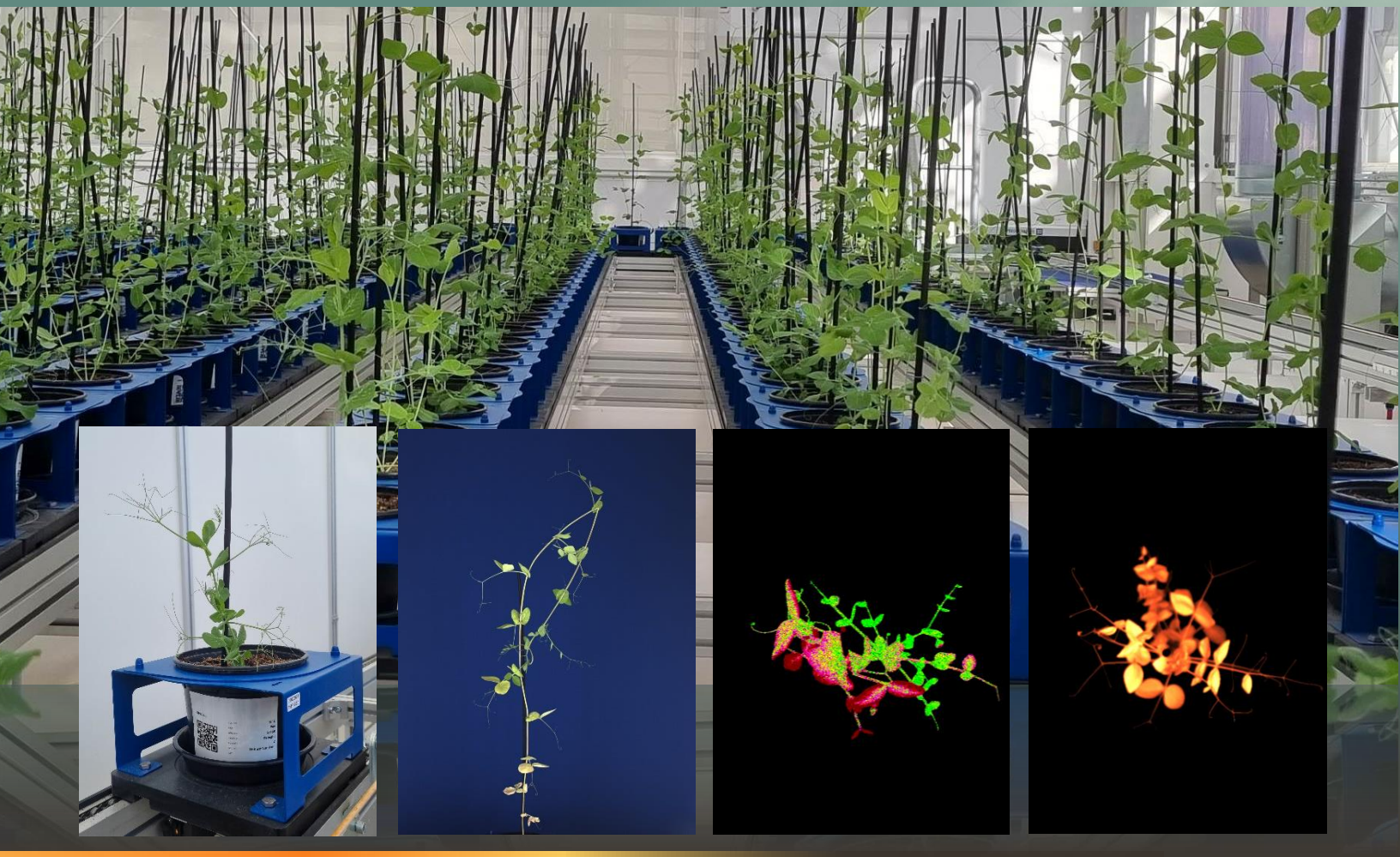
MSc Research Practice Crop Physiology

Centre for Crop Systems Analysis (CSA)

Netherlands Plant Eco-phenotyping Centre (NPEC)



Improving Non-invasive Biomass Estimation by Combining Different Types of Image-Based Phenotypic Traits:

A Study in Pea under Drought Stress

Narawitch Lertngim

Supervisors: Steven Driever & Rick van de Zedde

WAGENINGEN
UNIVERSITY & RESEARCH

Research topic:	Improving non-invasive biomass estimation by combining different types of image-based phenotypic traits: a study in pea under drought stress.
Credits:	24
Student name:	Narawitch Lertngim
Student number:	1052752
Supervisors:	Steven Driever
	Rick van de Zedde
Study Program:	Master of plant sciences
Specialization:	Crop sciences
Chair groups:	Centre for Crop Systems Analysis (CSA)
	Lab of Genetics/ Netherlands Plant Eco-phenotyping Centre (NPEC)
Research duration:	20th March to 7th July 2023 (16 weeks)

Abstract

Peas (*Pisum sativum L.*) are valuable protein sources for both human food and livestock feed, while also contributing to soil fertility. However, the higher frequency of global droughts has negatively impacted pea yield and protein content. To address this, breeding programs must prioritize enhancing pea adaptation and resilience to drought stress. In this regard, image-based high-throughput phenotyping platforms offer an efficient solution for screening large populations, enabling non-destructive acquisition of numerous phenotypic traits. Multimodal camera systems, such as (i) RGB imaging for architecture and morphology, (ii) spectral imaging for pigment contents, and (iii) chlorophyll fluorescence imaging for photosynthesis, provide diverse image-based phenotypic data. Such types of data have proven valuable for developing biomass estimation models, crucial for monitoring crop growth and yield predictions. However, existing studies predominantly focus on using a single type of camera system for biomass estimation. Since different camera systems detect distinct types of plant characteristics, the ability to combine different types of phenotypic traits could be a potential approach to enhance biomass estimation models. To explore this, a greenhouse experiment was conducted to evaluate 180 pea genotypes responses to drought using multimodal image-based phenotyping. Subsequently, biomass estimation models were developed using single type and multiple types of traits. The results indicated that pea plants exhibited greater biomass under control than drought conditions, consistent with digital biomass derived from RGB imaging. However, unexpected outcomes were observed in color, reflectance, and photosynthesis traits, potentially influenced by the plant's growth stage during measurement and the experimental water regimes. Combining different types of image-based phenotypic traits in multiple trait models yielded slight improvements in biomass estimation model accuracy (R^2 = 0.79, RMSE = 0.73) compared to models using a single trait type (R^2 = 0.76, RMSE = 0.68). This study highlights the limitations as well as the potential of combining different traits to enhance biomass estimation accuracy and discusses the approach for advancing plant phenotyping capabilities.

Keywords: biomass estimation / chlorophyll fluorescence / high-throughput phenotyping / image-based phenotypic traits / spectral imaging / pea

Acknowledgements

This research was completed with excellent support from various individuals. Firstly, I would like to express my deep gratitude to my research practice advisors, Steven Driever and Rick van de Zedde, who consistently provided me with valuable guidance. They offered me opportunities to learn and improve my hard and soft skills, dedicated their valuable time to track my progress, and encouraged me to conduct research from multiple perspectives and using various techniques. I firmly believe that the experiences, knowledge, and skills gained during this research practice will greatly benefit my future research career.

I would also like to extend my sincere appreciation to the researchers at the Netherlands Plant Eco-phenotyping Centre (NPEC) for their assistance in teaching, guiding, and advising me on high-throughput phenotyping facilities, image acquisition, data extraction, and data analysis. Moreover, their valuable comments and feedback greatly contributed to the enhancement of this research. I am grateful to Maryam Bagheri from Universidade Católica Portuguesa (UCP) for allowing me to participate in her experiment and providing invaluable information regarding experimental design and data analysis. Collaborating with NPEC and UCP was a remarkable and memorable experience. I would like to acknowledge my colleagues at the Centre for Crop Systems Analysis (CSA) for their kindness support during my research practice period. Furthermore, I am deeply grateful for the love, support, and encouragement from my family and friends.

Lastly, I would like to express my heartfelt appreciation to the National Center for Genetic Engineering and Biotechnology (BIOTEC), Thailand, for granting me the opportunity to pursue my studies and conduct research. Their financial support throughout my entire master's degree program, through the Royal Thai Government Scholarship Program offered by OCSC, has been invaluable.

Table of Contents

Chapter 1: Introduction

1.1 Impact of drought in pea cultivation

Pea (*Pisum sativum* L.) is a multipurpose crop and is ranked as the fourth most cultivated pulse crop in the world (Pandey et al., 2021). Peas are an important source of protein and have been used as livestock feed and human food for centuries (Rubiales et al., 2015). In addition, peas improve soil fertility and play a role in the nitrogen cycle by fixing atmospheric nitrogen to the soil through symbiotic interactions with soil-borne bacteria (Poore et al., 2018). In the current context of climate change, unpredictable weather patterns, reduced precipitation, and fluctuations in rainfall have led to more frequent droughts worldwide (Fahad et al., 2017; Jaghdani et al., 2021). As a result, pea yield and protein content are unstable, and the cultivated area of pea has decreased (Cernay et al., 2015). Under drought conditions, plants generally respond by closing their stomata to reduce water loss, which can decrease photosynthesis and lead to a reduction in biomass accumulation and yield production (Chaves et al., 2003; Xu et al., 2010). Therefore, it is important for breeding programs to improve pea adaptation and resilience to drought stress (Rubiales et al., 2015). Hence, the ability to identify these traits in large quantities of phenotypic data is required to keep up genomic data (Reynolds et al., 2019). High-throughput phenotyping platforms could be efficient tools for screening large populations, which would increase the acquisition of phenotypic data, reduce the gap between genomic and phenomic data and become helpful tools for breeding programs (Crain et al., 2018).

1.2 High-throughput phenotyping

A high-throughput phenotyping platform has great potential in improving the accuracy, efficiency, speed, and quality of determining crop growth and development compared to conventional phenotyping (Pabuayon et al., 2019). Additionally, it enables the acquisition of numerous phenotypic traits using digital and automated methods. Image-based high-throughput phenotyping assists in the computation of phenotypes by analyzing a large number of plants in a short time interval (Das Choudhury et al., 2016; Das Choudhury et al., 2018). Furthermore, the images provide spatial variation of information at the whole-plant level, which provides data that is difficult to measure using conventional phenotyping. The progress of current analytical techniques

has significantly contributed to high-throughput plant phenotyping (Campbell et al., 2018; Yang et al., 2020; Singh et al., 2021; Song et al., 2021). Image-based phenotyping, captured by multimodal cameras providing different valuable information of plants mechanisms. There are three types of camera systems in this experiment and often used in similar experiments elsewhere in image-based phenotyping, such as (i) RGB sideview imaging, (ii) spectral imaging (SI), and (iii) chlorophyll fluorescence (CF) imaging (Figure 1.1).

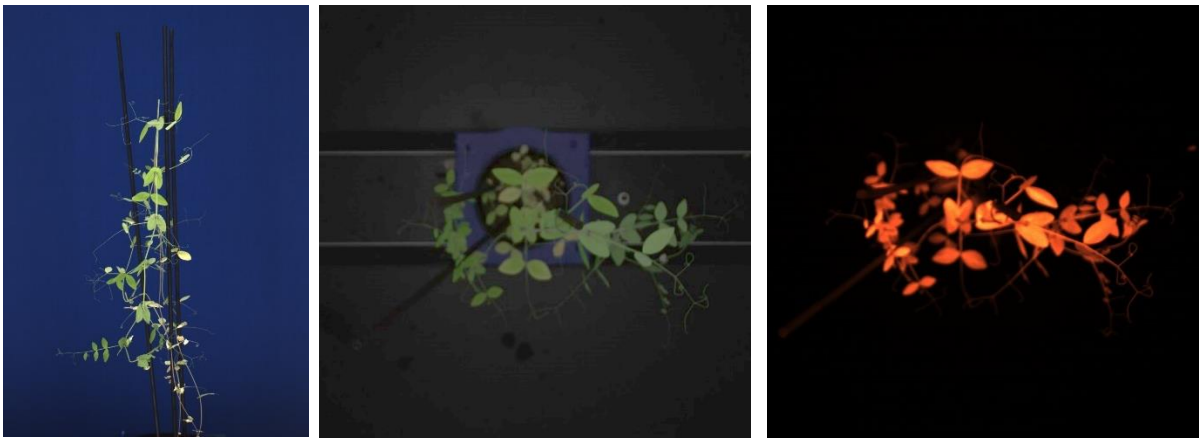


Figure 1.1 The images of pea (genotype *Sp094*) captured using different camera systems: RGB sideview imaging (left); spectral imaging (middle); chlorophyll fluorescence imaging (right).

RGB sideview imaging, the first type of image-based phenotyping, provides information on plant's shape and area. The extracted traits from these images can be linked to several conventional plant traits such as morphological and architectural traits. The number of plant pixels in an image can be interpreted as plant biomass, often multiple views of a plant are captured by a turn-table. Moreover, traits extracted from RGB sideview imaging can be linked to shape and density of plants, which is useful for monitoring plant growth and development (Anandan et al., 2020; Kim et al., 2020).

The second type is spectral imaging (SI), a technique that combines spectroscopy with digital imaging to scan multiple spectra and analyze discrete bands of wavelengths within electromagnetic spectrum. It provides diverse information on pigments (e.g. chlorophyll, carotenoids), water, nutrients, proteins, and carbohydrates, allowing for a more comprehensive understanding of plant

physiology and biochemistry. Spectral indices derived from different wavelengths can be used to study plant responses to stressors and estimate foliar traits and biomass. While SI provides chemical and physiological insights, it does not necessarily provide specific details on plant mechanisms (Polder et al., 2021; Michelon et al., 2023; Cotrozzi & Couture, 2020; Kim et al., 2011; Couture et al., 2013; Serbin et al., 2015; Kumar et al., 2021).

The third type is chlorophyll fluorescence (CF) imaging, which focuses on the functioning of photosystem II (PSII) and measures the light emitted by green plant tissue when illuminated with specific wavelengths. It provides information about photosynthesis and the photosynthetic activity of PSII. Additionally, CF imaging is effective in assessing plant stress and has been used to determine photosynthetic traits and detect early stress symptoms in high-throughput phenotyping (Baker, 2008; Porcar-Castell et al., 2014; Maxwell & Johnson, 2000; Murchie & Lawson, 2013; Kumar et al., 2021; Kalaji et al., 2017; Chaerle et al., 2003).

