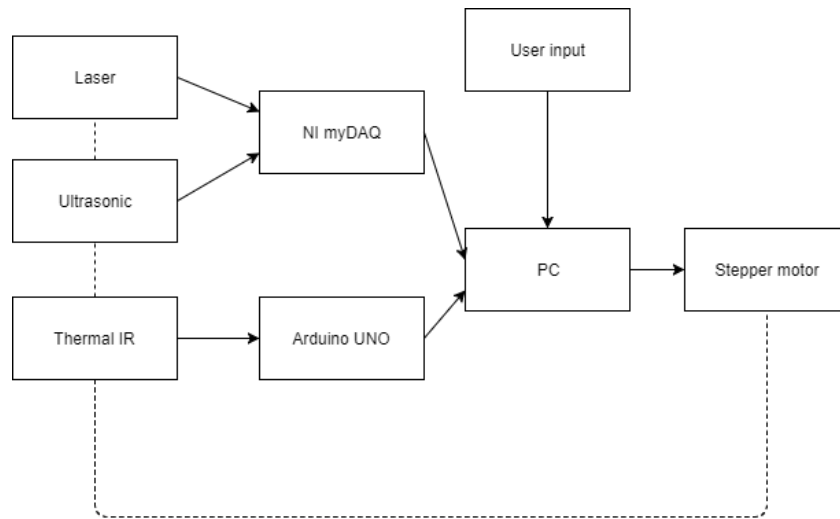


Fig. 1. Block diagram of the system



Fig. 2. Picture illustrating the complete device in the greenhouse setup

3.2. CONTROLLED ENVIRONMENT EXPERIMENT DESIGN

The controlled environment experiment took place within a greenhouse located at Macdonald Campus of McGill University, Montreal, Canada. On March 28th, 2017, four hydroponic beds were allocated with 29 plants each (15 lettuce and 14 kale) plus one of 28 plants (14 of each type), as depicted in Figure 3. This was the maximum number of plants each hydroponic bed could fit. Initially, all 5 beds were meant to have 29 plants, but there were not enough seedlings available. Each plant was placed in a 4”x4”x4” Rockwool cube. The cubes were distributed in 4 rows of 7, and in those beds with 29, there was a single cube on the edge. The crop varieties employed were romaine lettuce (*Lactuca sativa L. var. longifolia*) and red Russian kale (*Brassica napus subsp. pabularia*). They were germinated for two weeks in a growth chamber before being transferred into the hydroponic system. The beds were irrigated by an ebb and flow system, where the half-strength Hoagland solution was added periodically. Each bed had an independent ON/OFF controller to activate the water pumps with soil moisture sensors as feedback. Each bed had different lower and upper thresholds for the controller, as depicted in Table 1, to simulate the effect of water stress.

Table 1. Thresholds used to simulate water stress

Bed	Lower limit [%]	Upper limit [%]
1	50	100
2	50	85
3	50	75
4	10	100
5	25	85

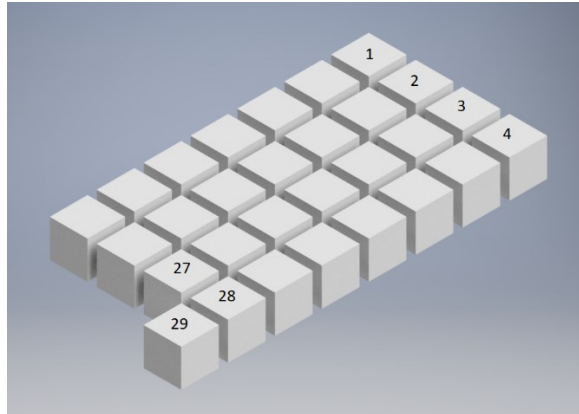


Fig. 3. Illustration of the Rockwool cubes' distribution in the hydroponic bed

On April 25th, the plants were harvested and the following properties were measured: fresh aboveground biomass, dry aboveground biomass, maximum height, maximum extended height, and the number of leaves. Since in all cases the measured biomass considers only the aboveground components, the term will no longer be specified. Both fresh and dry mass were measured using an electronic scale MXX-612 (Denver Instrument Inc., Bohemia, NY, USA). Maximum height was measured between the soil level (the top face of the Rockwool cube) and the highest point of the plant, without interfering with the plant's pose. Maximum extended height was measured between the soil level and the highest point of the plant, while manually straightening the plant as vertical as possible without breaking it. Intermediate harvesting was done at three earlier dates throughout the experiment where fewer plants were collected, as stated in Table 2. Figure 4 shows the state of the beds at 42 Days After Seeding (DAS).

Table 2. Schedule for data collection and plant harvesting

Date	DAS	Plants harvested per bed
April 14 th	32	5
April 17 th	35	N/A
April 18 th	36	4
April 19 th	37	N/A
April 20 th	38	N/A
April 21 th	39	6
April 22 th	40	N/A
April 24 th	42	N/A
April 25 th	43	14



Fig. 4. Picture illustrating experimental setup in greenhouse

The experiment design was a completely randomized design with temporal repeated measurements. It was assumed that the temporal repeated measurements were done over the beds instead of the plants. The setup was a part of a hydroponic system designed for a different experiment (Tikasz & Lefsrud, 2017). The treatment for this experiment was the water stress in each bed. Equation (1) describes this design, where Y_{ijt} denotes any of the measured plant properties i.e. mass and height, for bed i , plant j , and day t .

$$Y_{ijt} = \mu + a_i + \varepsilon_{ij}^{(1)} + c_t + (ac)_{it} + \varepsilon_{ijt}^{(2)} \quad (1)$$

In the above equation, μ is the overall mean, a_i is the effect of treatment i , c_t is the effect of time, $(ac)_{it}$ is the effect of the interaction between treatment and time, and $\varepsilon_{ij}^{(1)}$ and $\varepsilon_{ijt}^{(2)}$ are error terms. Under the scope of this experiment, the treatment was done solely to provide a wide range of biomass values to test the device. It was of interest to prove that the system can provide meaningful predictions across the whole range of values attained.

The developed device was used throughout the experiment according to the schedule shown in Table 2. Each sensor provided a measurement every 45 min. The motor was set to spin at a constant speed of $20^\circ/\text{s}$, providing a point every 0.9° that contains information from the three sensors. Consequently, a complete lap provided a profile of 400 points. The sensor holder was set at the maximum radius of 15 cm for all cases. In each bed, three locations were used, labeled Close, Middle, and Far, in terms of the ends of the beds. In each of these locations, around 5 laps were recorded with the device, covering both the kale and lettuce sides.

For the reported dates, other sensors were used to keep track of the ambient state, which were not part of the developed device. These sensors included thermal camera (C2, FLIR Systems Inc., Wilsonville, OR, USA), PAR/quantum sensor (QMSS-ELEC, Apogee Instruments Inc., Logan, UT, USA), ambient temperature and moisture (DHT22, Adafruit Industries, New York, NY, USA), and multispectral (Crop Circle ACS-430, Holland Scientific Inc., Lincoln, NE, USA). This last sensor, the Crop Circle, outputs the reflectance of the crop to light at different wavelengths: red, red-edge, and NIR. Additionally, two Vegetation Indexes (VI) are computed from those reflectance values: NDVI and NDRE. One measurement was taken per day per bed from each of the previously mentioned sensors.

3.3. CONTROLLED ENVIRONMENT ANALYSIS

The estimation of biomass M_{ijt} is usually based on canopy characteristics, e.g. plant height H_{ijt} , as presented in Equation (2). A regression, either linear or nonlinear, can be used to test how well a certain type of function fits the data. Some error ε_{ijt} 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.

$$M_{ijt} = f_1(H_{ijt}) + \varepsilon_{ijt}^{(3)} \quad (2)$$

The idea behind the device's design was to define a relationship from several height values, i.e. the height profile $\mathbf{h}_{ijt}$, and improve the ability to accurately predict biomass. This is expressed in Equation (3).

$$M_{ijt} = f_3(\mathbf{h}_{ijt}^{laser}, \mathbf{h}_{ijt}^{ultrasonic}) + \varepsilon_{ijt}^{(4)} \quad (3)$$

Furthermore, it may be the case that the biomass model described in Equation (3) can be improved by including other sensor measurements x_{ijt} , as stated by the following equation.

$$M_{ijt} = f_3(\mathbf{h}_{ijt}^{laser}, \mathbf{h}_{ijt}^{ultrasonic}, x_{ijt}^{(1)}, \dots, x_{ijt}^{(n)}) + \varepsilon_{ijt}^{(5)} \quad (4)$$

The goal was then to find some $f_2: \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}$ or $f_3: \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R} \times \dots \times \mathbb{R} \mapsto \mathbb{R}$ such that $\sigma_{(5)}^2 < \sigma_{(4)}^2 < \sigma_{(3)}^2$, under the assumption that the errors were normally distributed $\varepsilon_{ijt}^{(5)} \sim N(0, \sigma_{(5)}^2)$, $\varepsilon_{ijt}^{(4)} \sim N(0, \sigma_{(4)}^2)$ and $\varepsilon_{ijt}^{(3)} \sim N(0, \sigma_{(3)}^2)$; where m is the number of points recorded in a lap by either laser or ultrasonic sensors. One disadvantage of this approach was that the existence of a physical interpretation for the chosen f_2 could not be guaranteed.

The data were imported into MATLAB R2017a (MathWorks Inc., Natick, MA, USA) and SAS University Edition (SAS Institute Inc., Cary, NC, USA) for this analysis. With SAS, the procedure GLM (Generalized Linear Model) was used. First, the profiles were separated depending on the type of plant observed. Then, from the sets of laser and ultrasonic measurements, a scalar value was derived for each lap using different processing methods, which might be used as inputs to the function f_2 . Table 3 describes the alternatives considered. The values of 0, corresponding to spots with no crop (soil level), were filtered out for the median and mode. For the IR thermal measurement, only the mean value was computed. Finally, linear and exponential regressions were performed between the plant properties measured directly and the sensor measurements. 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 criterion. Finally, when a working model was achieved, the mean absolute percentage error (MAPE) was computed to indicate the precision of the system using Equation (5). This was not used as a criterion to choose between different models. When exponential regression was

used, instead of MAPE, the PE was computed as presented in Equations (6) and (7) for the cases of overestimation and underestimation, respectively.