1.3 Image-based biomass estimation model

Image-based phenotypic traits can also determine plant biomass without the need for a destructive harvest (Furbank et al., 2011; Neumann et al., 2015). This method is widely used in agriculture, and both linear and non-linear models are utilized to estimate biomass (Tackenberg, 2007; Golzarian et al., 2011; Rahaman et al., 2017). Linear models based on plant area are more effective than non-linear models, but their estimation error is too large for accurate biomass estimation (Leister et al., 1999; Golzarian et al., 2011). Apart from plant area derived from 2D RGB sideview imaging, there are other significant traits that also affect plant growth and biomass should be taken into consideration. Rahaman et al. (2017) proposed the linear biomass model as a function of plant area, plant compactness, and plant age. The results of that study confirmed that including more significant traits can improve biomass estimation accuracy of the model. However, the traits used in this study derived from the same camera systems which is 2D RGB sideview imaging. The interesting question is if the combination of other traits from different types of camera systems improves biomass estimation. Therefore, an alternative method for estimating biomass that considers other traits in addition to RGB sideview imaging is needed.

1.4 The combination of multiple traits for phenotyping improvement

RGB sideview imaging, spectral imaging traits (SI), and chlorophyll fluorescence (CF) imaging have been proved their effectiveness for phenotyping individually (Li et al., 2010; Feng et al., 2018). RGB sideview imaging has a strong relationship with plant architecture and morphology. SI information has an advantage with regard to monitoring general compositions of various plants mechanisms and nutrient contents, while CF has a strong relationship with photosynthesis, which are both related to biomass production. However, there are a few studies that focus on the combination of these different types of imaging using for plant phenotyping and biomass estimation. Since each type of imaging reflects different mechanisms of plants, the combination of different types of imaging could increase potential of phenotyping (Römer et al., 2012; Zhang et al., 202). Some studies have attempted to test the performances of combination of data acquired by different types of imaging. For example, at the agricultural-field level, the fusion of Light Detection and Ranging (LiDAR), a 3D shape analysis technique, and spectral data showed the highest biomass estimation accuracy of maize (Wang et al., 2016). Another study on *Fusarium* head blight of wheat indicated that the combination of SI and CF is more suitable for disease detection than using single data type methods (Mahlein et al., 2019). Instead of considering only traits derived from RGB imaging, it is possible that the combination of RGB imaging with either SI or CF or both could be a potential approach to improve the accuracy of biomass estimation, which could increase capability of plant phenotyping.

1.5 Research aims and research questions

This study aims to assess whether combination of RBG sideview imaging with spectral imaging traits (SI) and chlorophyll fluorescence (CF) improves the accuracy of biomass estimation models for pea plants under drought conditions. The hypothesis is that combining traits from RGB sideview imaging with either traits from SI or CF or both is expected to enhance biomass estimation accuracy. Each type of trait captures different plant characteristics related to biomass accumulation. RGB sideview imaging directly reflects morphology and architecture for plants. It is expected to have a high correlation with harvested biomass and contribute most to biomass estimation accuracy. SI captures reflectance from various plant nutrients and pigments, while CF captures fluorescence emitted from photosynthetic activity. Hence, these two types of imaging are

also expected to have correlations with harvested biomass, but not higher than traits from RGB sideview imaging. Combining all possible traits from RGB sideview imaging, SI, and CF is expected to yield the highest accuracy in biomass estimation. On the other hand, combining RGB sideview imaging with either SI or CF is expected to result in a lower biomass estimation accuracy, although it is still higher than using only RGB sideview imaging. This study will therefore investigate one main research question and two sub-research questions, which are:

Main research question:

1. Can the combination of RGB sideview traits with spectral topview imaging traits, and chlorophyll fluorescence traits improve the accuracy of biomass estimation model of pea under drought stress?

Sub research questions:

1.1 Is there a relationship among harvested biomass, RGB sideview traits, spectral topview imaging traits, and chlorophyll fluorescence traits in pea under drought stress?

1.2 Which image-based phenotypic traits is contributing most to the highest accuracy of biomass estimation model?

Chapter 2: Materials and methods

This research consists of four parts. First, a greenhouse experiment was conducted with 180 pea genotypes at two different levels of water supply. Second, all plants were measured for image-based traits using different camera systems, and the plants were destructively harvested for biomass traits. Third, imaging data was extracted, and the data was then used for the data analysis. The mean differences in each trait between plants grown in control and drought conditions was investigated. Additionally, the correlation between image-based traits and the biomass of the destructive harvest was determined. Fourth, biomass estimation models were formulated and evaluated. The conceptual research framework is shown in Figure 2.1, and the details of each method are described in the following sections.

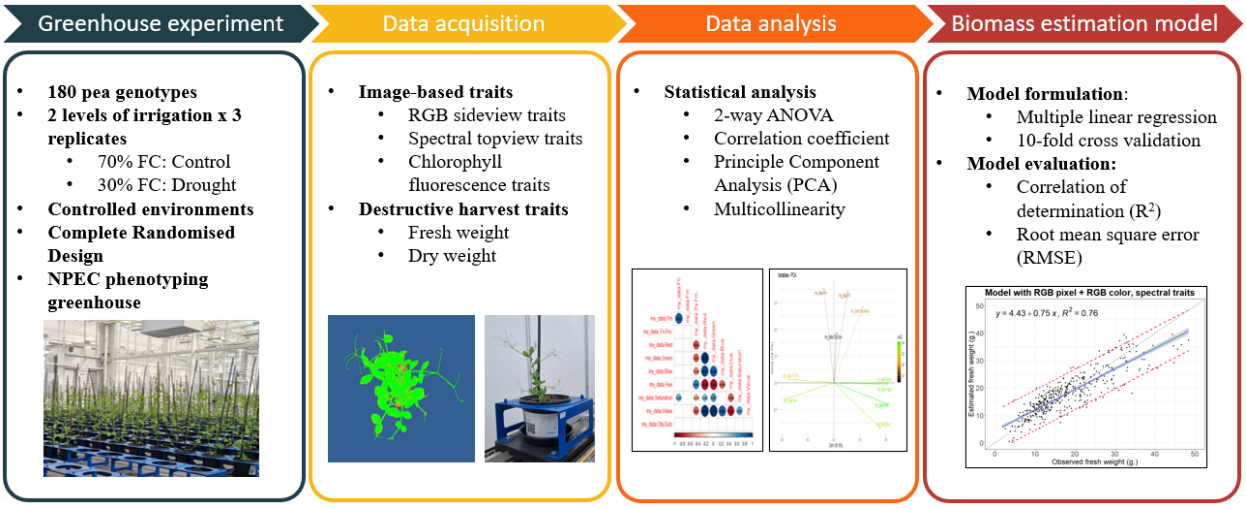


Figure 2.1 Research framework summarized in the approach. Short description of methodology. Abbreviations: FC: Field Capacity, NPEC: the Netherlands Plant Eco-phenotyping Centre.

2.1 Experimental design

A pot experiment was conducted in an environmentally controlled greenhouse (the Netherlands Plant Eco-phenotyping Centre (NPEC) greenhouse, Wageningen University and Research, the Netherlands). There were two factors in this study, which were pea genotypes and water irrigation.

Pea genotypes selected for this experiment based on their high level of protein in pea seeds were 180 genotypes in total. Two treatments of water irrigation were applied: control and drought. Since each treatment has 3 replications, there were 1,080 experimental units in this experiment. Two greenhouse compartments were used for the experiment and each compartment contained 540 pots. Each pot was randomly placed on the conveyer in the greenhouse compartments as completely randomized design. The greenhouse experiment was carried out from 6th March to 24th April 2023.

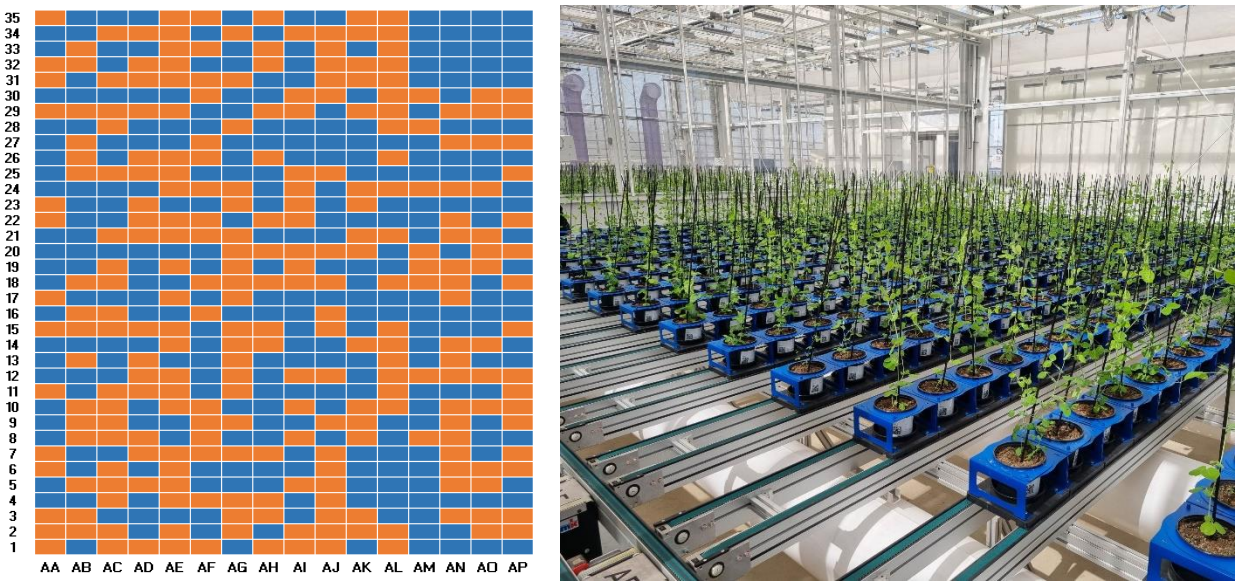


Figure 2.2 The experiment with completely randomized design (left) showed how the pots were placed each rectangular grid represent each pot; pea plant grown under control (blue) and drought (orange) condition; the practical pot arrangement (right).

The environment in the greenhouse was controlled with an average temperature of 22^{o}C during day and 18 ^{o}C at night, relative humidity was set at 60%. Natural light was used during the experiment with additional lighting supplied. No additional carbon dioxide was applied during the experiment. No nutrient solution was applied during the experiment. Each plant was grown in a pot of 17 cm diameter and 13 cm height filled with potting soil. During the experiment, plants were weighted every other day throughout the experiment to determine the amount of irrigation applied if needed. One to three sticks were put in the pot in order to support the pea plant structure and prevent them to grow out of the pot. Irrigation was applied when the amount of water was

lower than 70% of field capacity (FC) until the drought treatment started after three weeks of sowing. Plants were regularly irrigated to keep approximately 70% and 30% FC for control and drought conditions, respectively. Drought treatment was applied for three weeks then plants were imaged by sideview and topview camera systems. After that, plants were destructively harvested for the fresh weight and then were put in a force dry oven at 105^{o}C oven for at least 48 hours. After that, dry weight of the plant biomass was measured.

2.2 Phenotyping measurements and image processing

Phenotypic data was automatically collected using the conveyor system at NPEC greenhouse, providing high-throughput phenotyping conditions. Plants were measured with the following imaging: (i) RGB sideview imaging; (ii) spectral imaging (SI); (iii) chlorophyll fluorescence (CF) imaging. The measurement was taken once before harvest. The sideview whole plant color imaging was captured using a RGB sideview scan camera (1” Progressive-scan-CCD sensor with bayer filter; 3384 (H) x 2704 (V) pixel with 3.69 µm pixel size). Plants entering the RGB sideview cabinet were stopped on the turn-table. The rotation of the table was set at 6 times or every 60^{o} (0 ^{o}, 60 ^{o}, 120 ^{o}, 180 ^{o}, 240 ^{o}, and 300 ^{o}). Hence, six images were captured for each plant. Each image was masked for plant part and analysed plant parameters, such as digital biomass, convex hull, height, and solidity (Figure 2.3). The analysis of RGB sideview imaging was provided by the NPEC team, color information was not extracted from these sideview images.

Sideview digital biomass was calculated as the average of six sideview projected areas of plant parts, while height was measured using the pixel height of the plants. The convex hull, which is the smallest area that encloses the entire plant, provides information about the overall shape and size of plant structures. Solidity is the ratio of digital biomass to the convex hull area and represents the compactness of plant structures. The value of solidity ranges from 0 to 1, with higher values indicating a more compact or solid shape (Figure 2.4).

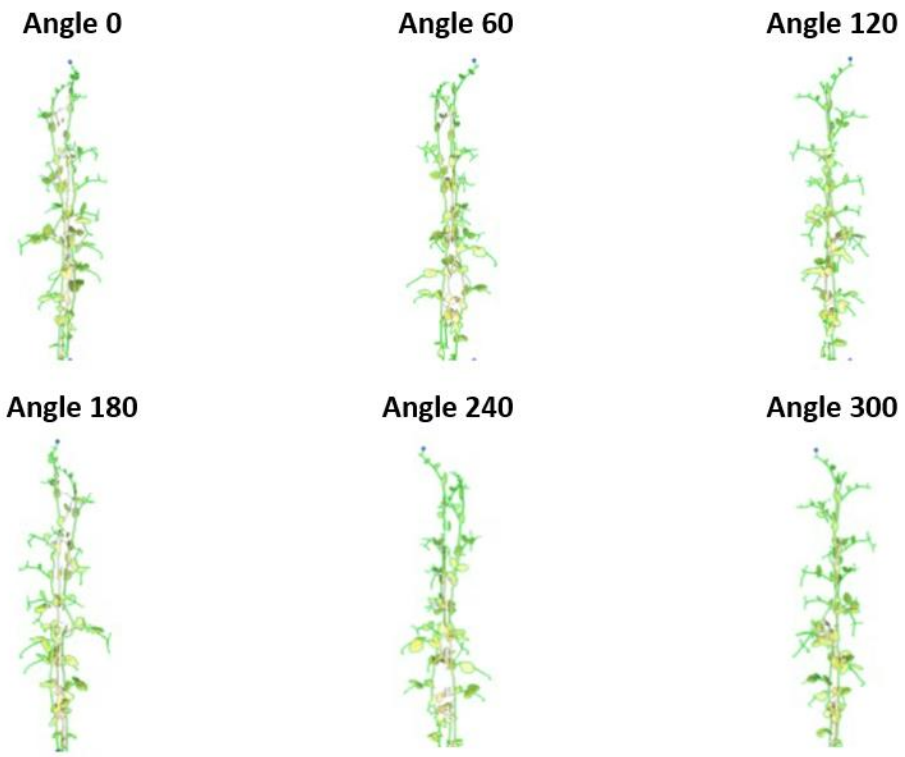


Figure 2.3 Example of an RGB sideview image with different 6 angles of the same pea (Genotype *Sp157* in control condition) for which digital biomass, height, convex hull, and solidity was calculated including the plant pixel area.