Table 3. Definition of processing methods for controlled environment experiment

Name in MATLAB	Description
Average crop height	Computes numeric definite integral and divides by the size of the integration interval
Average filtered crop height	Applies median filter with window size 10, then computes average crop height
Average recorded crop height	Computes arithmetic mean
Max	Finds the maximum value
Median	Removes all 0 values, then computes median
Mode	Removes all 0 values, then computes mode
Mean of max	Finds maxima between sets of 10 points, then averages them

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{actual_i - estimate_i}{actual_i} \right| \quad (5)$$

$$PE_{overestimation} = e^{RMSE} - 1 \quad (6)$$

$$PE_{underestimation} = 1 - e^{-RMSE} \quad (7)$$

3.4. DEVELOPMENT OF A FIELD SYSTEM

Several improvements were implemented before taking the device to a field environment. For instance, the sensor holder was changed to incorporate a fixed angle with respect to the vertical axis. The new holder is presented in Figure 5. The new holder design was sturdier and more rigid. By using shorter beams, the bending load was diminished, reducing the risk of mistakes in the sensor position due to deformation of the holder. In this configuration, the radius of the circular paths could be adjusted by changing the height of the

middle bar of the tripod. In any case, for the production field experiment, a fixed radius was used.

The laptop computer that was used to run the control system via LabVIEW was replaced by a Raspberry Pi 2 Model B (Raspberry Pi Foundation, Cambridge, UK) with a python script, shown in Appendix A. This change had multiple benefits for the implementation: use of free software, reduced weight, reduced cost, increased portability, increased battery life. About the latter, the power supply was taken from the same LiPo battery that powered the rest of the electronics rather than the computer internal battery; the power consumption of the Raspberry Pi was minimal when compared to a regular laptop computer. On the other hand, the disadvantage was the loss of the user interface. The new block diagram is shown in Figure 6.



Fig. 5. Picture illustrating the new holder

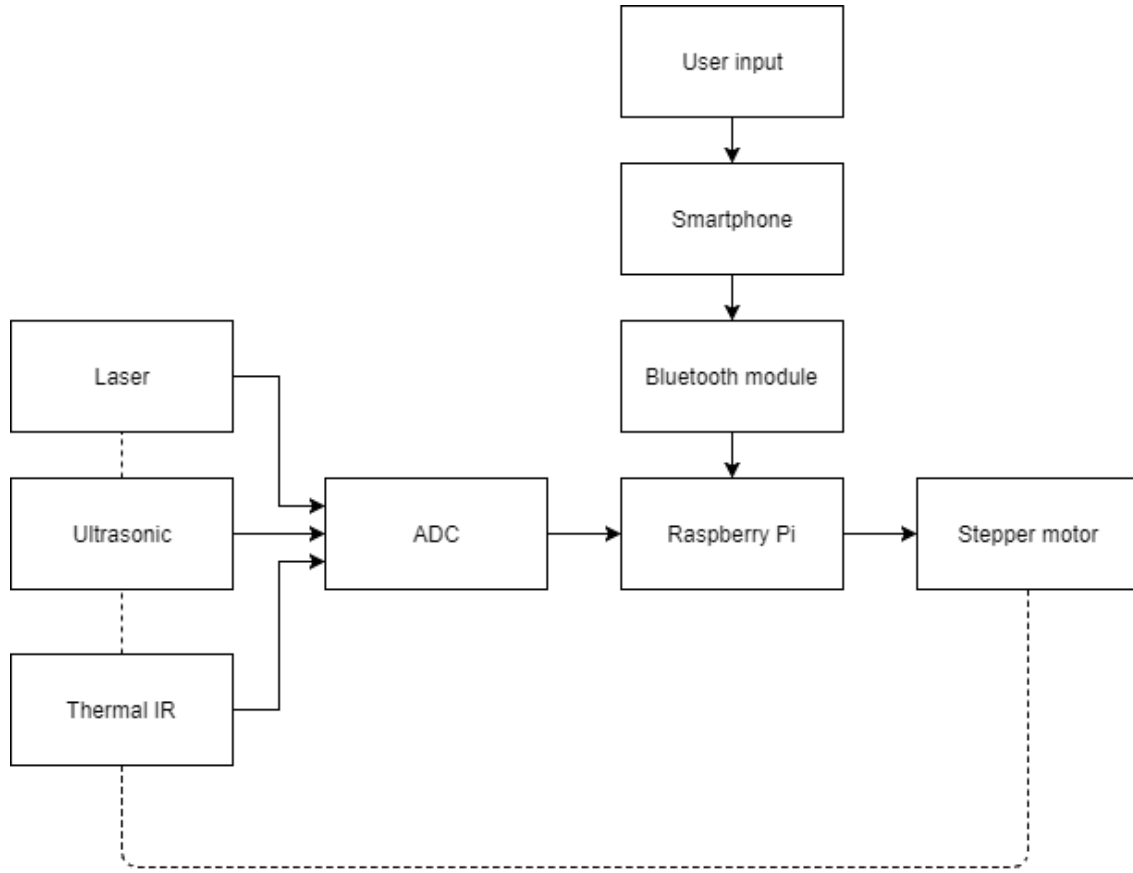


Fig. 6. Block diagram of modified system

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.

As a consequence of the change from laptop to Raspberry Pi, the NI myDAQ acquisition board and the Arduino UNO were replaced by two ADS1115 (Adafruit Industries, New York, NY, USA). These ADC offer two differential 16-bit channels that communicate with the Raspberry Pi via I2C protocol. Again, this helped to reduce the weight of the device.

Another modification was that the box containing the circuitry was fixed on top of the tripod. This step could only be done at this stage thanks to the reduced weight of the electronic components. In Figure 7, the modified device is presented.



Fig. 7. Picture illustrating setup for the production field experiment

3.5. PRODUCTION FIELD EXPERIMENT DESIGN

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 shorter time span, only requiring three days because of the time-consuming sampling of the fresh biomass. In this experiment, only fresh biomass was considered. A 50 cm x 50 cm 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.

Each of the sensors included in the device provided a measurement every 0.8° , creating a profile of 450 points per lap. For each location, about 6 laps were recorded, around 3 different centers inside the frame. From the set of complementary sensors used in the controlled environment experiment, only the ambient temperature and moisture (DHT22) was also used in the production field experiment. It was found that moisture was the best complement for the laser in the previous experiment.

3.6. PRODUCTION FIELD ANALYSIS

For the production field experiment, 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 that were taken around the same center point produced one height profile. Within each square frame, 3 different centers were used, as mentioned in Section 3.5.

Two approaches were taken to analyze the data from the production field experiment. The first approach (regression of profile-representative features) was the same as that used in the controlled environment experiment, where the function providing the best fit was considered. Here, the process was extended by considering a larger pool of processing methods, as well as functions that mathematically combined more than one of these methods. The regressions that were performed were linear, exponential, and polynomial up to degree 5. Table 4 contains the complete list of processing methods used in this approach. The code for these methods is presented in Appendices B to AA. The last two methods are the processing methods used by Su (2017).

Table 4. Definition of processing methods for production field experiment

Name in MATLAB	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 (8), 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 i^{th} plant (labeled H_i) depended only on the angular position, and not on the radius.

$$\begin{aligned}
 M_i &= \int \rho \, dV = \int_0^{R_l} \int_0^{2\pi} \int_0^{H_i(\theta, r)} r \, \rho(z, \theta, r) \, dz \, d\theta \, dr = \int_0^{R_l} \int_0^{2\pi} \int_0^{H_i(\theta)} r \, \rho(z) \, dz \, d\theta \, dr + \varepsilon_i \\
 &= \int_0^{R_l} r \, dr * \int_0^{2\pi} \int_0^{H_i(\theta)} \rho(z) \, dz \, d\theta + \varepsilon_i = \frac{R_l^2}{2} * \int_0^{2\pi} \int_0^{H_i(\theta)} \rho(z) \, dz \, d\theta + \varepsilon_i \quad (8)
 \end{aligned}$$

where M_i denotes the mass of the i^{th} 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 i^{th} plant. It is worth noting that Equation (8) can be rewritten as Equation (9) based on the Fundamental theorem of calculus. Here the M_{ijt} notation was dropped to improve readability, but it would also hold if added consistently.

$$M_i = \frac{R_l^2}{2} * \int_0^{2\pi} \int_0^{H_i(\theta)} \rho(z) \, dz \, d\theta + \varepsilon_i = \frac{R_l^2}{2} * \int_0^{2\pi} f(H_i(\theta)) \, d\theta + \varepsilon_i \quad (9)$$

for some function $f: \mathbb{R} \mapsto \mathbb{R}$ 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 (9) becomes Equation (10).

$$D_i = \frac{M_i}{A} = \frac{\frac{R_l^2}{2} * \int_0^{2\pi} f(H_i(\theta)) d\theta + \varepsilon_i}{\pi R^2} = \frac{1}{2\pi} \int_0^{2\pi} f(H_i(\theta)) d\theta + \varepsilon_i^{(6)} \quad (10)$$

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 $\rho(z) = \rho \rightarrow D_i = \frac{\rho}{2\pi} \int_0^{2\pi} H_i(\theta) d\theta = a * Integral + b$, 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 $f(x) = x^2 \rightarrow D_i = \frac{1}{2\pi} \int_0^{2\pi} [H_i(\theta)]^2 d\theta = a * Energy + b$, 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. Appendices AB and AC present the code used for this approach, with a specific example of a candidate function.

4. RESULTS AND DISCUSSION

4.1. CONTROLLED ENVIRONMENT EXPERIMENT

Significant differences were found over time and with water stress (bed). Consequently, a diverse set of biomass and canopy properties was present, providing an adequate setup for the evaluation of the designed device. Figures 8 and 9 compare the growth of the plants across the beds.