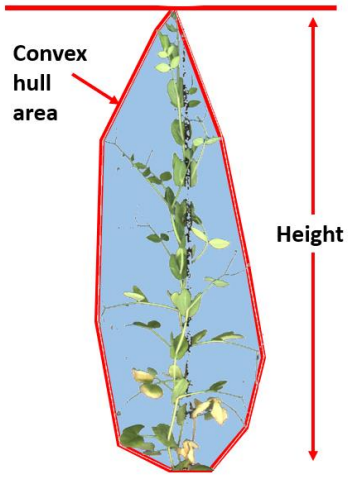


Figure 2.4 Pea plant with overlaid representative measured sideview traits.

The topview SI and CF imaging were captured using CropReporter camera systems (PhenoVation B.V., Wageningen, The Netherlands). The CropReporter^{TM} consists of cabinet with a camera system that houses controller computer, charge-coupled device (CCD) camera with optical filter wheel and focusing unit, integrated high-intensity red light-emitting diodes (LEDs) for excitation of the photosynthesis. All images are captured with the same lens (10 Mp 8 mm lens, 200 Lp mm-1 resolution, 400 – 1000 nm spectral range) and CCD camera (1.3 Mp, 1296 x 996 pixels), with real 14-bit signal resolution. Plants were imaged at approximately 140 cm distance from the camera. The output is 16-bit RAW format, and automatic analysis of RGB, SI, and CF imaging was performed by DA^{TM} software (PhenoVation B.V., Wageningen, The Netherlands) (Lazarević et al., 2021).

Before CF measurement, plants were dark-adapted for 30 minutes. For the excitation of photosynthesis, 4,000 $\mu molm^{-2}s^{-1}$ red LED light was used. The integration time for capturing the chlorophyll fluorescence image was 200 µs. The minimum chlorophyll fluorescence (F_0) and maximum chlorophyll fluorescence (F_m) images were captured after 10 µs and 800 ms, respectively (Lazarević et al., 2021). Only dark-adapted measurement was taken in this experiment. Measured F_0 and F_m was used for calculation of the maximum quantum yield for PSII (F_v/F_m) (Genty et al., 1989). Then plants were again illuminated for SI. Spectral images were captured at reflectance of colors: red (R_{Red}-640 nm), green (R_{Green}-550 nm), and blue (R_{Blue}-475 nm), providing RGB topview imaging. These spectral reflectance in visible region indicates alterations in pigment content. Then hue (HUE), saturation (SAT), and value (VAL) were automatically calculated by Data Analyzer (DA) software (PhenoVation B.V., Wageningen, The Netherlands) (Lazarević et al., 2021).

HSV color spaces (Hue, Saturation, and Value) is defined as a circular cylinder with the HUE as the angle around the cylinder (with red being 0°, green 120° and blue 240°), SAT as the distance from the center line of the cylinder and VAL as the height of the point of base of cylinder (FutureLearn, 2023; Figure 2.5). In plant phenotyping, the HUE component of the HSV color system can be used to detect and classify different color variations in plants. On the other hand, SAT refers to the intensity or purity of a color, providing insights into the intensity or saturation of pigments in plant tissues. Additionally, VAL represents the brightness of a color and can be utilized to assess the light reflectance properties of plant surfaces (Lazarević et al., 2022).

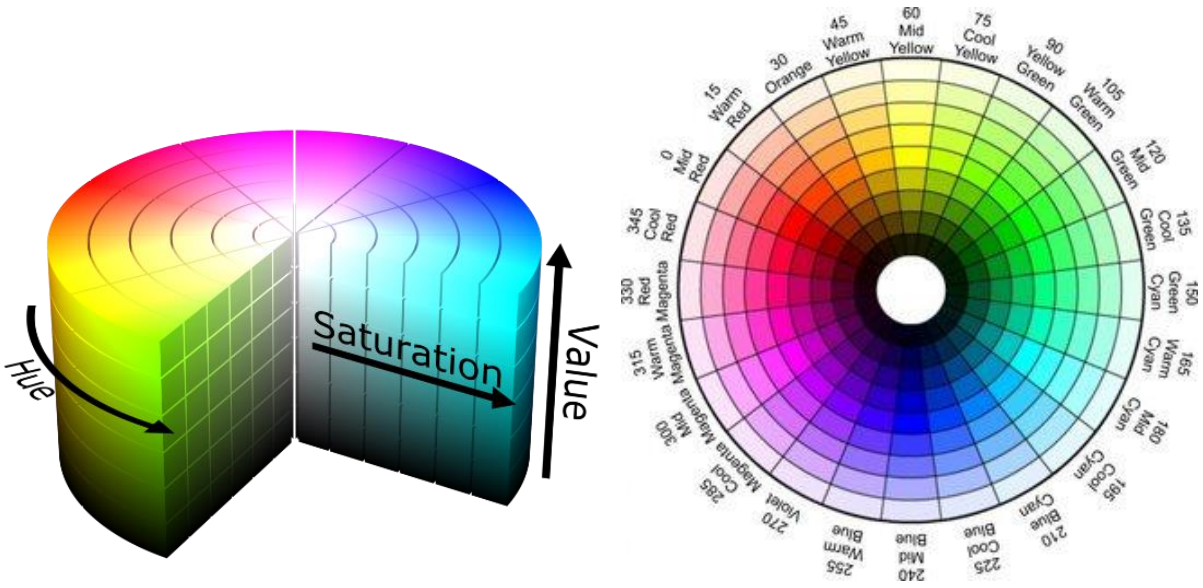


Figure 2.5 The HSV (Hue, Saturation, and Value) color model mapped to a cylinder (Source: https://commons.wikimedia.org/wiki/File:HSV_color_solid_cylinder.png).

For data extraction from images captured by CropReporter, CF was set at minimum 1500 and minimum object size was 10 pixels in order to mask plant area, remove sticks, and remove noise from the images.

All measured and calculated traits from both RGB sideview camera system and the CropReporter were categorized into three types; (i) RGB sideview traits; (ii) topview SI traits; (iii) CF traits (Table 2.1).

Table 2.1. List of traits derived from RGB sideview imaging and CropReporter camera systems categorized by types of imaging with abbreviations, wavelength for measurement or equation for calculation, and the reference if appropriate.

Abbreviation	Trait	Unit	Camera system
RGB sideview traits			
DB_Top	Topview digital biomass	cm^2	CropReporter
DB_Side	Sideview digital biomass	cm^2	RGB Sideview
CH	Sideview convex hull area	cm^2	RGB Sideview
Height	Height	cm	RGB Sideview
Solidity	Sideview solidity	-	RGB Sideview
Spectral imaging traits			
HUE	Hue	degree	CropReporter
VAL	Value	-	CropReporter
SAT	Saturation	-	CropReporter
R_{Red}	Reflectance in red color	-	CropReporter
R_{Green}	Reflectance in green color	-	CropReporter
R_{Blue}	Reflectance in blue color	-	CropReporter
Chlorophyll fluorescence traits			
F_0	Minimum fluorescence of dark-adapted leaves	-	CropReporter
F_m	Maximum fluorescence of dark-adapted leaves	-	CropReporter
F_v/F_m	Maximum quantum yield of PSII photochemistry	-	CropReporter

2.3 Trait selection for model formulation

To select traits for forming biomass estimation model, multicollinearity was considered to confirm that there was no one or more pairs of predictors is highly correlated. In regression analysis, multicollinearity led to unreliable inferences about the effects of the predictor variables on the response variables (Bhandari, 2023). To solve this multicollinearity problem, correlation among traits was considered. The very high positive or negative correlation, with correlation coefficient was greater than 0.7 or lower than -0.7, was considered as multicollinearity. If any pair of traits show a high positive or negative correlation, one of these traits will be excluded from the model

formulation based on their correlation with harvested biomass. The trait that had a higher correlation coefficient with harvested biomass was used for model formulation. The consideration was continued until there was no pair of traits showed high positive or negative correlation.

Then the biomass estimation models were formulated by using multiple linear regression. Multicollinearity was checked again by using variance inflation factor (VIF). The lower VIF values showed less possibility of multicollinearity in the model. The acceptable VIF values was lower than 2.5 (Johnston et al., 2018). The formulated models that all traits had VIF values lower than 2.5 were acceptable and were used for biomass estimation. If at least one of the predictors in the model had higher 2.5 of VIF values, one of the predictors will be eliminated based on correlation of determination (R^2) of the model until there was no predictors had higher 2.5 of VIF values.

2.4 Biomass estimation models

There were two main types of biomass estimation models formulated in this study, which were the biomass estimation model with single and multiple types of traits. The models were formulated by using multiple linear regression models, where harvested biomass was a response variable and selected image-based traits were predictors. For models with single type of traits, there were three biomass estimation deriving from three types of traits (Table 2.1) and are shown in the Equation 2.1 to 2.3:

Model 1: RGB sideview trait model

$$Y_1 = \beta_{1,0} + \sum_{i=1}^{n} \beta_{1,i} RGB_i + \varepsilon_1 \quad \text{(Equation 2.1)}$$

Model 2: Spectral imaging trait model

$$Y_2 = \beta_{2,0} + \sum_{i=1}^{m} \beta_{2,i} SI_i + \varepsilon_2 \quad \text{(Equation 2.2)}$$

Model 3: Chlorophyll fluorescence trait model

$$Y_3 = \beta_{3,0} + \sum_{i=1}^{p} \beta_{3,i} CF_i + \varepsilon_3 \quad \text{(Equation 2.3)}$$

Where Y_i is estimated biomass of model i, β is coefficient,

RGB is RGB sideview trait, ε is residuals,

SI is spectral imaging trait,

CF is chlorophyll fluorescence trait,

n, m, p is number of RGB, $color$, and CF traits, respectively.

Then the combination of different types of image-based traits were studied. The biomass estimation models with multiple traits were also formed by using multiple linear regression and using the same procedure for trait selection. There were four possible combinations of multiple traits which are shown in Equation 2.4 to 2.7.

Model 4: RGB sideview + Spectral imaging trait model

$$Y_4 = \beta_{4,0} + \sum_{i=1}^{n} \beta_{4,i} RGB_i + \sum_{j=1}^{m} \beta_{4,j} SI + \varepsilon \quad \text{(Equation 2.4)}$$

Model 5: RGB sideview + Chlorophyll fluorescence trait model

$$Y_5 = \beta_{5,0} + \sum_{i=1}^{n} \beta_{5,i} RGB_i + \sum_{k=1}^{p} \beta_{5,k} CF_i + \varepsilon \quad \text{(Equation 2.5)}$$

Model 6: Spectral imaging + Chlorophyll fluorescence trait model

$$Y_6 = \beta_{6,0} + \sum_{j=1}^{m} \beta_{6,j} SI_i + \sum_{k=1}^{p} \beta_{6,k} CF_i + \varepsilon \quad \text{(Equation 2.6)}$$

Model 7: RGB sideview + Spectral imaging + Chlorophyll fluorescence trait model

$$Y_7 = \beta_{7,0} + \sum_{i=1}^{n} \beta_{7,i} RGB_i + \sum_{j=1}^{m} \beta_{7,j} SI_i + \sum_{k=1}^{p} \beta_{7,k} CF_i + \varepsilon \quad \text{(Equation 2.7)}$$

Once the collections of selected predictors from each model was determined and biomass estimation was formed, the models were evaluated by using 10-fold cross validation. The data was randomly separated into 10 groups of approximately equal sample size. Nine groups were used as training data set, while the remaining group was a testing data set. The root mean squared error ($RMSE_i$) was then computed on the observations in the held-out fold. This procedure was repeated 10 times; each time, a different group of data set was treated as a testing data set. This process results in 10 estimates of the testing error. Then the 10-fold cross validation estimate was computed by averaging these values (Equation 2.8):

$$RMSE = \frac{1}{10} \sum_{i=1}^{10} RMSE_i \quad \text{(Equation 2.8)}$$

RMSE was determined in both training data and testing data. In addition, correlation of determination (R^2) was also used to evaluate model performance together with RMSE. The best performance biomass estimation model was considered by the lowest RMSE and highest R^2.

2.5 Data and Statistical analysis

A two-way ANOVA for completely randomized design followed by Fisher's protected least significant difference (LSD) test was used to determine the statistical significance of differences in means of image-based traits, and harvested biomass among treatments with p-value < 0.05. The Pearson correlation coefficient and principal component analysis (PCA) was used to investigate relationship among traits and consider multicollinearity in the study with p-value < 0.05. Multiple linear regression analysis was used to form biomass model via the image-based traits as predictors and harvested biomass as a response variable. The variance inflation factor (VIF) was considered to confirm multicollinearity problem and excluding predictors until all predictors had VIF values lower than 2.5. After that, 10-fold cross validation was used to evaluate biomass estimation models by considering correlation of determination (R^2), root mean square error (RMSE) of both training and testing data. All statistical analysis part was conducted by RStudio version 4.3.1 with 'agricolae', 'multicompView', 'factoextra', 'vip', and 'tidyverse' (RStudio Team 2020).

Chapter 3: Results

To study the performance of the biomass estimation model by using a combination of image-based phenotypic traits, 180 pea genotypes grown in control and drought conditions were assessed to investigate differences in biomass traits and image-based traits (Figure 3.1). Then, biomass and image-based traits were used to build biomass estimation models with single or multiple types of traits. The results are represented in five sections: (3.1) the effect of drought on biomass, (3.2) the effect of drought on image-based phenotypic traits, (3.3) relationship among phenotypic traits, (3.4) biomass estimation model formulation with single type of image-based trait, and (3.5) biomass estimation model formulation with multiple types of image-based traits.

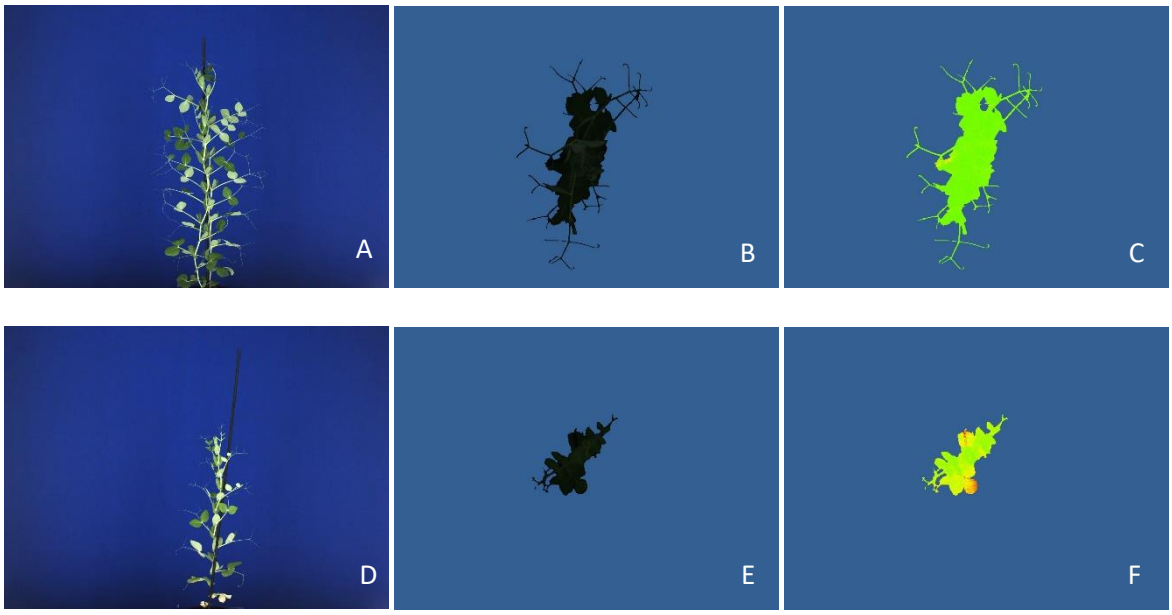


Figure 3.1 Pea (genotype *Sp154*) was grown under control conditions (A, B, C) and drought conditions (D, E, F) at approximately six weeks after sowing. The images were captured using different camera systems: RGB sideview camera systems (A, D); topview spectral imaging via CropReporter (B, E); topview chlorophyll fluorescence (C, F). In the case of chlorophyll fluorescence, a shade of green to yellow color indicated a decrease in photosynthetic rate.

3.1 The effect of drought on biomass

There was no significant interaction effect between genotype and treatment on both total fresh (FW) and dry weight (DW) harvested at the end of the experiment ($p < 0.001$, Figure 3.2, and Table 3A.1.1). FW and DW showed significant differences between the control and drought conditions ($p < 0.001$, Figure 3.2, and Table 3A.1.1). Pea plants grown under control conditions had higher FW and DW, approximately 22 g and 4 g, respectively, while those grown under drought conditions had approximately 12 g and 2 g of FW and DW (Figure 3.2). Additionally, a significant effect of genotype was also observed in this study ($p < 0.001$, Table 3A.1.1). Variation was observed in FW and DW with approximate values in each genotype ranging between 5 and 50 g, and 0.5 and 9 g, respectively (Figure 3.3).

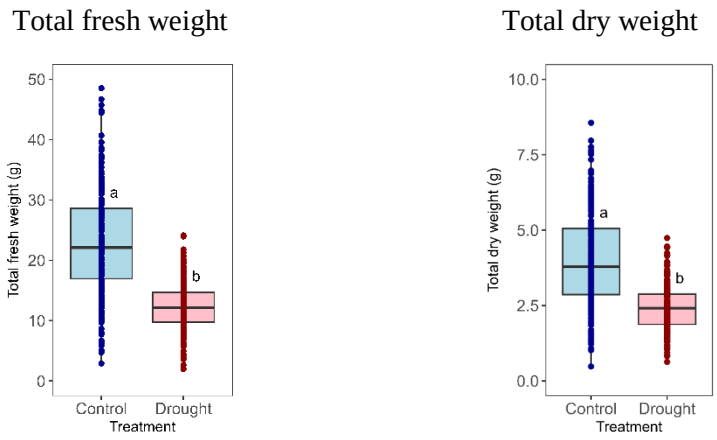


Figure 3.2 Box plots of total fresh weight (left) and total dry weight (right) of pea grown under control (blue) and drought (red) conditions. The letters above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

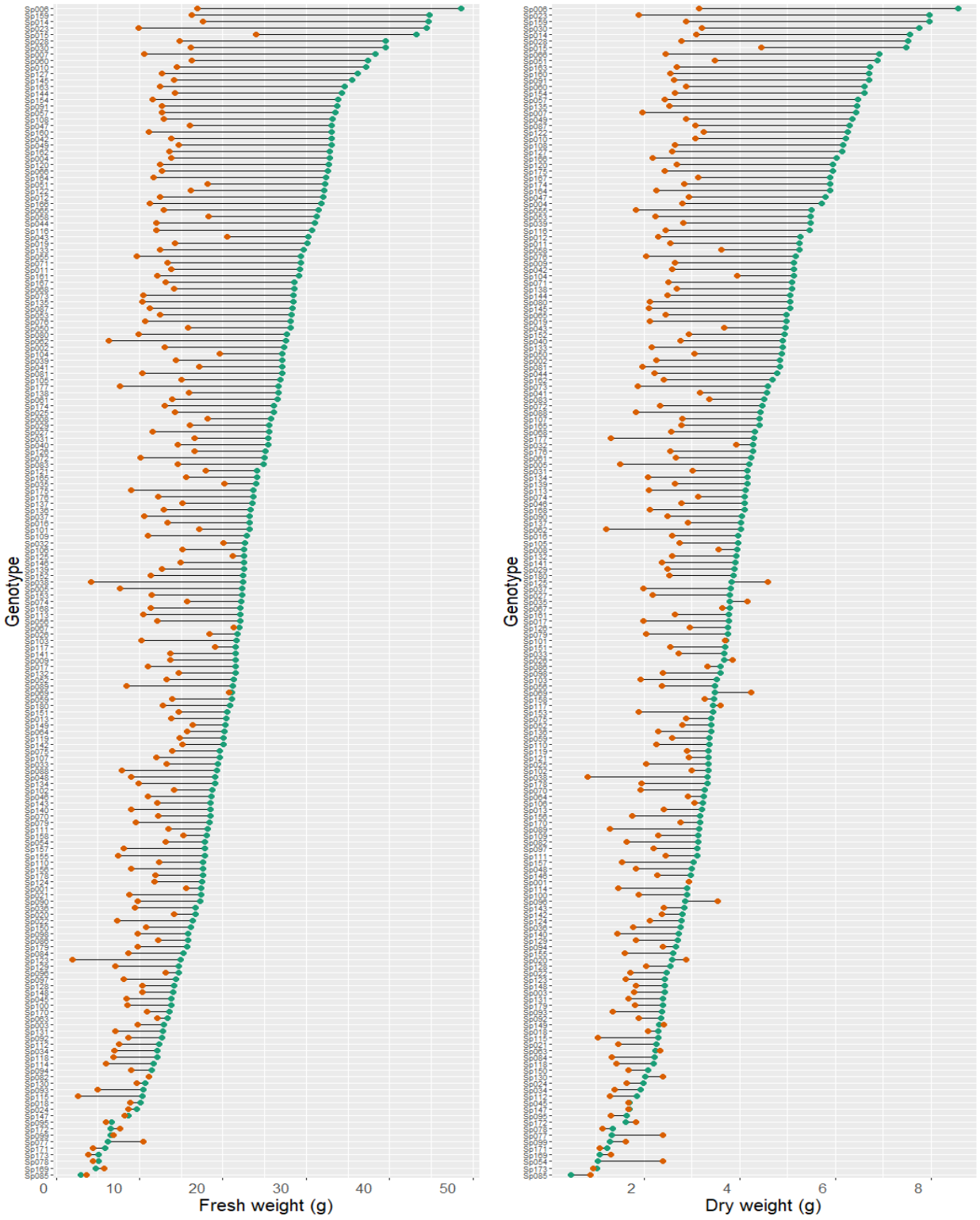


Figure 3.3 Dot plots showing variations of total fresh weight (left) and total dry weight (right) of all genotypes of pea grown under control (green) and drought (orange) conditions, ordering from highest weight to lowest weight in control condition.

3.2 The effect of drought on image-based phenotypic traits

Image-based phenotypic traits were categorized into three main types based on the types of sensors used in the measurements: (3.2.1) RGB sideview traits; (3.2.2) Spectral imaging (SI) traits (3.2.3) Chlorophyll fluorescence (CF) traits.

3.2.1 RGB sideview traits

There were five RGB sideview traits: topview digital biomass (DB_Top), average sideview digital biomass (DB_Side), convex hull (CH), height, solidity. An interaction effect was observed between pea genotypes and treatments on all traits ($p < 0.05$, Figure 3.4, and Table 3A.1.2). The significant effect of treatment was also found in all traits ($p < 0.001$, Figure 3.4, and Table 3A.1.2). Pea plants grown under control conditions exhibits a larger DB_Top and DB_Side compared to those grown under drought conditions (Figure 3.4). DB_Top was approximately average of 120 and 80 cm^2 in control and drought conditions, respectively, while DB_Side was of approximately 300 and 200 cm^2, in control and drought conditions respectively (Figure 3.4). Plant height and CH were also significantly higher in control than drought conditions ($p < 0.001$, Figure 3.4, and Table 3A.1.2). Only solidity was significantly lower in controlled plants than drought-stressed plants ($p < 0.001$, Figure 3.4, and Table 3A.1.2).

Additionally, the sideview digital biomass, convex hull, height, and solidity at six different angles, as well as the median, minimum, and maximum of these traits, showed significant effects of genotypes, treatments, and their interaction ($p < 0.001$, Appendix 3A.2). The trend of each side, median, minimum, and maximum was similar to average, with significantly higher in plants grown under control conditions than drought conditions ($p < 0.001$, Appendix 3A.2).

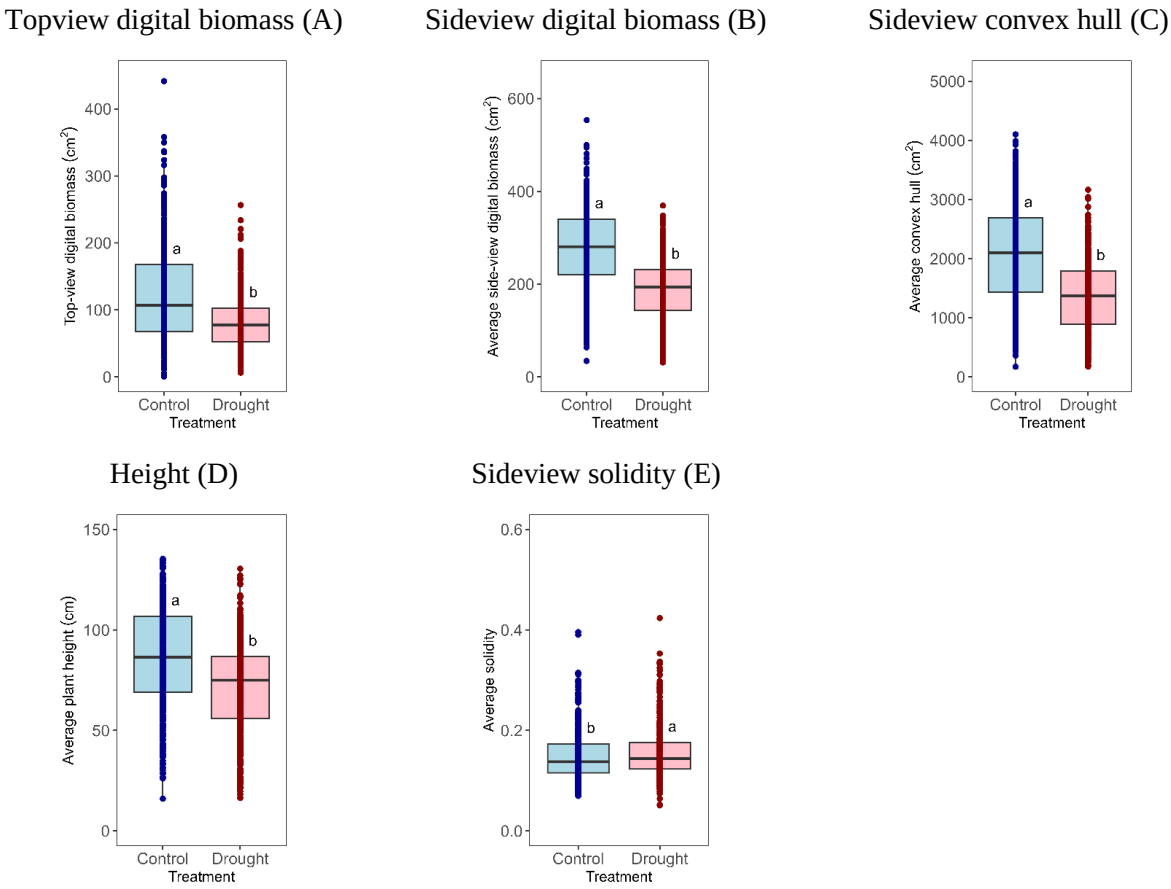


Figure 3.4 Box plots of topview digital biomass (A), average sideview digital biomass (B), average sideview convex hull (C), average plant height (D), average sideview solidity (E) of pea grown under control (blue) and drought (red) conditions. The letters above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

3.2.2 Spectral imaging traits

In this study, six spectral imaging (SI) traits were collected, including spectral reflectance in red (R_{Red}), spectral reflectance in green (R_{Green}), spectral reflectance in blue (R_{Blue}), hue (HUE), saturation (SAT), value (VAL), The results revealed significant effects of genotypes, treatments, and their interaction on all SI traits ($p < 0.05$, Figure 3.5, and Table 3A.1.3).