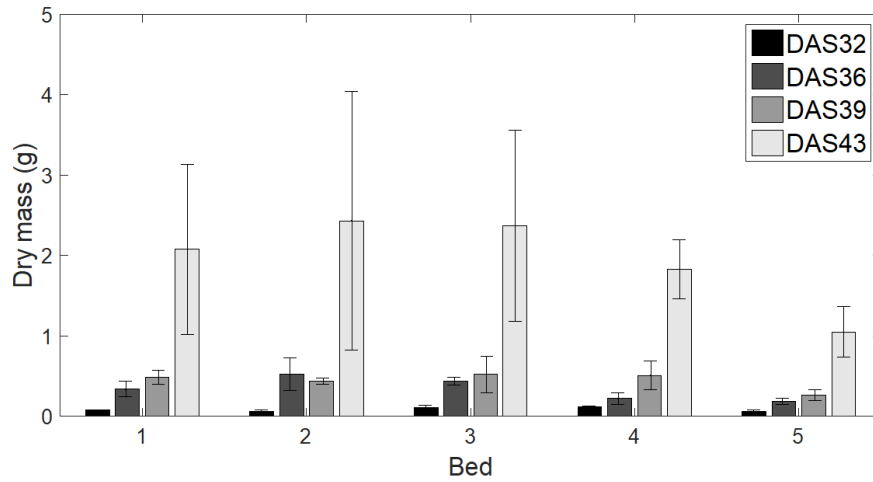


Fig. 8. Evolution of average lettuce dry biomass against bed

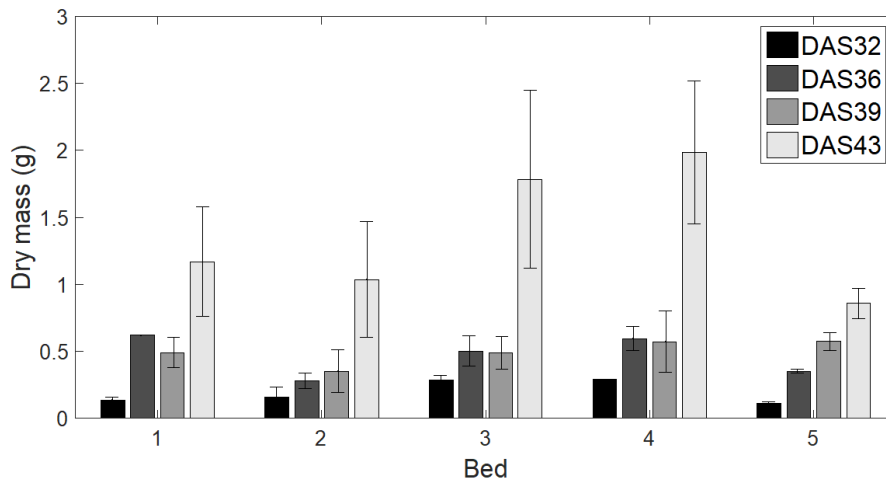


Fig. 9. Evolution of average kale dry biomass against bed

Similar results were obtained for fresh mass, maximum height and number of leaves. It was observed that the best performance for kale was in bed 4, while for lettuce, beds 2 and 3 were similar to each other. For both plants, bed 5 had the lowest yield recorded. It is believed that the mass increases monotonically, despite cases like the kale dry mass in bed 1, and it just happened that the samples collected were particularly large by chance. This could be explained by the increasing the variability in the measurements with time; as well, the samples for the intermediate harvesting were relatively small.

Additionally, by considering the mass against the manually measured maximum plant height, it was found that an exponential model provided the best fit. This proved the existence of function f_1 from Equation (2) and gives hints about the possible form of f_2 . In the following steps of the analysis, the exponential case was considered. Figures 10 and 11 present this regression for lettuce and kale, respectively.

All the beds were included alike in these regressions. Better fitting could be found by considering each bed separately. Also, improved results were found when using maximum extended height, increasing the R^2 values to 0.93 and 0.94 for lettuce and kale, respectively.

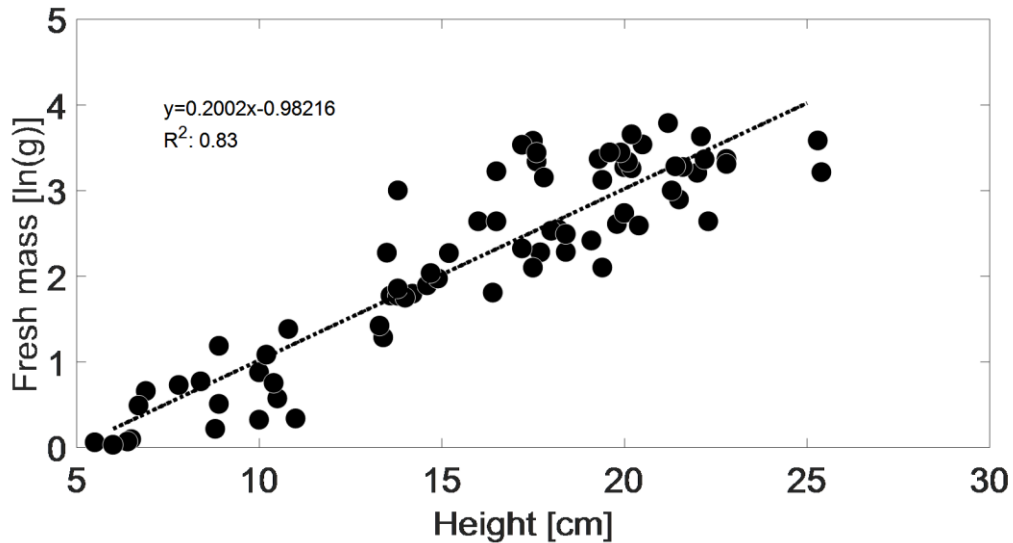


Fig. 10. Exponential regression of lettuce fresh biomass vs maximum height

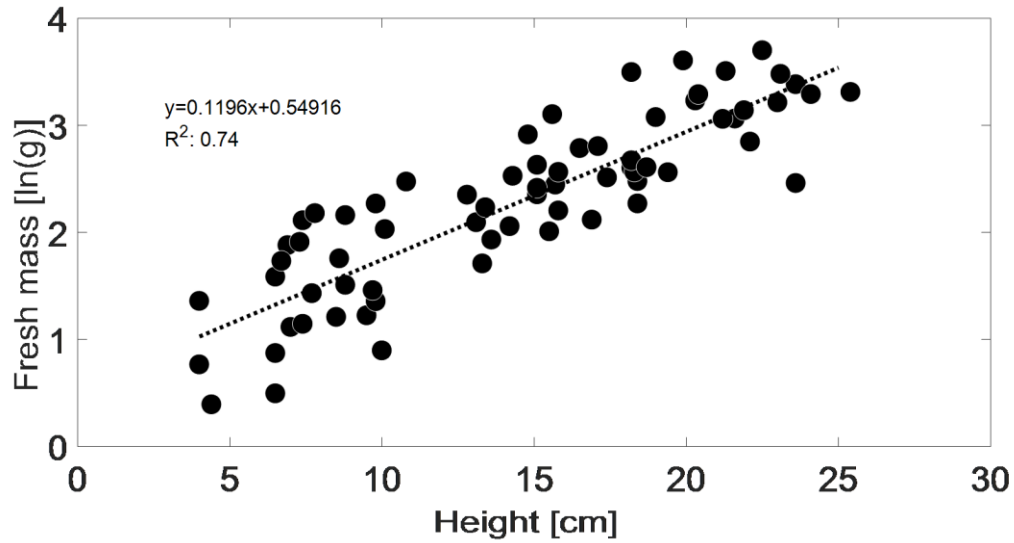


Fig. 11. Exponential regression of kale fresh biomass vs maximum height

As mentioned in the previous sections, a profile of plant height measurements was taken along a circular path. Figures 12 and 13 illustrate an example of a lap for bed 1 in DAS 36 and bed 2 in DAS 43, and some of the processing methods evaluated. With 0 being the ground level, some negative values were found due to gaps between the Rockwool blocks, which can be seen as holes in the soil. These negative values were replaced by zeros before applying the different processing methods mapping to single scalars. In an average sample, these values less than or equal to 0 made up 5% of the available data. Later, the scalars were averaged with those of the same bed in the day.



Fig. 12. Picture illustrating circular path over the crops at middle location for (a) DAS 36 and (b) DAS 43

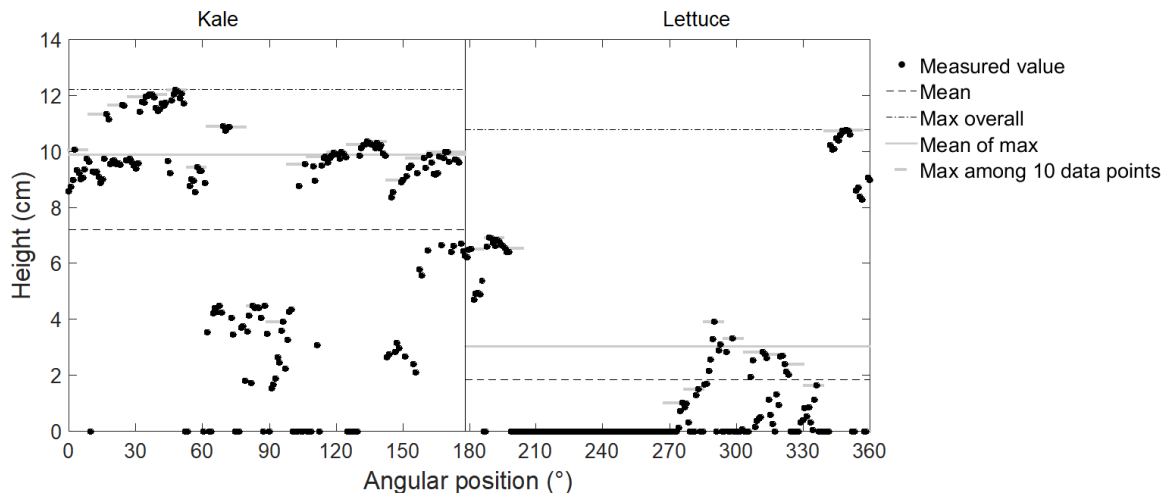


Fig. 13. Example of height profile as recorded by ultrasonic sensor and calculated processing methods

Table 5 provides the R^2 values found for the exponential regression, based on each of the considered processing methods. The best case is highlighted with bold font.

Table 5. Coefficients of determination from exponential regression between the biomass and processed ultrasonic height measurements for each processing method

Function	Fresh lettuce	Dry lettuce	Fresh kale	Dry kale
Average crop height	0.80	0.75	0.63	0.56
Average filtered crop height	0.69	0.63	0.54	0.50
Average recorded crop height	0.80	0.74	0.56	0.53
Max	0.69	0.62	0.55	0.52
Median	0.80	0.75	0.48	0.47
Mode	0.69	0.63	0.65	0.58
Mean of max	0.78	0.71	0.57	0.48

Figures 14 and 15 present the best fit for each case. Similar to what was presented for the ultrasonic sensor, Figure 16 shows an example of a profile of measurements for bed 4 in different days of the experiment. The profile from ultrasonic is retained for comparison.