Pea plants grown under control conditions exhibited higher R_{Red}, R_{Green}, and R_{Blue} compared to those grown under drought conditions (Figure 3.5). Specifically, under control conditions, the average R_{Red}, R_{Green}, and R_{Blue} were 2,300, 3,500, and 1,600, respectively, while approximately 2,100, 3,200, and 1,500 were observed in pea plants grown under drought conditions (Figure 3.5).

A similar trend was observed on SAT and VAL, with higher values in pea plants grown under control conditions compared to those grown under drought conditions. The SAT in pea plants grown under control conditions was approximately 0.55, whereas it was around 0.50 in plants grown under drought conditions (Figure 3.5). Similarly, under control conditions, pea plants had slightly higher values for the parameter of VAL compared to drought conditions, with a difference of approximately 0.05 (Figure 3.5). In contrast, HUE was the only trait that had higher values in pea plants grown under drought conditions compared to those grown under control conditions, with approximate values of 105 and 100, respectively (Figure 3.5).

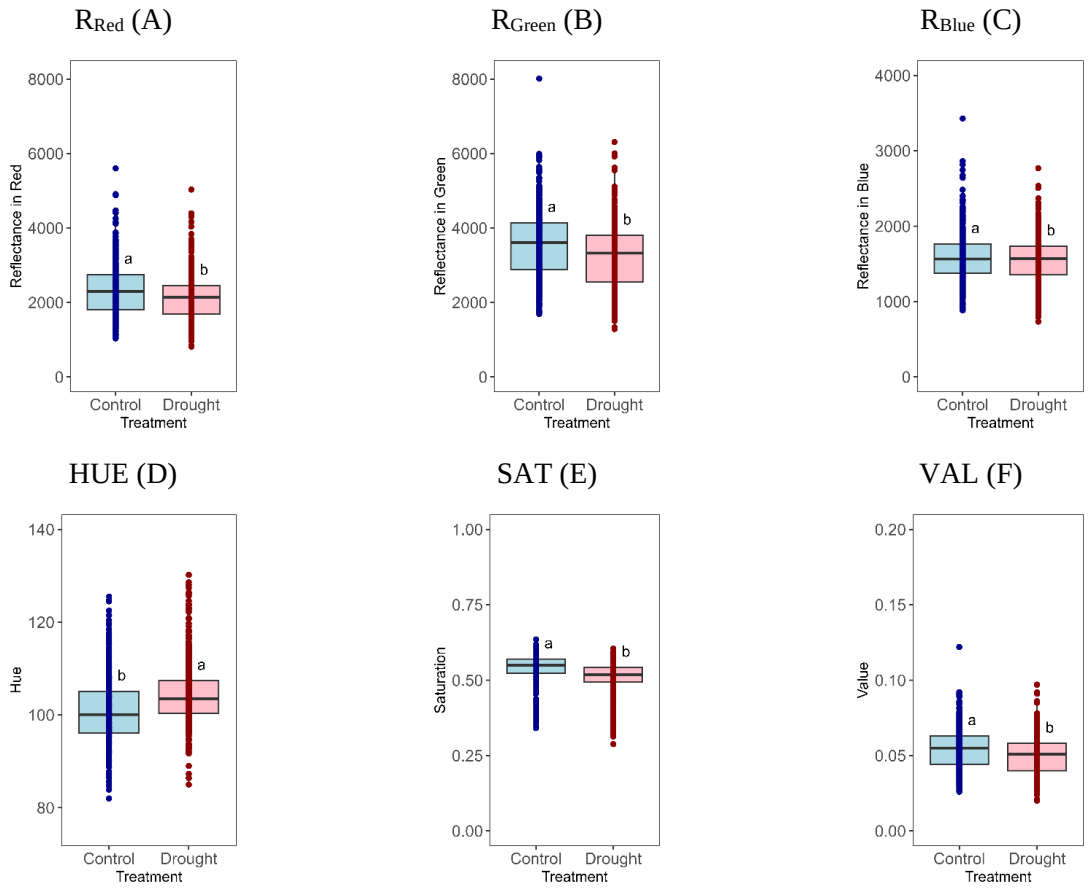


Figure 3.5 Box plots of spectral imaging traits of pea grown under control (blue) and drought (red) conditions: reflectance of red color (R_{Red}: A), reflectance of green color (R_{Green}: B), reflectance of blue color (R_{Blue}: C), hue (HUE: D), saturation (SAT: E), and value (VAL: F). The letters above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

3.2.3 Chlorophyll fluorescence traits

There were three parameters for chlorophyll fluorescence (CF) traits: minimum and maximum fluorescence of dark-adapted leaves (F_0 and F_m), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m). There was no significant interaction effect between genotypes and treatments observed in any of the chlorophyll fluorescence traits ($p > 0.05$, Figure 3.6, and Table 3A.1.4). However, significant effects of genotype and treatment were found in F_0 and F_v/F_m ($p < 0.05$, Figure 3.6, and Table 3A.1.4). Regarding F_0, pea plants grown under control conditions had higher values compared to those grown under drought conditions, with approximately 1,400 and 1,300, respectively (Figure 3.6). On the other hand, F_v/F_m in pea plants grown under drought

conditions showed significantly higher values than those grown under control conditions, with approximately 0.77 and 0.76, respectively (Figure 3.6). Furthermore, only a significant effect of genotype was observed in F_m ($p < 0.001$, Figure 3.6, and Table 3A.1.4), while there was no significant effect of treatment on F_m ($p = 0.066$, Figure 3.6, and Table 3A.1.4), and the average value of F_m was approximately 6,400 in both treatment conditions.

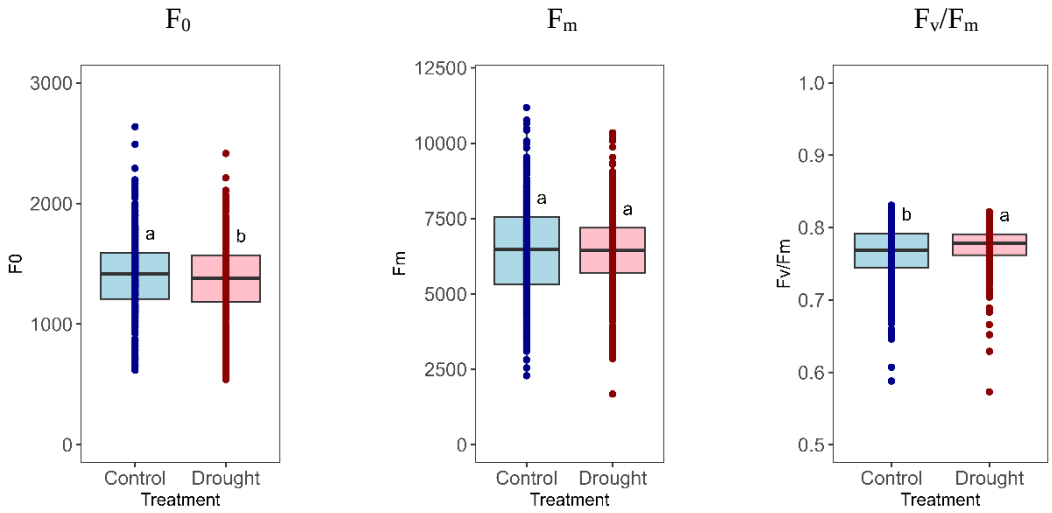


Figure 3.6 Box plots of chlorophyll fluorescence traits of pea grown under control (blue) and drought (red) conditions: minimum fluorescence of dark-adapted leaves (F_0, left), maximum fluorescence of dark-adapted leaves (F_m, middle), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m, right). The letters above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

3.3 Relationship among phenotypic traits

The criteria for evaluating correlation coefficient and its interpretation can be found in Appendix 3A.3. A high positive correlation was observed between RGB sideview traits and harvested biomass in both fresh (FW) and dry weight (DW), specifically topview digital biomass (DB_Top) and sideview digital biomass (DB_Side), with indicating a high positive correlation with harvested biomass ($r > 0.7$, Figure 3.7). Convex hull (CH) and height also showed moderate positive correlation with FW and DW, while solidity had a very low negative correlation (Figure 3.7). Low positive correlations were found between chlorophyll fluorescence (CF) traits and harvested

biomass, with the highest correlation coefficient belonging to the maximum fluorescence of dark-adapted leaves (F_m) at approximately 0.38 and 0.44 for FW and DW, respectively (Figure 3.7). On the other hand, most spectral imaging (SI) traits showed no significant correlation with biomass, except for saturation (SAT) and reflectance in green color (R_{Green}), which exhibited a very low correlation (Figure 3.7).

Correlations were also observed among different types of image-based phenotypic traits. Only DB_Top showed a high positive correlation with the minimum fluorescence of dark-adapted leaves (F_0) and F_m, while DB_Top showed a very low positive with the maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m). The other traits in RGB sideview traits showed very low positive correlation with F_0 and F_m, and very low negative correlation with F_v/F_m, except for solidity that showed the opposite direction of correlation (Figure 3.7). RGB sideview traits, except for solidity, showed low positive correlation with almost SI traits except for HUE, while solidity and HUE had a low positive correlation (Figure 3.7). Only SAT showed moderate positive correlations with both F_0 and F_m, while exhibiting a low negative correlation with F_v/F_m (Figure 3.7). Furthermore, there was a moderate negative correlation between F_v/F_m and most SI traits, except for HUE, which had a moderate positive correlation (Figure 3.7).

Within the same types of image-based phenotypic traits, there was moderate and high positive correlations among RGB sideview traits, except for solidity which had a negative correlation to the others (Figure 3.7). A very high positive correlation was found among R_{Red}, R_{Green}, R_{Blue}, and value (VAL), while HUE had a high negative correlation with other traits in SI traits (Figure 3.7). Finally, a high positive correlation was found between F_0 and F_m, while low negative and positive correlations were observed between F_v/F_m and F_0, and F_m, respectively (Figure 3.7).

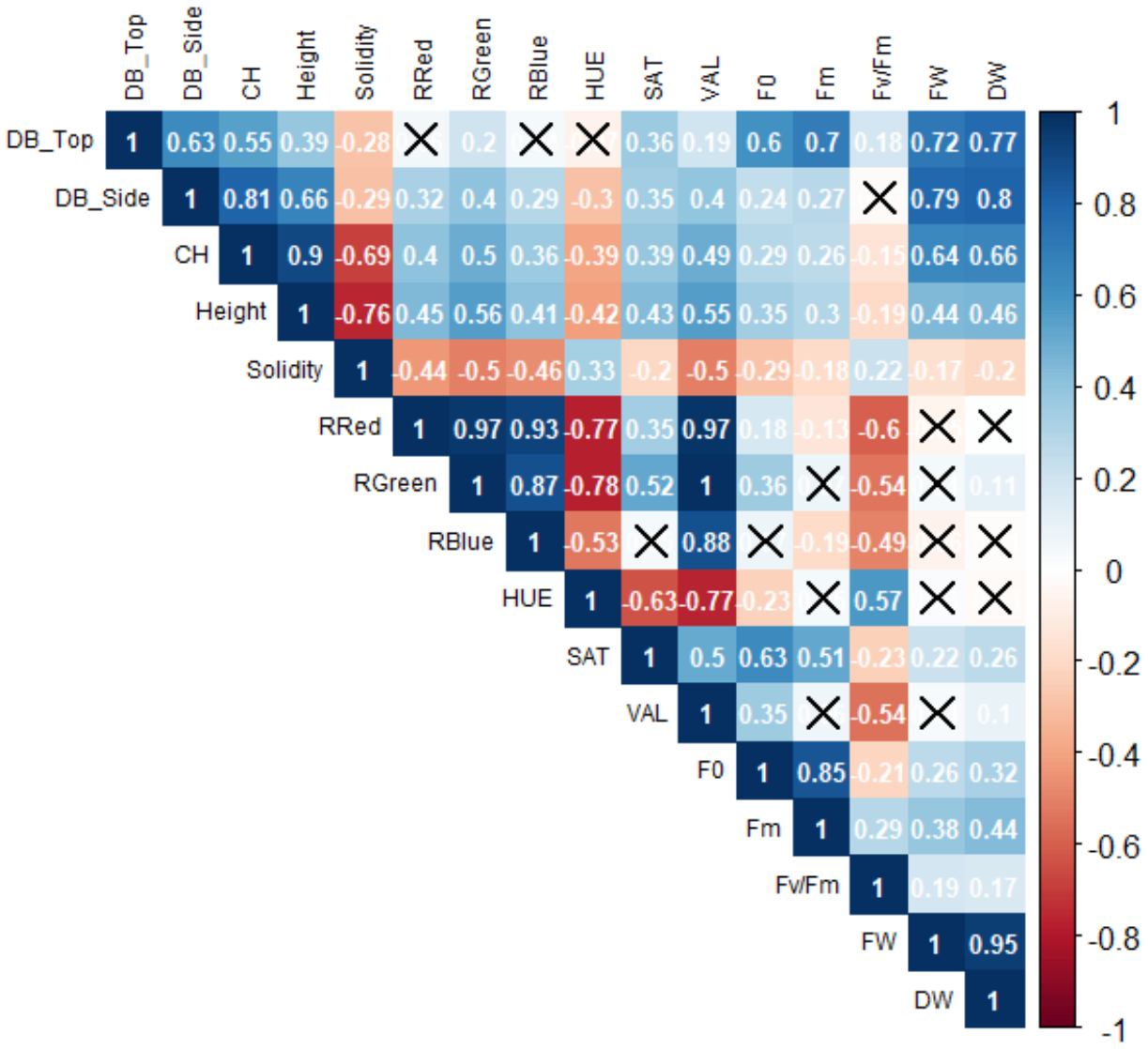


Figure 3.7 Correlation matrix for among all image-based traits, total fresh and dry weight of pea grown under control and drought conditions. Positive and negative correlations are illustrated in blue and red colors, respectively. Color shade is proportional to the correlation coefficients, with their value shown in the color intensity bar. Cross mark (X) shows that there was no significant correlations between two parameters; topview digital biomass (DB_Top), sideview digital biomass (DB_Side), sideview convex hull (CH), height, solidity, reflectance of red color (R_{Red}), reflectance of green color (R_{Green}), reflectance of blue color (R_{Blue}), hue (HUE), saturation (SAT), value (VAL), minimum fluorescence of dark-adapted leaves (F_0), maximum fluorescence of dark-adapted leaves (F_m), maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m), total fresh weight (FW), and total dry weight (DW).

A principal component analysis (PCA) was conducted on all image-based phenotypic traits, resulting in two principal components (PCs) being considered. These two PCs accounted for nearly 90.8% of the cumulative variance in all the traits (Figure 3.8, Appendix 3A.4). Overall, traits categorized in the first PC (Dim1) were mostly SI traits and F_v/F_m, while RGB sideview traits were mostly categorized in the second PC (Dim2). The Dim1 explained 74.1% of the variance and consisted of height, solidity, R_{Red}, R_{Green}, R_{Blue}, HUE, VAL, and F_v/F_m. Notably, solidity, HUE, and F_v/F_m showed similar trends, while the other traits in Dim1 exhibited opposite directions (Figure 3.8, Appendix 3A.4). The Dim2 accounted for 16.7% of the variance and included DB_Top, DB_Side, CH, SAT, F_0, and F_m. Almost the traits in Dim2 displayed a similar direction except for F_m (Figure 3.8, Appendix 3A.4).

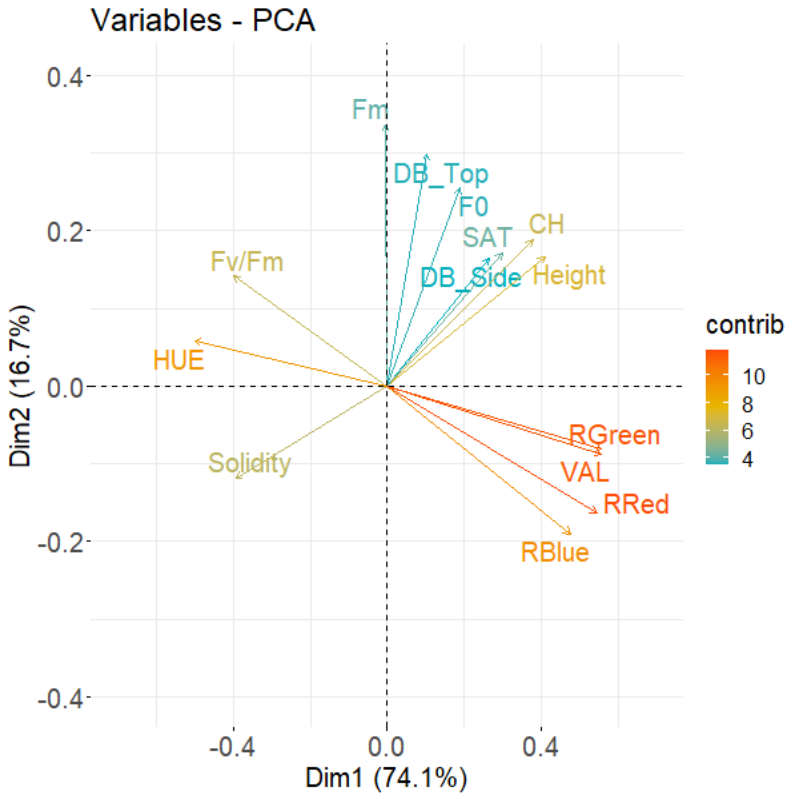


Figure 3.8 Trait loading scores of all traits for principal component analysis; topview digital biomass (DB_Top), sideview digital biomass (DB_Side), convex hull (CH), height, solidity, reflectance of red color (R_{Red}), reflectance of green color (R_{Green}), reflectance of blue color (R_{Blue}), hue (HUE), saturation (SAT), value (VAL), minimum fluorescence of dark-adapted leaves (F_0), maximum fluorescence of dark-adapted leaves (F_m), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m).

3.4 Biomass estimation models formulation with single type of image-based traits

There were three biomass estimation models based on different types of image-based phenotypic traits: (i) a model based on RGB sideview traits, (ii) a model based on spectral imaging (SI) traits, and (iii) a model based on chlorophyll fluorescence (CF) traits. To select parameters for model formulation, multicollinearity was assessed by considering correlation coefficients and the variance inflation factor (VIF). Correlation coefficients higher than 0.7 (Songara, 2022) and VIF values exceeding 2.5 indicated the presence of considerable multicollinearity (Johnston et al., 2018). The set of parameters used for model formulation was chosen to include the maximum number of traits while avoiding any multicollinearity problems and providing the highest correlation of determination (R^2). The detail for parameter selection was shown in the appendix (Appendix 3A.5)

Regarding RGB sideview traits, the model that yielded the best performance without encountering any multicollinearity issues was the one that incorporated both topview (DB_Top) and sideview digital biomass (DB_Side) (Table 3A.5.1). The biomass estimation model utilizing RGB sideview traits can be represented by equation 3.1. For the model based on SI traits, a multicollinearity problem arose when all parameters were included in the model formulation. However, the optimal model in this scenario was constructed using two traits, namely hue (HUE) and saturation (SAT), and it can be expressed through equation 3.2. Lastly, the model utilizing CF traits that exhibited no multicollinearity problems and achieved the highest R^2, was developed using the maximum fluorescence of dark-adapted leaves (F_m) and the maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m). The equation representing this model is equation 3.3.

Biomass estimation model with RGB sideview traits:

$$y = a_{1,0} + a_{1,1}DB\_Top + a_{1,2}DB\_Side \quad \text{(Equation 3.1)}$$

Biomass estimation model with spectral imaging traits:

$$y = a_{2,0} + a_{2,1}HUE + a_{2,2}SAT \quad \text{(Equation 3.2)}$$

Biomass estimation model with chlorophyll fluorescence traits:

$$y = a_{3,0} + a_{3,1}F_m + a_{3,2}F_v/F_m \quad \text{(Equation 3.3)}$$

Where y is biomass traits (total fresh or dry weight) in gram (g.),

DB_Top is topview digital biomass (cm^2)

DB_Side is average of sideview digital biomass (cm^2)

HUE is hue,

SAT is saturation,

F_m is maximum fluorescence of dark-adapted leaves,

F_v/F_m is maximum quantum yield of photosystem II (PSII) photochemistry,

$a_{i,j}$ is coefficient.

After parameter selection for each model, fresh weight (FW) and dry weight (DW) were considered and used to form biomass estimation models. The 10-fold cross-validation was used to investigate the model performance. The average correlation of determination (R^2) and root mean square error (RMSE) of the training and testing datasets of the training and testing datasets were evaluated (Table 3.1).

Table 3.1 Correlation of determination (R^2) and root mean square error (RMSE) of training and testing dataset of fresh and dry weight biomass estimation model with a single type of trait.

Types of image-based phenotypic traits	R^2	RMSE (g.)	
		Training error	Testing error
RGB sideview traits			
Fresh weight	0.70	4.78	4.76
Dry weight	0.76	0.73	0.72
Spectral imaging traits			
Fresh weight	0.09	8.28	8.31
Dry weight	0.10	1.42	1.42
Chlorophyll fluorescence traits			
Fresh weight	0.15	8.01	8.04
Dry weight	0.19	1.34	1.35

The results showed that all the biomass estimation models were better in DW estimation than in FW estimation (Table 3.1). Among the three biomass estimation models, the model built from RGB sideview traits had a much better performance than the other models in both FW and DW estimation. The RGB sideview trait model provided the highest R^2 and the lowest RMSE in both training and testing errors (Table 3.1). The model built from CF traits showed the second-best performance in formulating biomass estimation models, with R^2 values of approximately 0.15 and 0.20 for FW and DW, respectively (Table 3.1). The model derived from SI traits had the lowest R^2. However, there was a slight difference in RMSE between the model from SI and CF traits model (Table 3.1). The results also showed that the model based on RGB sideview traits had a stronger capability than those based on SI and CF traits in estimating biomass, as evidenced by the correlation between observed and estimated biomass (Figure 3.9).

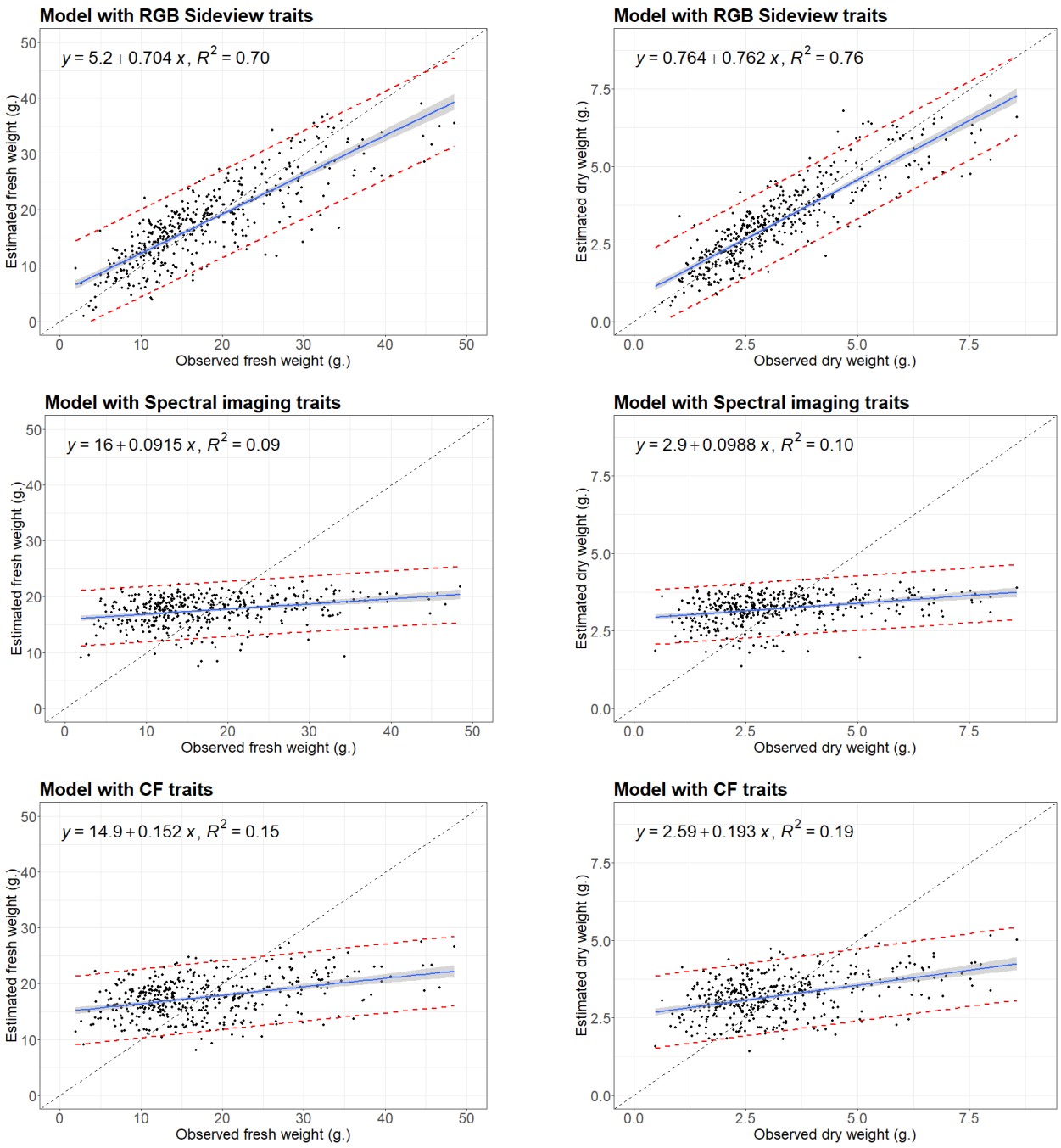


Figure 3.9 The scatterplot of observed fresh weight (left) and dry weight (right) in gram (g.) versus estimated biomass in gram (g.) using RGB sideview traits (top), spectral imaging (SI) traits (middle), and chlorophyll fluorescence (CF) traits (bottom). Blue line shows a linear regression line; Grey-shaded area shows confidence interval at 95%; Red-dashed lines show the lower and upper of prediction interval. Black-dashed line show the 1:1 line.

3.5 Biomass estimation model formulation with multiple types of image-based traits

Biomass estimation models were formulated using multiple types of image-based phenotypic traits in four scenarios: (i) a combination of RGB sideview and spectral imaging (SI) traits, (ii) a combination of RGB sideview and chlorophyll fluorescence (CF) traits, (iii) a combination of SI and CF traits, and (iv) a combination of all traits. The initial models were formulated by combining parameters from the models with single types of traits. Similar to the models with single types of traits, the parameters used in each model were selected by considering the correlation among traits and variance inflation factor (VIF) values. One parameter was excluded from the model until there were no multicollinearity issues (Appendix 3A.6).

The model derived from the combination of RGB sideview, and SI traits initially included four variables: topview digital biomass (DB_Top), average sideview digital biomass (DB_Side), HUE, and saturation (SAT), as there were no multicollinearity issues. These four parameters were then included in the model formulation, and the model can be expressed by equation 3.4. For the model combining RGB sideview and CF traits, three parameters were used due to multicollinearity. The maximum fluorescence of dark-adapted leaves (F_m) was excluded from this scenario, and the model can be expressed by equation 3.5. Similarly, the model combining SI and CF traits showed multicollinearity with four variables. Excluding HUE from the model provided a higher R^2 compared to excluding SAT and resolved the multicollinearity problem. Therefore, the biomass estimation model in this case utilized SAT, F_m, and the maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m), and can be represented by equation 3.6. Lastly, when combining all traits, the model initially included all six parameters, but multicollinearity was found. To address this, the number of parameters was reduced until the model consisted of four: DB_Top, DB_Side, HUE, and F_v/F_m, represented by equation 3.7.

Biomass estimation model with RGB sideview + spectral imaging traits:

$$y = a_{4,0} + a_{4,1}DB\_Top + a_{4,2}DB\_Side + a_{4,3}HUE + a_{4,4}SAT \quad \text{(Equation 3.4)}$$

Biomass estimation model with RGB sideview + chlorophyll fluorescence traits:

$$y = a_{5,0} + a_{5,1}DB\_Top + a_{5,2}DB\_Side + a_{5,3}F_v/F_m \quad \text{(Equation 3.5)}$$

Biomass estimation model with spectral imaging + chlorophyll fluorescence traits:

$$y = a_{6,0} + a_{6,1}SAT + a_{6,2}F_m + a_{6,3}F_v/F_m \quad \text{(Equation 3.6)}$$

Biomass estimation model with all traits:

$$y = a_{7,0} + a_{7,1}DB\_Top + a_{7,2}DB\_Side + a_{7,3}HUE + a_{7,4}F_v/F_m \quad \text{(Equation 3.7)}$$

Where

y	is biomass traits (total fresh or dry weight) in gram (g.),
DB_Top	is topview plant pixel from CropReporter,
DB_Side	is sideview plant pixel,
HUE	is hue,
SAT	is saturation,
F_m	is maximum fluorescence of dark-adapted leaves,
F_v/F_m	is maximum quantum yield of photosystem II (PSII) photochemistry,
$a_{i,j}$	is coefficient.