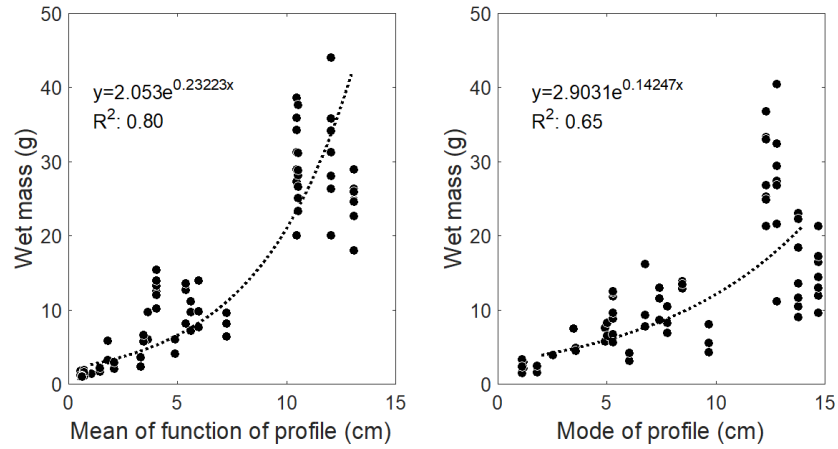


Fig. 14. Best model of exponential regression for (a) lettuce and (b) kale fresh biomass based on ultrasonic

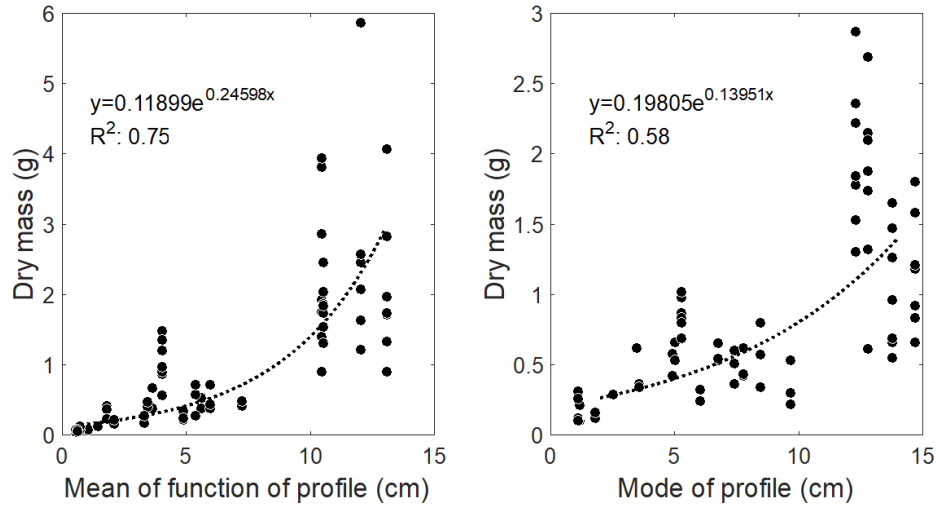


Fig. 15. Best model of exponential regression for (a) lettuce and (b) kale dry biomass based on ultrasonic

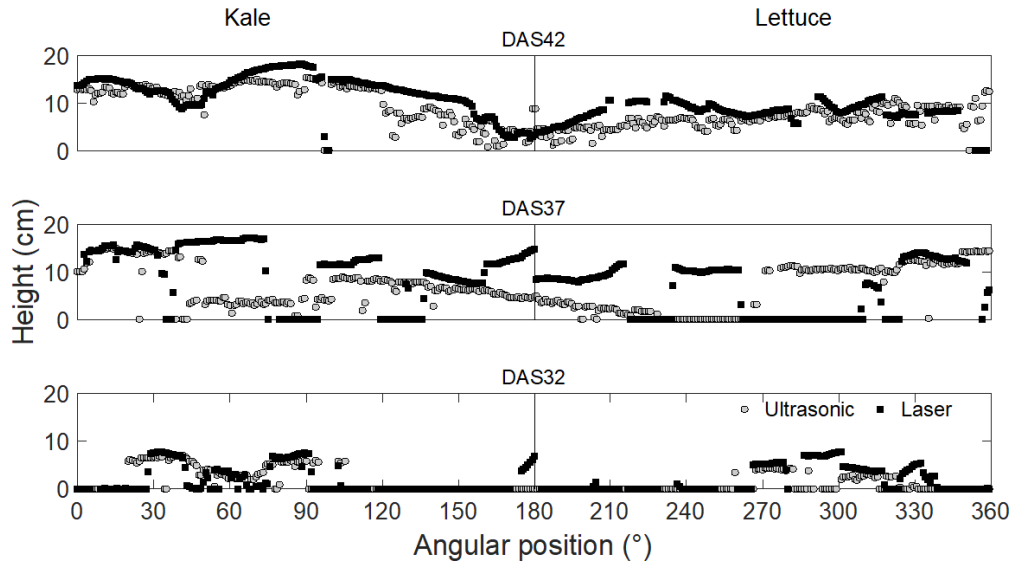


Fig. 16. Example of height profiles as recorded by laser and ultrasonic sensors for different days of the experiment

The smaller spot size of the laser allowed for the identification of canopy structures that might have passed unseen by the ultrasonic sensor, as well as provided a better determination of the ground level. The ultrasonic signal was noisy compared to the laser measurement; however, the laser had more failed readings. This could happen for a variety of reasons, among them, that the reflected laser beam did not reach the lens or excessive

sunlight saturated the sensor. Table 6 shows the R^2 values for the regression based on the laser, and highlights the best processing method in bold letters. Figures 17 and 18 show the calibration step based on the laser measurements, with the best fitting processing methods.

Table 6. Coefficients of determination from exponential regression between the biomass and processed laser height measurements for each processing method

Function	Fresh lettuce	Dry lettuce	Fresh kale	Dry kale
Average crop height	0.87	0.80	0.51	0.50
Average filtered crop height	0.75	0.66	0.67	0.58
Average recorded crop height	0.88	0.81	0.60	0.48
Max	0.75	0.66	0.52	0.44
Median	0.87	0.80	0.56	0.43
Mode	0.75	0.66	0.69	0.61
Mean of max	0.78	0.73	0.61	0.49

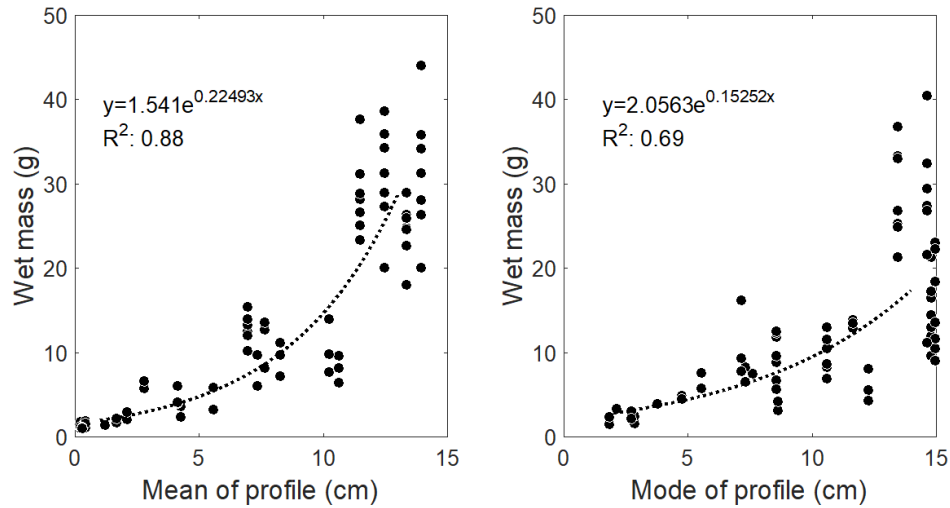


Fig. 17. Best model of exponential regression for (a) lettuce and (b) kale fresh biomass based on laser

At this point, the advantage of considering the information from both sensors at the same time was evaluated. From the previous sections, it was known what functions worked best for each sensor, as reported in Table 7, and labeled as f_k and f_l . Then, the two sensors

were combined in the way depicted in Equation (11). This was an extension of the exponential regression.

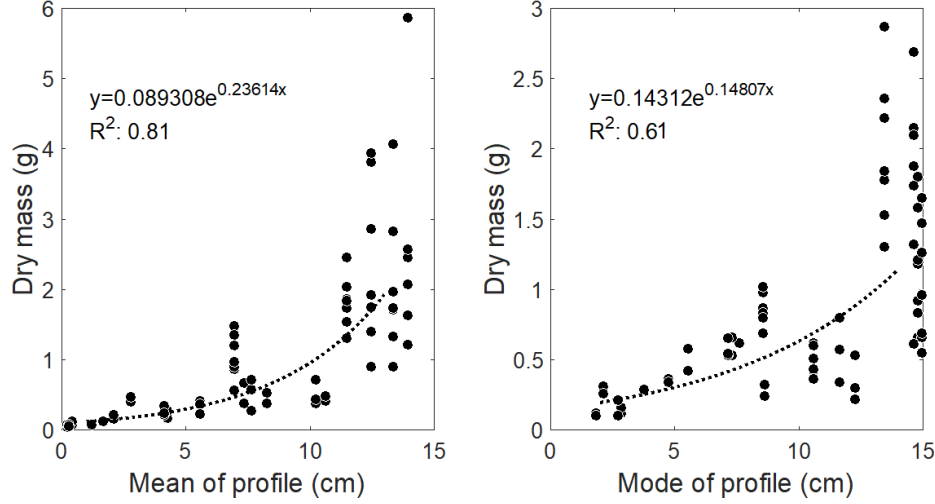


Fig. 18. Best model of exponential regression for (a) lettuce and (b) kale dry biomass based on laser

Table 7. Processing methods found to produce the best results in exponential regression

Case	Method for laser	Method for ultrasonic
Lettuce fresh mass	Average recorded crop height	Average crop height
Lettuce dry mass	Average recorded crop height	Average crop height
Kale fresh mass	Mode	Mode
Kale dry mass	Mode	Mode

$$\ln[f_2(\mathbf{h}_{ijt}^{laser}, \mathbf{h}_{ijt}^{ultrasonic})] = p + qf_k(\mathbf{h}_{ijt}^{laser}) + rf_l(\mathbf{h}_{ijt}^{ultrasonic}) + sf_k(\mathbf{h}_{ijt}^{laser})f_l(\mathbf{h}_{ijt}^{ultrasonic}) \quad (11)$$

With this procedure, Table 8 was built. This table allows for a comparison of the performance of using the ultrasonic sensor, the laser, or both sensors. It was found that, while having both sensors offers the best results, the difference against having only the laser was not very large. There was a clear difference between the laser and ultrasonic sensors, with the advantage favoring the former. The RMSE values were found in the logarithmic scale.

Table 8. Comparison of fitting in exponential regression between laser, ultrasonic and both

Case	Ultrasonic		Laser		Both	
	R^2	RMSE [ln(g)]	R^2	RMSE [ln(g)]	R^2	RMSE [ln(g)]
Lettuce fresh mass	0.80	0.511	0.88	0.394	0.90	0.356
Lettuce dry mass	0.75	0.623	0.81	0.541	0.83	0.523
Kale fresh mass	0.65	0.483	0.69	0.456	0.71	0.443
Kale dry mass	0.58	0.545	0.61	0.529	0.62	0.525

The following steps were taken supposing that the function in Equation (11) may be further improved by adding other types of sensors, as anticipated in Equation (4). Table 9 presents correlation values for the different sensors other than the laser and ultrasonic sensor. For brevity, they are not presented in the usual form of a symmetric matrix.