Four biomass estimation models were formulated and evaluated based on their performance, assessed through the correlation of determination (R^2), and root mean square error (RMSE). Consistent with the model utilizing a single type of image-based phenotypic traits, all the models demonstrated better performance in estimating dry weight (DW) compared to fresh weight (FW), as indicated by higher R^2 values and lower RMSE values (Table 3.2).

Table 3.2 Correlation of determination (R^2) and root mean square error (RMSE) of training and testing dataset of fresh and dry weight biomass estimation model with multiple types of traits.

Types of image-based phenotypic traits	R^2	RMSE (g.)	
		Training error	Testing error
RGB sideview + spectral imaging traits			
Fresh weight	0.76	4.31	4.30
Dry weight	0.79	0.68	0.68
RGB sideview + chlorophyll fluorescence traits			
Fresh weight	0.72	4.61	4.61
Dry weight	0.77	0.71	0.71
Spectral imaging + chlorophyll fluorescence traits			
Fresh weight	0.16	7.98	8.03
Dry weight	0.20	1.34	1.35
All traits			
Fresh weight	0.76	4.31	4.31
Dry weight	0.79	0.68	0.68

The results indicated that the model combining RGB sideview and SI traits outperformed the other models. It achieved the highest R^2 and the lowest RMSE compared to the others (Table 3.2). The biomass model incorporating all traits performed well, ranking second and closely matching the performance of the combination of RGB sideview and SI traits. They exhibited almost similar values of R^2 and RMSE for the training error. However, the model with all traits showed slightly higher values of RMSE for the testing error (Table 3.2). The model incorporating RGB sideview and CF traits also demonstrated a high R^2 and low RMSE (Table 3.2). In contrast, the model combining SI and CF traits exhibited the lowest R^2 and the highest RMSE. The differences in these

values were substantial compared to the other three models (Table 3.2). The results also showed that the models that had parameters from RGB sideview traits had a stronger capability than those based on the combination of SI and CF traits in estimating biomass in both FW and DW, as evidenced by the correlation between observed and estimated biomass (Figure 3.10 and Figure 3.11).

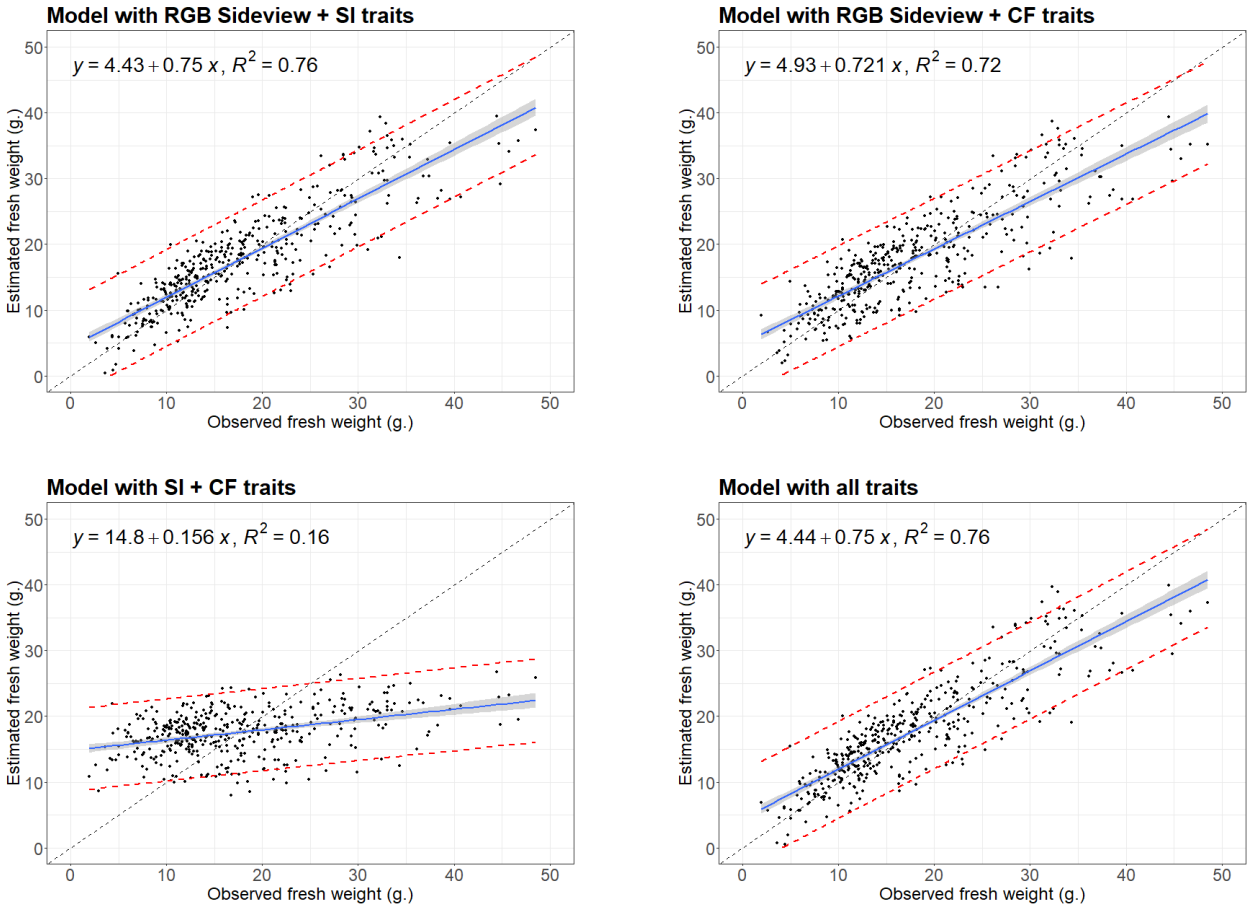


Figure 3.10 The scatterplot of observed fresh weight (FW) in gram (g.) versus estimated FW in gram (g.) using multiple image-based traits; RGB sideview spectral imaging (SI) traits (top-left), RGB sideview and CF traits (top-right), SI and CF traits (bottom-left), and all traits (bottom-right). Blue line shows a linear regression line; Grey-shaded area shows confidence interval at 95%; Red-dashed lines show the lower and upper of prediction interval; Black-dashed line show the 1:1 line.

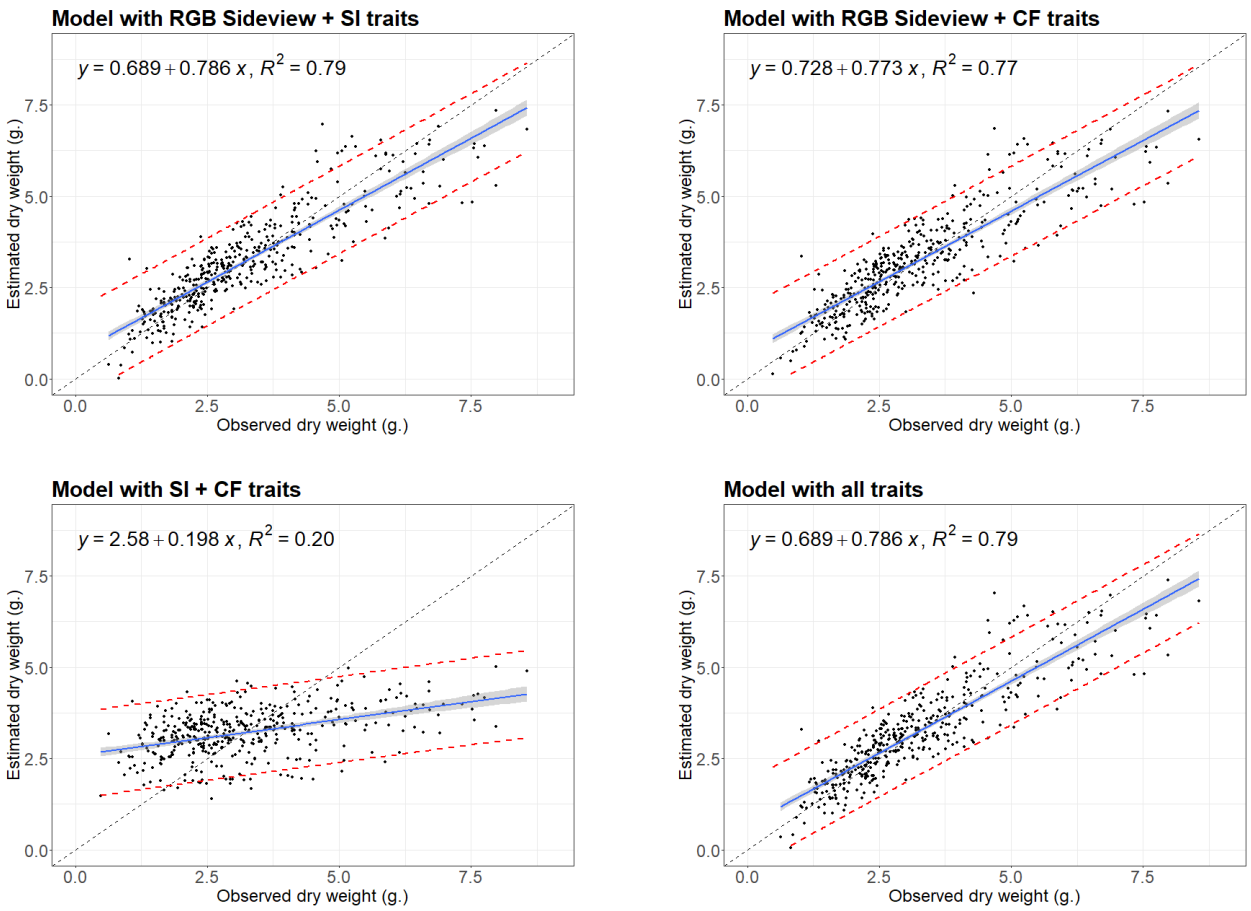


Figure 3.11 The scatterplot of observed dry weight (DW) in gram (g.) versus estimated DW in gram (g.) using multiple image-based traits; RGB sideview traits, and spectral imaging (SI) traits (top-left), RGB sideview traits and chlorophyll fluorescence (CF) traits (top-right), SI and CF traits (bottom-left), and all traits (bottom-right). Blue line shows a linear regression line; Grey-shaded area shows confidence interval at 95%; Red-dashed lines show the lower and upper of prediction interval; Black-dashed line show the 1:1 line.

Chapter 4: Discussion

This study describes the use of image-based phenotyping to investigate pea's responses under drought conditions and improve biomass estimation using the combination of different types of non-invasive image-based phenotypic traits. The study focused on three types of image-based phenotypic traits: (i) RGB sideview traits; (ii) spectral imaging (SI) traits; (iii) chlorophyll fluorescence (CF) traits. The first type of trait represented the morphology and architecture of pea plants, while alteration of pigments and photosynthetic activity were related to the second and third trait type, respectively. Our hypothesis was that the biomass estimation could be enhanced by incorporating RGB sideview traits with either SI or CF traits or both in the biomass estimation model. This is because the combination of different types of traits covers both the morphological and physiological mechanisms of plants, which ultimately affect plant biomass. Partially supporting our hypothesis, the biomass estimation model with multiple types of image-based phenotypic traits showed a slight improvement in accuracy. Among the tested models, the combination of RGB sideview traits and SI traits exhibited the highest accuracy in the biomass estimation model. It was expected that all-traits model would achieve the highest accuracy, while the all-traits model was ranked the second-best model. The combination of RGB sideview traits with CF traits resulted in lower accuracy compared to the previously mentioned models, but it still outperformed the model derived from RGB sideview traits alone, as expected. Additionally, the models generated without RGB sideview traits displayed substantially lower accuracy in biomass estimation, highlighting the crucial role of RGB sideview in achieving the highest accuracy. These findings are further discussed in terms of the ability of different types of image-based traits to investigate plant responses to drought and the performance of the developed biomass estimation.

4.1 More pronounced changes in morphological traits than in other traits in drought

Analysis of the RGB sideview imaging obtained from RGB sideview imaging and the camera system CropReporter, such as topview digital biomass (DB_Top), average sideview digital biomass (DB_Side), convex hull area (CH), plant height, revealed that drought led to a significant decrease in all these measured morphological traits, while the trait solidity increased in drought condition. The study by Kim et al. (2020) showed that rice grown under drought stress had more

compact structure, resulting in lower CH and higher solidity. This is because drought limited plant growth and expansion, resulting in plants growing densely in available space. The same pattern of results was also found in our study since pea plants grown in drought had more compact structure than in control conditions (Figure 3.4). This result was supported by the fact that pea plants subjected to drought conditions showed significantly lower total fresh weight (FW) and dry weight (DW) compared to those grown under control conditions ($p < 0.001$, Table 3A.1, and Figure 3.2).

Furthermore, a high positive correlation, with correlation coefficients exceeding 0.7, was observed between morphological traits derived from RGB sideview traits and total harvested biomass. Notably, DB_Side exhibited a higher correlation coefficient than DB_Top for both FW and DW (Figure 3.7). According to the results, sideview images are more effective at determining biomass of pea plants than topview images, since pea plants grown in the pot vertically climbed the stick by using tendrils and less spread in horizontal direction. The advantages of sideview over topview images for plant phenotyping was described by Hati et al. (2023) that sideview images can capture vertical shoot growth, while top-view images do not always capture all the features due to overlapping leaves. Hence the performance of sideview and top view images may be different when apply to other plant species since they have different architecture, morphology, and structure.

The strong correlations between RGB sideview traits and harvested biomass have also been reported in numerous studies utilizing different plant species, stressors, and automated imaging phenotyping systems. For instance, studies involving safflower under drought (Joshi et al., 2021), wheat under nitrogen deficiency (Banerjee et al., 2020), rice under salinity (Campbell et al., 2015), and pea seedlings in control environments (Nguyen et al., 2018) showed remarkably high correlations between estimated shoot biomass derived from imaging platforms and harvested biomass, with correlation coefficients ranging from approximately 0.85 to 0.95, using the Scanalyzer 3D digital imaging platform. These high correlations between image-based pixel and harvested traits in different plant species, stressors, and high throughput facilities illustrated the consistency and ability of RGB sideview traits for interpreting harvested biomass.