Table 9. Pearson correlation values between plant biomass and other measured properties

Sensor measurement	ln(lettuce freshmass)	ln(lettuce drymass)	ln(kale freshmass)	ln(kale drymass)
IR thermal*	-0.564	-0.655	-0.486	-0.617
IR thermal Std. Dev.	-0.195	-0.196	0.002	-0.050
PAR/Quantum	-0.278	-0.343	-0.061	-0.189
Moisture*	0.800	0.795	0.790	0.755
Ambient temperature*	-0.898	-0.913	-0.830	-0.808
Red-Edge*	-0.897	-0.901	-0.855	-0.850
NIR*	0.905	0.899	0.836	0.796
Red*	-0.827	-0.831	-0.750	-0.750
NDRE*	0.500	0.527	0.510	0.589
NDVI	0.051	0.108	0.108	0.251

*: Indicates variables whose correlation have p-values below 0.0001

By producing a regression model after having added an extra measurement besides the laser and ultrasonic, it was possible to test the improvement that each type of sensor produced. Table 10 shows the obtained R^2 values. Only those sensors with a significant correlation were evaluated.

For lettuce, the most important effect was achieved when adding the moisture measurement into the estimation. For kale, on the other hand, it occurred when Red-Edge was added. Moisture would be selected, due to its reduced size and cost as compared to the multispectral sensor. No major changes in the design would be required to incorporate the moisture sensor into the device.

Table 10. Coefficients of determination from exponential regression between the biomass and combination of laser and ultrasonic height measurements with other measured properties

Case	No extra	IR as extra	Temperature as extra	Moisture as extra	Red-Edge as extra	NIR as extra
Lettuce fresh mass	0.90	0.91	0.93	0.94	0.93	0.92
Lettuce dry mass	0.83	0.84	0.87	0.88	0.88	0.88
Kale fresh mass	0.71	0.71	0.75	0.77	0.83	0.80
Kale dry mass	0.62	0.64	0.73	0.74	0.78	0.76

In Table 11, the calibration RMSE and Percentage Errors (PE) for underestimation and overestimation are reported. While the former was based on logarithmic scale, the latter was with respect to grams. These were used to show how each step provided progress in further improving the estimations. Figures 19 and 20 present the estimated values against the actual mass values found in the experiment, using the final approach where the moisture measurements were included.

Table 11. Summary of results from exponential regressions

Case	Approach	R^2	R^2_{adj}	RMSE [ln(g)]	PE Under [%]	PE Over [%]
Lettuce fresh mass	Maximum height	0.83	0.83	0.469	37.4	59.8
	Ultrasonic	0.80	0.80	0.511	40.0	66.6
	Laser	0.88	0.88	0.394	32.6	48.3
	Laser & Ultrasonic	0.90	0.90	0.356	30.0	42.8
	Moisture as extra	0.94	0.93	0.295	25.5	34.3
Lettuce dry mass	Maximum height	0.75	0.74	0.624	46.4	86.7
	Ultrasonic	0.75	0.75	0.623	46.4	86.5
	Laser	0.81	0.81	0.541	41.8	71.8
	Laser & Ultrasonic	0.83	0.82	0.523	40.7	68.8
	Moisture as extra	0.88	0.88	0.437	35.4	54.7
Kale fresh mass	Maximum height	0.75	0.74	0.413	33.8	51.1
	Ultrasonic	0.65	0.65	0.483	38.3	62.1
	Laser	0.69	0.68	0.456	36.6	57.7
	Laser & Ultrasonic	0.71	0.70	0.443	35.8	55.8
	Moisture as extra	0.77	0.75	0.402	33.1	49.5
Kale dry mass	Maximum height	0.60	0.59	0.535	41.5	70.8
	Ultrasonic	0.58	0.58	0.545	42.0	72.5
	Laser	0.61	0.60	0.529	41.1	69.8
	Laser & Ultrasonic	0.62	0.61	0.525	40.9	69.1
	Moisture as extra	0.74	0.72	0.441	35.7	55.5

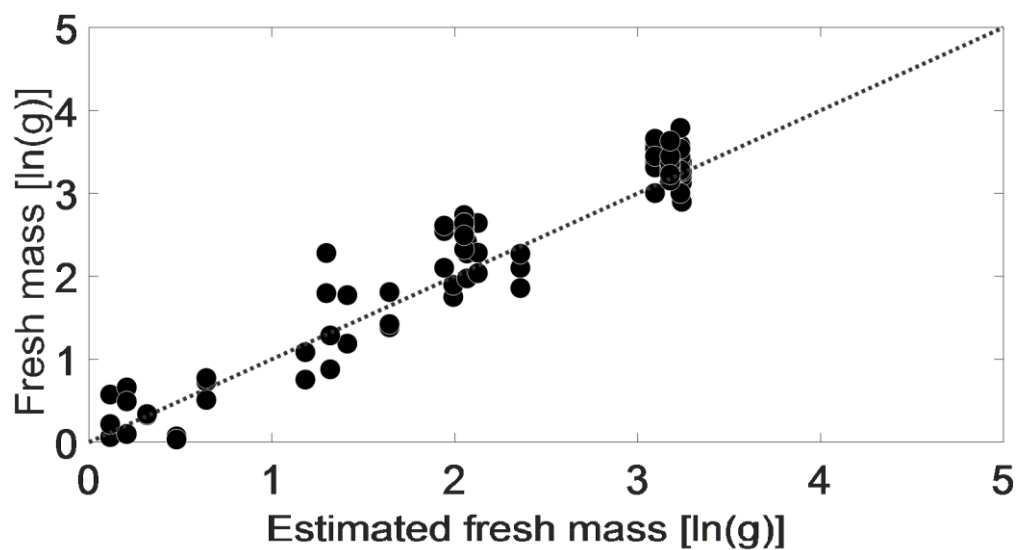


Fig. 19. Actual fresh biomass against fresh biomass predicted by the device for lettuce

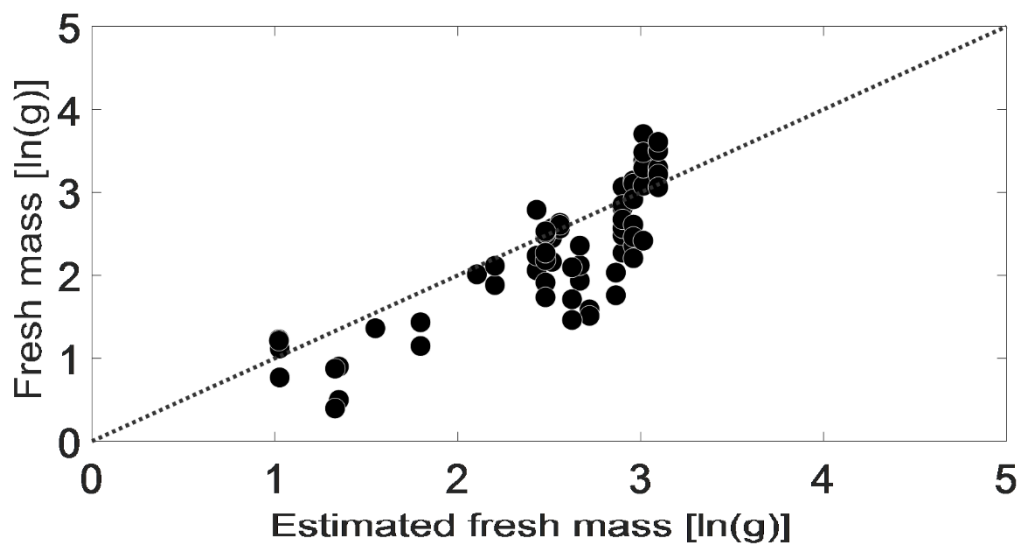


Fig. 20. Actual fresh biomass against fresh biomass predicted by the device for kale

4.2. PRODUCTION FIELD EXPERIMENT

Figure 21 presents the fresh biomass yield manually measured within the square frame for all dates. This yield was simply the biomass divided by the area of the square frame (0.25 m^2) and transformed to t/ha units. These data have a mean value of 10.77 t/ha and standard deviation of 7.93 t/ha. The difference between the maximum and minimum values was 32.88 t/ha, proving that a varied range of biomass yield had been recorded. It seemed to not be normally distributed, but rather similar to a log-normal. Whatever the case, the exact distribution was not relevant for the analysis. A drawback of the experiment was the absence of samples in the range of 20 t/ha to 30 t/ha.

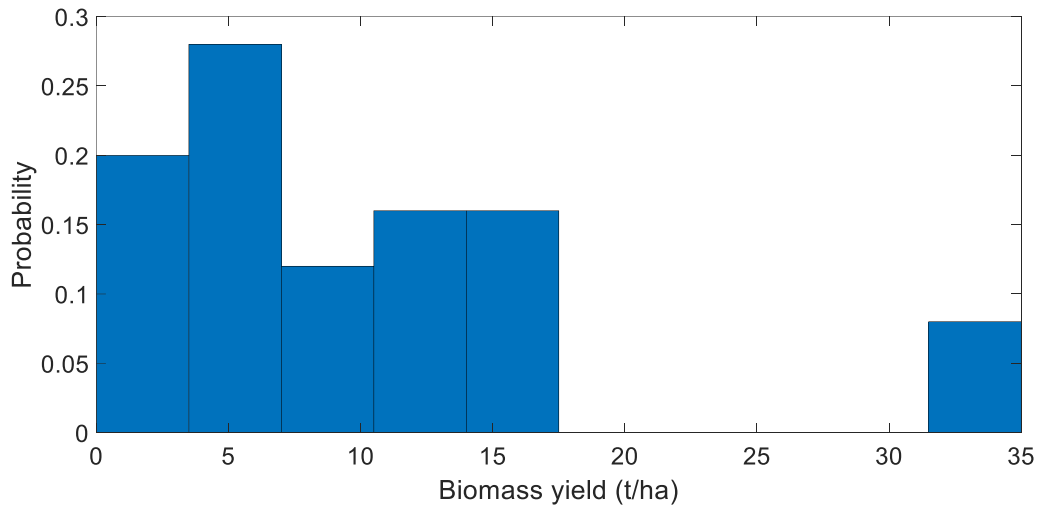


Fig. 21. Normalized histogram of fresh biomass yield

Given the results found in section 4.1, it was decided to consider only the laser measurements. Next, taking the first approach, regression of profile-representative features, the same as in the controlled environment experiment, Table 12 was built with the processing methods described in Table 4. The RMSE values are in t/ha. The best performing method was Lower envelope integral with linear regression. Figure 22 illustrates the best performing regression at this stage.

Table 12. Summary of results from regressions by processing method

Name in MATLAB	Linear			Quadratic			Exponential		
	R^2	R^2_{adj}	RMSE [t/ha]	R^2	R^2_{adj}	RMSE [t/ha]	R^2	R^2_{adj}	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	-0.03	8.67	0.02	-0.07	8.82	0.01	-0.03	8.66
Energy	0.51	0.49	6.10	0.52	0.48	6.19	0.52	0.50	6.07
Variance	0.01	-0.03	8.69	0.01	-0.08	8.88	0.01	-0.03	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	-0.04	8.72	0.02	-0.07	8.82	0.00	-0.04	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	-0.04	8.73	0.00	-0.09	8.92	0.00	-0.04	8.73
Derivative variance	0.00	-0.04	8.72	0.15	0.07	8.25	0.00	-0.04	8.73
Derivative energy	0.00	-0.04	8.73	0.14	0.06	8.29	0.00	-0.04	8.73
Count	0.02	-0.02	8.64	0.24	0.17	7.76	0.01	-0.03	8.67
Peak count	0.00	-0.04	8.72	0.03	-0.06	8.77	0.00	-0.04	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	-0.03	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

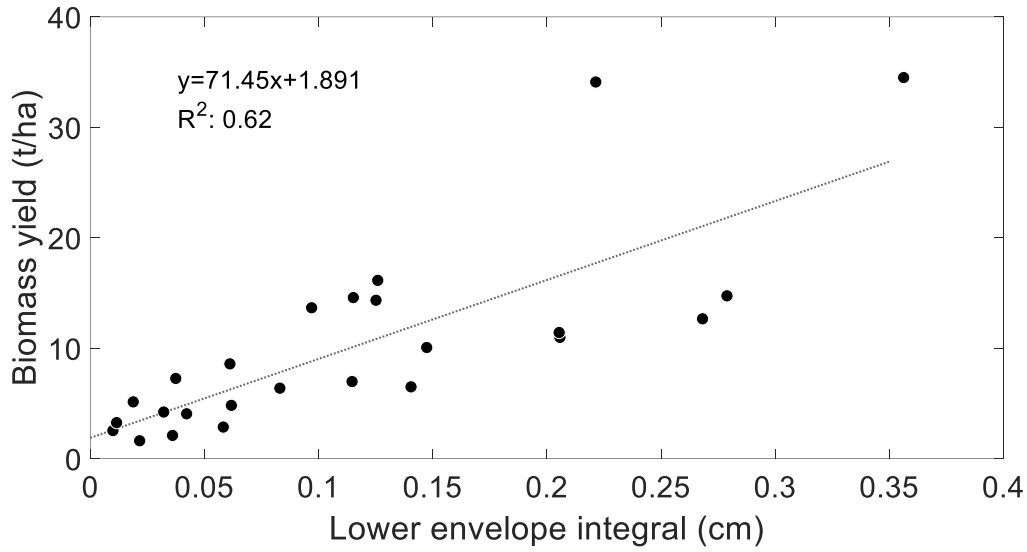


Fig. 22. Linear regression of biomass and lower envelope integral

The resulting RMSE of 5.38 t/ha with that processing method was 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 in order 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 13 and the following equations summarize the most relevant findings.

Table 13. Summary of results from regressions by equation

Equation number	R^2	R^2_{adj}	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

$$f(\mathbf{h}_i^{laser}) = a + b \frac{[Lower\ envelope\ energy]}{[Upper\ envelope\ energy]} + c[Mean\ of\ max] \frac{[Lower\ envelope\ energy]}{[Upper\ envelope\ energy]} \quad (12)$$

$$f(\mathbf{h}_i^{laser}) = a + b[SFDR] + c[Lower\ envelope\ energy]^2 + d[SFDR]^2 \quad (13)$$

$$\begin{aligned} f(\mathbf{h}_i^{laser}) = & a \frac{[Integral]^2}{[Max]^2} + b \frac{[Average\ recorded\ crop\ height]^2}{[Variance]^2} \\ & + c \frac{[Integral]^2 [Average\ recorded\ crop\ height]}{[Max]^2 [Variance]} \\ & + d \frac{[Integral] [Average\ recorded\ crop\ height]^2}{[Max] [Variance]^2} \\ & + e \frac{[Average\ recorded\ crop\ height]^3}{[Variance]^3} \end{aligned} \quad (14)$$

$$\begin{aligned} f(\mathbf{h}_i^{laser}) = & a[Lower\ envelope\ energy] + b[Lower\ envelope\ energy]^2 \\ & + c[Lower\ envelope\ energy]^3 + d[SFDR] + e[SFDR]^2 + f[SFDR]^3 \end{aligned} \quad (15)$$

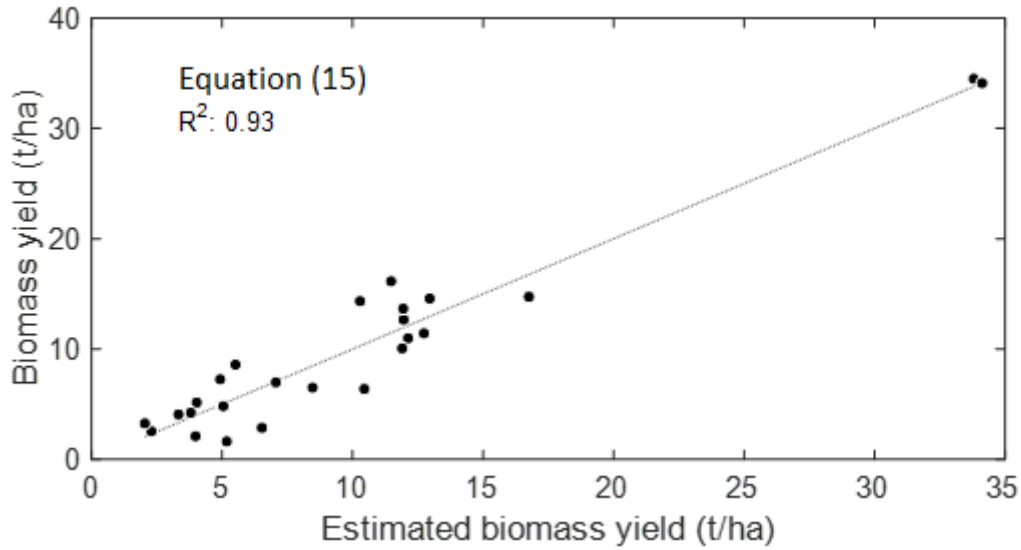


Fig. 23. Best fit achieved with the first approach

The best fitting regression, expressed in Equation (15) and used to build Figure 23, relied on two processing methods that are relatively hard to find physical meaning. Nonetheless, an intuition about what SFDR and Lower envelope energy mean can be provided for 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 could produce a more reliable estimation. In the context of height profiles, low frequency components referred to large-scale sections of the canopy, as opposed to sections of the profile with larger local variability.

Regarding the second approach, inference of a canopy density function, the performance of different density functions is summarized in Table 14. In order to achieve results comparable to those of the first approach, a relatively large number of parameters was required. For example, the best fitting of a density function used 9 parameters, as specified in Equation (16). Figures 24 and 25 illustrate this density function, while Figure 26 shows the fitting of the data using this approach. It was expected that the density at lower parts of the canopy was higher, since the stems are heavier than the leaves. The elements of this approach have a clear physical meaning.

Table 14. Summary of results from density functions

Density function	R^2	R^2_{adj}	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

$$f(H_i(\theta)) = a * \tanh\left(b * \frac{H_i(\theta) + c}{H_i(\theta) + d}\right) \frac{H_i(\theta) - e}{H_i(\theta) + f} + g * \sin(p(H_i(\theta) + l)) \quad (16)$$

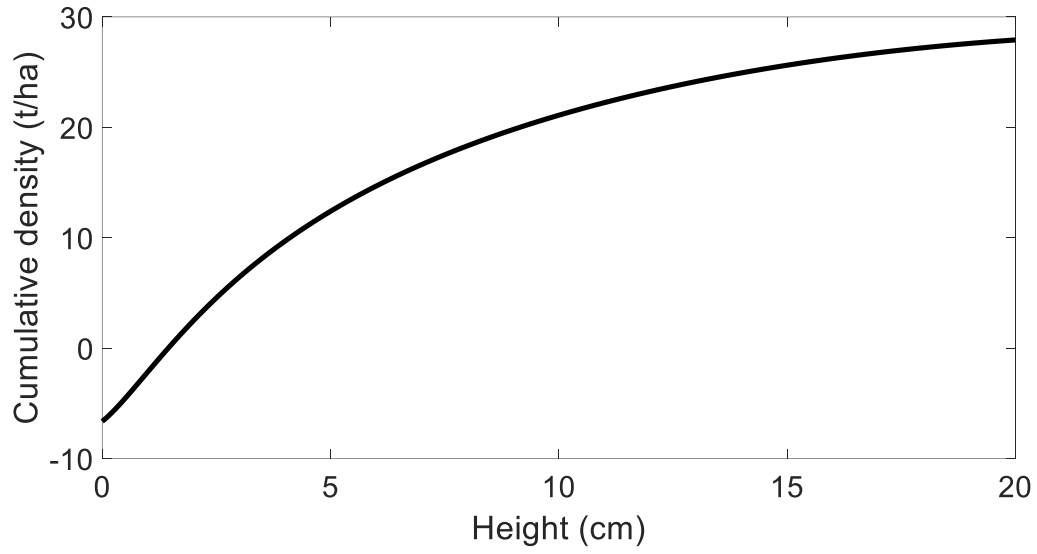


Fig. 24. Best fit for cumulative density function

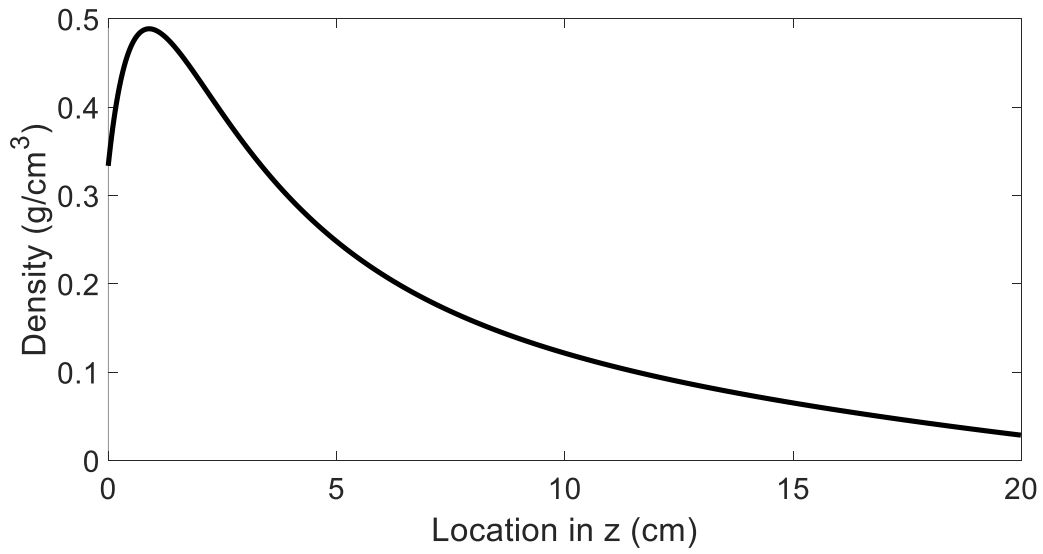


Fig. 25. Best fit for density function

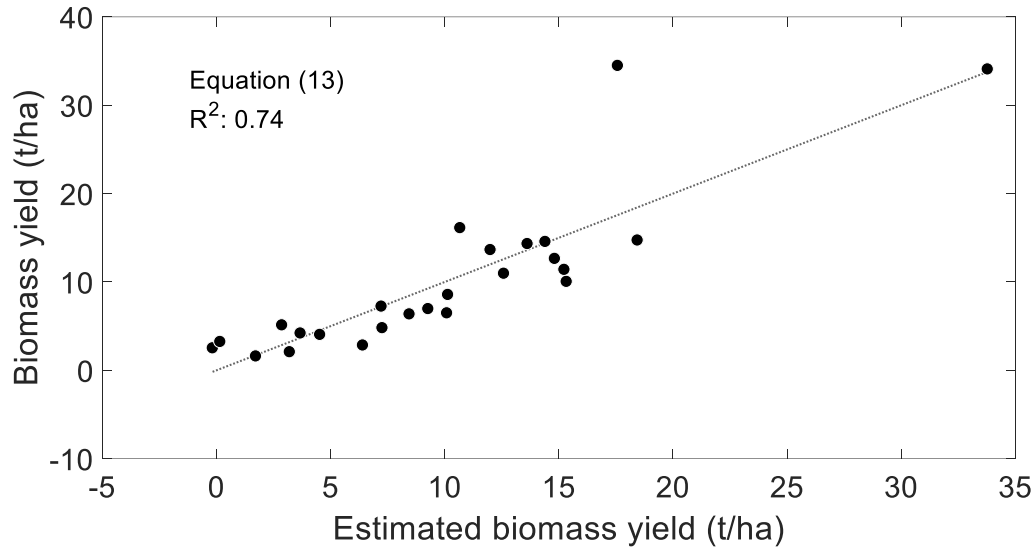


Fig. 26. Best fit achieved with the second approach

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 15 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 seemed 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 was significantly reduced. The case of having a logistic cumulative density function produced the best results in terms of adjusted R^2 , and the RMSE was 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 regarding the determination of the outlier status of the higher yield points. Thus, it was assumed that the higher yield was produced by the natural variability of the crop, and was retained in the final analysis.

Table 15. Summary of results from density functions without higher yield points

Density function	R^2	R^2_{adj}	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 16 compares the best case of both approaches. The percentage error was considerably high for both cases, which was troubling, but was partially explained by the large range covered by the measured biomass. Whatever the case, the coefficient of determination was significant, showing that the laser measurements explained the biomass behavior. There was a possibility that the uncertainty of the manual biomass measurement was also playing a role in this discrepancy. The moisture sensor information remained 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 equation (17).

$$[final\ estimate] = a * [first\ or\ second\ approach] + b * [moisture] + c \quad (17)$$

Table 16. Summary of best results by approach

Approach	R^2	R^2_{adj}	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

5. SUMMARY AND CONCLUSIONS

The model described in Equation (10), when the moisture measurements were added, was found to be the best fit for the controlled environment experiment. The calibration RMSE was around $0.3 \ln(g)$ and $0.4 \ln(g)$ for the estimation of fresh biomass in lettuce and kale, respectively. Similarly, the RMSE was $0.44 \ln(g)$ for dry biomass in both lettuce and kale. The percentage error was between 25.5% and 55.5% for the final calibration. These results proved that the developed device is a viable tool for measuring biomass in an easy and fast way. There were strong indicators that the device is not only more convenient than traditional manual measurement of plant height, i.e. with ruler or measuring tape, but also more accurate in the estimation of biomass.

During the development of this controlled environment experiment, it was found that the laser sensor outperforms the ultrasonic sensor, and the integration of both provides a method to further improve the estimation of biomass, especially when accompanied by moisture measurements. Computing the average crop height, average recorded crop height, and mode proved to be simple yet useful transformations for the signal processing step. This resulted in better fitting than other processing methods. It was also clear that this was dependent on the crop species under consideration.

The production field experiment streamlined the setup and featured a more thorough analysis method. Two different approaches were considered, providing R^2 of 0.94 and 0.78, respectively. These values were comparable to those of the controlled environment experiment, showing that the performance was not diminished in an open field environment. The modifications done on the prototype produced even more convenient handling of the data collection. This would result in faster and more frequent sampling becoming available for the final users.

6. REFERENCES

- Arnó, J., Escolà, A., Vallès, J., Sanz, R., Masip, J., Palacín, J., & Rosell, J. (2009). Use of a ground-based LIDAR scanner to measure leaf area and canopy structure variability of grapevines. *Precision Agriculture '09*.
- Biskup, B., Scharr, H., Schurr, U., & Rascher, U. (2007). A stereo imaging system for measuring structural parameters of plant canopies. *Plant, Cell and Environment*.
- Borgelt, S. C. (1993). Sensing and measurement technologies for site specific management. *Proceedings of Soil Specific Crop Management*. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.
- Buelvas, R., & Adamchuk, V. (2017). Crop canopy measurement using laser and ultrasonic sensing integration. *2017 ASABE Annual International Meeting*. Spokane, WA.
- Catchpol, W. R., & Wheeler, C. J. (1992). Estimating plant biomass: A review of techniques. *Australian Journal of Ecology*.
- Ehlert, D., Heisig, M., & Adamek, R. (2009). Test of a high-end laser rangefinder scanner in agriculture. *Precision Agriculture '09*.
- Ehlert, D., Horn, H.-J., & Adamek, R. (2008). Measuring crop biomass density 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*.
- Elaksher, A. F., Bhandari, S., Carreon-Limones, C. A., & Lauf, R. (2017). Potential of UAV LiDAR systems for geospatial mapping. *Lidar Remote Sensing for Environmental Monitoring 2017*. San Diego.
- Engström, L., Lindén, B., Söderström, M., Börjesson, T., Nissen, K., Lorén, N., . . . Hagner, O. (2009). Determination of canopy properties of winter oilseed rape using remote sensing techniques in field experiments. *Precision Agriculture '09*.
- Escolà, A., Arnó, J., Sanz, R., Camp, F., Masip, J., Solanelles, F., . . . Planas, S. (2009). Sensing tree canopy parameters in real time for precision fructiculture/horticulture applications: methodology set-up and first results. *Precision Agriculture '09*.

- 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*.
- Fricke, T., Richter, F., & Wachendorf, M. (2011). Assessment of forage mass from grassland swards by height measurement using an ultrasonic sensor. *Computers and Electronics in Agriculture*, 142-152.
- Fulton, J. P., Sobolik, C., Shearer, S., Higgins, S., & Burks, T. (2009). Grain yield monitor flow sensor accuracy for simulated varying field slopes. *Applied Engineering in Agriculture*, 15-21.
- 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*.
- Jonasson, S. (1988). Evaluation of the point intercept method for the estimation of plant biomass. *Nordic Society Oikos*, 101-106.
- 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*.
- Moorthy, I., Miller, J. R., Hu, B., Chen, J., & Li, Q. (2008). Retrieving crown leaf area index from an individual tree using ground-based LiDAR data. *Canadian journal of remote sensing*.
- 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*.
- Omasa, K., Hosoi, F., & Konishi, A. (2006). 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany*.
- Plant, R. E. (2001). Site-specific management: the application of information technology to crop production. *Computers and Electronics in Agriculture*, 9-29.
- 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*.
- Tang, L., & Shao, G. (2015). Drone remote sensing for forestry research and practices. *Journal of Forestry Research*.
- Tikasz, P., & Lefsrud, M. (2017). Replacing hydroponic nutrient solution with compost tea made from animal manure. *2017 ASABE Annual International Meeting*. Spokane: ASABE.
- Tilly, N., Aasen, H., & Bareth, G. (2015). Fusion of Plant Height and Vegetation Indices for the estimation of barley biomass. *Remote Sensing*.
- Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., & Peña, J. M. (2015). High-Throughput 3-D Monitoring of Agricultural-Tree Plantations with Unmanned Aerial Vehicle (UAV) Technology. *PLOS ONE*.
- 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*.
- Vansichen, R., & DeBaerdemaeker, J. (1993). A measurement technique for yield mapping of corn silage. *Journal of Agricultural Engineering Research*, 1-10.
- Zaman, Q. U., Schumann, A. W., & Miller, W. M. (2005). Variable rate nitrogen application in Florida citrus based on ultrasonically-sensed tree size. *Applied Engineering in Agriculture*.
- Zaman, Q. U., Schumann, A. W., & Miller, W. M. (2006). Estimation of citrus fruit yield using ultrasonically-sensed tree size. *Applied Engineering in Agriculture*.
- Zarco-Tejada, P. J., Diaz-Varela, R., Angileri, V., & Loudjani, P. (2014). Tree height quantification using very high resolution imagery acquired from an unmanned aerial

vehicle (UAV) and automatic 3D photo-reconstruction methods. *European Journal of Agronomy*.

APPENDICES

APPENDIX A: Python script run in Raspberry Pi

```
#!/usr/bin/env python
```

```
import time
```

```
import serial
```

```
import numbers
```

```
import os.path
```

```
import Adafruit_DHT
```

```
import Adafruit_ADS1x15
```

```
from subprocess import call
```

```
from zaber.serial import BinarySerial, BinaryDevice, BinaryCommand
```

```
def newLogFile():
```

```
    filename='file1.txt'
```

```
    while(os.path.isfile('/home/pi/Laser/'+filename)):
```

```
        filename='file'+str(int(filename[4:-4])+1)+'.txt'
```

```
    return open('/home/pi/Laser/'+filename,'w')
```

```
def readFast adc1, adc2, GAIN, RATE):
```

```
    # - 0 = Channel 0 minus channel 1
```

```
    # - 1 = Channel 0 minus channel 3
```

```
    # - 2 = Channel 1 minus channel 3
```

```
    # - 3 = Channel 2 minus channel 3
```

```
    valueLaser=60-
```

```
0.1*(200*0.00018751*adc1.read_adc_difference(0,gain=GAIN,data_rate=RATE)-600.7)
```

```
valueUltrasonic=2.54*(3.5+68*0.000187*adc2.read_adc_difference(3,gain=GAIN,data_rate=RATE)/4.5)
```

```
valueThermal=100*0.000187*adc2.read_adc_difference(0,gain=GAIN,data_rate=RATE)+  
0.2
```

```
    return [valueLaser,valueUltrasonic,valueThermal]
```

```
serBluetooth=serial.Serial('/dev/ttyAMA0',9600)
```

```
logFile=newLogFile()
```

```
sensor = Adafruit_DHT.DHT22
```

```
pin = 23
```

```
adc1 = Adafruit_ADS1x15.ADS1115()
```

```
adc2 = Adafruit_ADS1x15.ADS1115(address=0x49)
```

```
GAIN = 2/3
```

```
RATE = 475
```

```
speed=73
```

```
direction=True
```

```
full_lap=12800
```

```
updatePosition=BinaryCommand(1,60)
```

```
flag=True
```

```
while(flag):
```

```
    serBluetooth.write('idle\n')
```

```
    message=serBluetooth.readline().strip()
```

```
    print(message)
```

```
    if message=='new':
```

```
        serBluetooth.write(logFile.name)
```

```
        serBluetooth.write(str(os.path.getsize(logFile.name)))
```

```
        logFile.close()
```

```

logFile=newLogFile()
elif message=='test':
    data=[]
    for k in range(10):
        sensorOutput=readFast(adc1,adc2,GAIN,RATE)
        data+= [sensorOutput]
        serBluetooth.write('1:{:.2f} | {:.2f} | {:.2f}\n'.format(sensorOutput[0],
                                                                    sensorOutput[1],
                                                                    sensorOutput[2]))

        time.sleep(0.005)
elif message=='lap':
    serBluetooth.close()
    data=[]
    logFile.write('New Lap\n')
    port=BinarySerial('/dev/ttyUSB0',timeout=200)
    device=BinaryDevice(port, 1)
    lap=False
    initialPosition=device.stop().data
    currentPosition=initialPosition
    while(abs(currentPosition-initialPosition)<=full_lap):
        if(direction):
            device.move_vel(speed)
        else:
            device.move_vel(-1*speed)
        currentPosition=device.send(updatePosition).data
        sensorOutput=readFast(adc1,adc2,GAIN,RATE)
        data+= [sensorOutput+[currentPosition]]
    device.stop()
    port.close()
    serBluetooth=serial.Serial('/dev/ttyAMA0',9600)
    for i in range(5):

```

```

serBluetooth.write('1: {} | {} | {} | {} \n'.format(str(data[-5+i][0]),
                                                    str(data[-5+i][1]),
                                                    str(data[-5+i][2]),
                                                    str(data[-5+i][3])))

serBluetooth.write(str(len(data)))
for point in data:
    logFile.write('1: {}, {}, {}, {} \n'.format(str(point[0]),
                                                str(point[1]),
                                                str(point[2]),
                                                str(point[3])))
serBluetooth.write('finished')
direction=not direction
elif message=='humid':
    for i in range(10):
        s=Adafruit_DHT.read(sensor,pin)
        if isinstance(s[0],numbers.Number):
            valueHumidity=s[0]
            valueTemperature=s[1]
        else:
            valueHumidity=0
            valueTemperature=0
        serBluetooth.write('2: {} | {} \n'.format(str(valueHumidity),
                                                    str(valueTemperature)))
        logFile.write('2: {}, {} \n'.format(str(valueHumidity),
                                                    str(valueTemperature)))
        time.sleep(1)
elif message=='stop':
    if not logFile.closed:
        logFile.close()
    flag=False
elif message=='off':

```

```
serBluetooth.close()
if not logFile.closed:
    logFile.close()
call("sudo nohup shutdown -h now", shell=True)
```

APPENDIX B: MATLAB code for computing Integral in first approach

```
function [ output ] = JIntegral( vector, angles )  
A=vector;  
A(isnan(A))=0;  
output=trapz(pi*angles/180,A);  
end
```

APPENDIX C: MATLAB code for computing Average recorded crop height in first approach

```
function [ output ] = JMean( vector )  
A=vector;  
A (A<0)=0;  
output=nanmean (A) ;  
end
```

APPENDIX D: MATLAB code for computing Max in first approach

```
function [ output ] = JMax( vector )  
output=max(vector);  
end
```


APPENDIX E: MATLAB code for computing Energy in first approach

```
function [ output ] = JEnergy( vector, angles )  
A=vector;  
A(isnan(A))=0;  
output=trapz(pi*angles/180,A.^2);  
end
```

APPENDIX F: MATLAB code for computing Variance in first approach

```
function [ output ] = JVar( vector )  
A=vector;  
A(A<0)=0;  
output=nanvar(A);  
end
```

APPENDIX G: MATLAB code for computing Average trimmed crop height in first approach

```
function [ output ] = JTrimmean( vector )  
output=trimmean(vector,50);  
end
```

APPENDIX H: MATLAB code for computing Trimmed variance in first approach

```
function [ output ] = JTrimvar( vector )
A=vector;
aux=A(A>prctile(A,25));
output=nanvar(aux(aux<prctile(A,75)));
end
```

APPENDIX I: MATLAB code for computing Energy wavelet in first approach

```
function [ output ] = JEnergyW( vector, angles )
A=vector;
A(isnan(A))=0;
n=5;
w='haar';
[c,l]=wavedec(A,n,w);
thr=[max(abs(c(sum(l(1:5))+1:sum(l(1:6))))),max(abs(c(sum(l(1:4))+1:sum(l(1:5))))),0.75*max(abs(c(sum(l(1:3))+1:sum(l(1:4))))),0.5*max(abs(c(sum(l(1:2))+1:sum(l(1:3))))),0.25*max(abs(c(l(1))+1:sum(l(1:2)))))]);
[Af,~,~,~,~]=wdencmp('lvd',c,l,w,n,thr,'s');
output=trapz(pi*angles/180,Af.^2);
end
```

APPENDIX J: MATLAB code for computing Lower envelope in first approach

```
function [ output ] = JLmean( vector )  
[~,aux]=envelope(vector,10,'peaks');  
output=nanmean(aux);  
end
```

APPENDIX K: MATLAB code for computing Lower envelope integral in first approach

```
function [ output ] = JLintegral( vector, angles )  
[~,aux]=envelope(vector,10,'peaks');  
output=trapz(pi*angles/180,aux);  
end
```

APPENDIX L: MATLAB code for computing Lower envelope energy in first approach

```
function [ output ] = JLeenergy( vector, angles )  
[~,aux]=envelope(vector,10,'peaks');  
output=trapz(pi*angles/180,aux.^2);  
end
```


APPENDIX M: MATLAB code for computing Upper envelope in first approach

```
function [ output ] = JUmean( vector )  
[aux,~]=envelope(vector,10,'peaks');  
output=nanmean(aux);  
end
```

APPENDIX N: MATLAB code for computing Upper envelope energy in first approach

```
function [ output ] = JUenergy( vector, angles )  
[aux,~]=envelope(vector,10, 'peaks');  
output=trapz(pi*angles/180, aux.^2);  
end
```

APPENDIX O: MATLAB code for computing Derivative in first approach

```
function [ output ] = JDmean( vector, angles )  
derivative=diff(vector)./diff(pi*angles/180);  
aux=derivative(derivative>prctile(derivative,25));  
output=nanmean(aux(aux<prctile(derivative,75)));  
end
```

APPENDIX P: MATLAB code for computing Derivative variance in first approach

```
function [ output ] = JDvar( vector, angles )
derivative=diff(vector)./diff(pi*angles/180);
aux=derivative(derivative>prctile(derivative,25));
output=nanvar(aux(aux<prctile(derivative,75)));
end
```

APPENDIX Q: MATLAB code for computing Derivative energy in first approach

```
function [ output ] = JDenergy( vector, angles )
derivative=diff(vector)./diff(pi*angles/180);
thr=prctile(derivative,[25,75]);
derivative(derivative<thr(1))=0;
derivative(derivative>thr(2))=0;
derivative(isnan(derivative))=0;
output=trapz(pi*angles(1:end-1)/180,derivative.^2);
end
```

APPENDIX R: MATLAB code for computing Count in first approach

```
function [ output ] = JCount( vector )  
aux=findchangepts(vector,'MinThreshold',1);  
output=length(aux);  
end
```

APPENDIX S: MATLAB code for computing Peak count in first approach

```
function [ output ] = JPcount( vector )  
aux=findpeaks(vector);  
output=length(aux);  
end
```

APPENDIX T: MATLAB code for computing Frequency in first approach

```
function [ output ] = JFreq( vector )  
output=medfreq(vector);  
end
```


APPENDIX U: MATLAB code for computing Bandwidth in first approach

```
function [ output ] = JPBW( vector )  
output=powerbw(vector);  
end
```

APPENDIX V: MATLAB code for computing SFDR in first approach

```
function [ output ] = JSFDR( vector )  
output=sfdr(vector);  
end
```

APPENDIX X: MATLAB code for computing SNR in first approach

```
function [ output ] = JSnr( vector )  
output=snr(vector);  
end
```

APPENDIX Y: MATLAB code for computing THD in first approach

```
function [ output ] = JTHD( vector )  
output=thd(vector);  
end
```

APPENDIX Z: MATLAB code for computing Mean of max in first approach

```
function [ output ] = JMeanmax( vector )
n=floor(length(vector)/10);
aux=zeros(n,1);
for i=1:n
    aux(i)=max(vector(10*(i-1)+1:10*i));
end
output=mean(aux);
end
```

APPENDIX AA: MATLAB code for computing Pseudo-max in first approach

```
function [ output ] = JPsmax( vector )  
output=prctile(vector,95);  
end
```

APPENDIX AB: MATLAB code defining error function for second approach

```
function [ e ] = errorD( mv, mystruct, params )
a=params(1);
b=params(2);
c=params(3);
d=params(4);
e=params(5);
f=params(6);
g=params(7);
h=params(8);
l=params(9);
positions=fieldnames(mystruct);
mestimate=zeros(length(positions)-2,1);
for i=2:length(positions)-1
   iaux=zeros(1,size(mystruct.(positions{i}),2));
    for j=1:size(mystruct.(positions{i}),2)
        xx=mystruct.(positions{i})(:,j);

maux(j)=trapz(pi*mystruct.x/180,a*tanh(b*(xx+f)./(xx+h)).*(xx
+c)./(xx+d)+e*sin(l*(xx+g)))/(2*pi);
    end
    mestimate(i-1)=nanmean(iaux,2);
end
e=rms(mv-mestimate);
end
```

APPENDIX AC: MATLAB script in second approach

```
clc;
clear all;
load('dataj2.mat');
load('shortForTrain.mat');
positions=fieldnames(JData);
ai=91.6872;
bi=0.4049;
ci=-1.5183;
di=2.822;
ei=6.1028;
fi=1.9798;
gi=0.1399;
hi=3.1413;
li=0.1;
[coef,fval]=fminsearch(@(u)
errorD(dd,JData,u),[ai,bi,ci,di,ei,fi,gi,hi,li]);
rmse=fval;
a=coef(1);
b=coef(2);
c=coef(3);
d=coef(4);
e=coef(5);
f=coef(6);
g=coef(7);
h=coef(8);
l=coef(9);
mestimate=zeros(length(positions)-2,1);
for i=2:length(positions)-1
    mau=zeros(1,size(JData.(positions{i}),2));
    for j=1:size(JData.(positions{i}),2)
```



```

xx=JData.(positions{i})(:,j);

maux(j)=trapz(pi*JData.x/180,a*tanh(b*(xx+f)/(xx+h)).*(xx+c)
./(xx+d)+e*sin(l*(xx+g)))/(2*pi);
    end
    mestimate(i-1)=nanmean(maux,2);
end
r2=1-sum((dd-mestimate).^2)/sum((dd-mean(dd)).^2);

```