In other traits, the analysis of spectral imaging (SI) traits and chlorophyll fluorescence (CF) traits showed unexpected results. SI traits revealed a decrease in the reflectance of red (R_{Red}), green (R_{Green}), and blue (R_{Blue}), as well as saturation (SAT) and value (VAL), under drought conditions and, hue (HUE) was the only trait that increased in response to drought (Figure 3.5). Surprisingly,

these results partially contradict a previous experiment conducted by Lazarević et al. (2022) on *M. sinensis*, which also investigated the effect of drought stress on plant growth using the same camera system, CropReporter. In the study of *M. sinensis*, R_{Red} and R_{Blue} increased under drought conditions, indicating a decrease in photosynthetic pigments and an increase in yellow and brown pigments, which are typical characteristics of senescing leaves resulting from a variety of stresses (Li et al., 2014). Another contradiction was observed in HUE, which was significantly higher in the drought condition compared to the control condition in this study (Figure 3.5). Typically, leaf color has a HUE range from 60° to 120°, representing the gradient from yellow to green (Rezzouk et al., 2020; Genangeli et al., 2023). Therefore, plants exposed to drought stress should exhibit a lower HUE than those in the control condition. Kim et al. (2020) established criteria for distinguishing plants grown under control and drought conditions by using HUE which proved to be effective in differentiating between the rice grown under control and drought conditions (Kim et al., 2020).

Furthermore, surprising results were also found with regard to the CF traits. F_v/F_m in pea plants grown in drought conditions was significantly higher than in those grown under control conditions (Figure 3.6). Although F_v/F_m is widely used as a physiological parameter for estimating plant performance under stressful conditions, several authors have reported that it is not sensitive to mild or moderate water stress (Bukhov and Carpentier, 2004; Massacci et al., 2008). The reaction centers of PSII are not affected by drought until the stress becomes severe, due to the presence of various photoprotective mechanisms (Sommer et al., 2023). Lazarević et al. (2022) found that F_v/F_m in *M. sinensis* only decreased after prolonged and severe water deficiency stress, which was as long as two or three weeks of withholding irrigations. Similar results were observed in *Arabidopsis* plants grown under drought conditions, where a significant difference in F_v/F_m was observed after seven days of drought stress (Yao et al., 2018). In other pea experiment, a significant decrease of F_v/F_m, was found after four and five days without irrigation (Couchoud et al., 2020). Another type of water regime in drought treatment was to maintain low amount of water in drought condition. For example, in hybrid poplar experiment, water was maintained at 100% and lower than 50% of field capacity (FC) for control and drought treatment, respectively, and the experiment was applied for several weeks to show the significant different in F_v/F_m between control and drought conditions (Guo et al., 2010). An experiment in pigeon pea maintained 100% and 50% FC during the whole experiment but withheld irrigation one week before measurement (Narina et al.,

2014). Compared to our study, water was still applied during the drought treatment in the whole experiment, maintaining 30% FC for drought and 70% FC for the control condition which may not have been severe enough to cause damage to photosystem II (PSII) and affect F_v/F_m under drought conditions as expected.

Additionally, the irrigation method could be another factor contributing to the unexpected results with regard to F_v/F_m. There were two compartments in this experiment and pot weight from each compartment was determined approximately every other day to irrigate and maintain 70% FC and 30% FC for the control and drought treatment, respectively. The pot weight and cumulative weight of water added showed that water was applied in drought was more consistent than in control since there was a sharply fluctuating pattern of pot weight in control condition (Figure 4A.1.1). The pattern showed that the pot weight of plants grown in control conditions reduced faster than the those grown in drought conditions (Figure 4A.1.1), meaning that plants grown under control conditions used more water for transpiration consistent with the sharply increasing in the amount of cumulative water added in plants grown under conditions. Hence, the frequency of irrigation may not have been sufficient for maintaining water at a constantly higher level in the control treatment, and the transpiration of plants grown under control conditions could have been limited by the soil water availability at times, leading to a decrease in transpiration rate and subsequently a decrease in photosynthetic rate. The decrease in photosynthesis rate could have affected the decrease in reflectance of SI and CF traits observed in this study (Figure 3.5 and Figure 3.6). In addition, the pot weight of the control conditions almost reached the weight of 30% FC as in drought treatment. This pattern of pot weight was observed about a week before the image acquisition. This suggests that the severity of drought in this experiment varied among plants. Therefore, the frequency and amount of irrigation could have affected transpiration and other physiological mechanisms of plants grown under control treatment, resulting in the unexpectedly lower F_v/F_m for control plants found in this experiment.

Generally, the changes in SI traits indicate alterations in pigment content and plant senescence (Lazarević et al., 2022). In this study, another potential explanation for unexpected results in SI and CF traits is the senescence of pea plants grown under control conditions due to reaching the reproductive phase. From the pot weight and cumulative water added data, pea plants grown under control conditions received more water than those in the drought treatment (Figure 4A.1.1). Hence,

they reached their reproductive stage faster than those grown under drought conditions. This is further supported by the observation that some pea plants in control conditions were already flowering and producing pea seeds before imaging (Figure 4A.2.1). Plant senescence and stress can be visible as a loss of photo pigments, such as anthocyanins, xanthophylls, carotenoids, and chlorophyll, which characterize leaf color (Gregersen et al., 2014; Winker-Shirley, 2015). Experiments monitoring pea growth have found that the average of three vegetation indices, such as green-red vegetation index (GRVI), normalized difference vegetation index (NDVI), and normalized difference red-edge index (NDRE), decreased as plants approached physiological maturity and senescence (Vargas et al., 2019). Similar results were observed in an experiment monitoring rice growth using vegetation indices (VIs) (Ang et al., 2020). Moreover, a very high positive relationship between VIs and SPAD was found in this experiment, indicating a decrease in VIs related to chlorophyll content in the leaves (Ang et al., 2020). Consistent with our results, the reflectance of colors decreased, and the lower HUE, representing a less green gradient color, in pea plants grown under control conditions could be the consequence of changes in chlorophyll pigments due to senescence after reaching the reproductive stage.

To summarize, the results from this experiment showed more pronounced changes in RGB sideview traits related to morphological traits, such as DB_Top and DB_Side, while SI and CF traits showed unexpected results contrary to expectations. Reduction in plant biomass is a typical response to drought (Farooq et al., 2009), as it reduces metabolite requirements (Chaves et al., 2003) and plays a crucial role in the regulation of heat dissipation and transpiration (Blum, 2005). Several possible reasons for these unexpected results in spectral traits and F_v/F_m include the limitation of water regimes during the experiment, the frequency of irrigations, and the stage of plant growth and development during image acquisition. Further investigations and adjustments in experimental design and conditions may provide insights into the underlying factors contributing to these unexpected observations.

4.2 RGB sideview traits providing the highest accuracy for biomass estimation model

An image-based phenotypic trait has been applied to estimate biomass in agriculture and ecology. This approach enables the non-destructive assessment of plant response to the environment over time and allows the determination of plant biomass without the need to harvest the entire plant

(Rahaman et al., 2017). Several biomass estimation models based on image-based phenotypic traits have been derived using linear methods (Golzarian et al., 2011). In this study, biomass estimation models were developed using multiple linear regression methods. Each model was derived from different three types of image-based phenotypic traits, including RGB sideview, spectral imaging (SI), and chlorophyll fluorescence (CF) traits. The results showed that using RGB sideview traits as predictor variables yielded the best biomass estimation model, with an approximate R^2 value of 0.7 and RMSE values of approximately 4.7 and 0.7 for fresh (FW) and dry weight (DW), respectively (Table 3.1). These results were consistent with previous studies, demonstrating high performance and accuracy of biomass estimation model deriving from RGB sideview traits (Tackenberg, 2007; Golzarian et al., 2011). Furthermore, there are various ways to improve biomass estimation models by using other morphological traits. For example, a study on common pea utilized plant height and width instead of plant area to form the biomass estimation model (Barboza et al., 2023). Another research on rice suggested including compactness in the model to account for the compact structure of rice plants (Elangovan et al., 2023). However, predicting biomass using only RGB sideview traits poses a significant challenge due to the diverse crop architectures of different plant species (Elangovan et al., 2023).

There was a substantial difference in the accuracy of biomass estimation models derived from RGB sideview traits compared to other traits. The model derived from CF traits ranked as the second-best biomass estimation model. However, its performance and accuracy were comparable to the model derived from SI traits (Table 3.1). Although SI and CF traits are related to photo pigments that affect photosynthesis and physiological mechanisms of plants, weak correlations were found between harvested biomass and CF traits, while there was almost no significant correlation between harvested biomass and SI traits in this study (Figure 3.7). These findings provide supportive evidence for the lower performance of biomass estimation models derived from these two traits.

Although many studies have utilized SI traits to study drought responses in plants, only a few directly interpret the results from visible reflectance. The common approach involves converting the color and other regions of reflectance into vegetation indices, which are widely used in plant phenotyping to estimate plant biophysical and biochemical traits (Koh et al., 2022). Vegetation indices (VIs) derived from the combination of RGB and near infrared (NIR) spectra have shown

effective results in detecting plant responses to drought in various species such as pea and rice (Nemeskéri et al., 2015; Kim et al., 2020; Javornik et al., 2023; Tafesse et al., 2022). Since this study only imaged the visible spectrum, it is limiting to quantify the various common types of VIs for plants' responses to drought. Therefore, considering a wider range of wavelengths, including visible and NIR, could be a potential approach to improve the understanding of plant responses under drought stress, rather than relying solely on the visible spectrum. Using VIs for estimating biomass has been successful in several research studies. For example, a field experiment on spring and winter pea using different types of VIs for biomass estimation models revealed a high positive correlation between harvested and estimated biomass (Vargas et al., 2019). There were other similar experiments were conducted on different plant species, such as wheat, oat, common bean, soybean, and rice, that also a high positive correlation between harvested and estimated biomass (Wang et al., 2016; Coelho et al., 2018; Carneiro et al., 2022; Barboza et al., 2023; Elangovan et al., 2023).

RGB sideview traits are generally used for biomass estimation models, as they directly reflect the projected plant area of plants from different views, which is closely related to biomass. The performance of each view of images depends on plant architecture and structure. Therefore, using RGB sideview traits to derive biomass estimation models is a promising and effective approach. However, there are other traits that can play a role in plant growth and influence accumulated biomass. For example, SI and CF traits are related to alterations in leaf pigments, which in turn affect photosynthesis processes. Hence, combining different types of image-based phenotypic traits could improve the performance of biomass estimation models. In conclusion, while RGB sideview traits or plant area serve as effective predictors for biomass estimation models, other traits related to pigment alteration and photosynthesis processes can also contribute to plant growth and accumulated biomass. The incorporation of a diverse range of image-based phenotypic traits, including VIs derived from various wavelengths, presents an opportunity to enhance the performance and understanding of biomass estimation models in different plant species and under various environmental conditions.

4.3 The slight improvement of biomass estimation model with multiple types of image-based traits

In this study, the biomass estimation models were enhanced by combining multiple types of image-based traits instead of relying on a single type. The results demonstrated that the combination of multiple image-based traits slightly improved the accuracy of the biomass estimation models when RGB sideview traits were included in the developed model. The best models were achieved by combining RGB sideview and spectral imaging (SI) traits yielding an approximate R^2 value of 0.75 and 0.79, and an RMSE in testing data value of 4.30 and 0.68 for fresh weight (FW) and dry weight (DW), respectively (Table 3.2). The second-best model was achieved by combining all traits yielding the same values of R^2 for both FW and DW, and an RMSE in training data. Only RMSE in testing data was slightly different which value of 4.31 (Table 3.2). While, using RGB sideview trait alone had an R^2 value of approximately 0.70 and 0.76, and an RMSE value of 4.78 and 0.73 for FW and DW, respectively. Similarly, the model combining RGB sideview and chlorophyll fluorescence (CF) traits also displayed improved accuracy in the biomass estimation model (Table 3.2). Only the combination of SI and CF traits did not exhibit improvement, likely due to their very low correlations with harvested biomass compared to the RGB sideview traits (Figure 3.7).

There have been several studies exploring the combination of different types of traits to enhance biomass estimation and improve phenotyping accuracy. For instance, a study on common bean utilized plant height, width from destructive harvest, and vegetative indices (VIs), revealing that the model combining these traits yielded the highest accuracy compared to models using either trait alone (Barboza et al., 2023). However, it is important to note that this combination still required the inclusion of destructively harvested traits. In another study on soybean, the fusion of multispectral and thermal data was found to provide the best results for biomass estimation (Maimaitijiang et al., 2017). Additionally, this combination demonstrated the best estimates for nitrogen concentration, chlorophyll a content, and leaf area index (LAI) predictions, with the fusion of RGB and thermal data proving most effective (Maimaitijiang et al., 2017). Research focused on the combination of SI and CF to identify rice under stress conditions indicated that fusing visible and near-infrared spectral traits with CF traits had the potential to enhance plant stress type identification performance (Zhang et al., 2022). Furthermore, the combination of different traits has shown improvement in plant disease detection techniques. A study on wheat found that combining CF with SI traits enhanced disease detection capabilities for yellow rust in

wheat (Moshou et al., 2005). Although the combination of different types of image-based phenotypic traits for biomass estimation has received limited attention, most studies investigating trait combinations have demonstrated improvements in various areas, including disease detection, and nitrogen and chlorophyll content estimation. These findings highlight the potential of employing combinations of different image-based phenotypic traits to enhance the efficiency of high-throughput phenotyping, leading to increased accuracy.

4.4 Future aspects

The high-throughput plant phenotyping platform is a powerful tool for enhancing and accelerating phenotyping, which is a bottleneck in plant breeding programs. The usefulness of image-based phenotypic traits has been confirmed by numerous studies conducted on different plant species under various stress conditions. However, recent research suggests that the true bottleneck lies not in data acquisition but in data analysis of the vast datasets generated by high-throughput phenotyping in a short period (Campbell et al., 2018; Yang et al., 2020). This study aimed to improve biomass estimation models by combining multiple types of image-based phenotypic traits using the dataset obtained from the experiment. Such an approach has the potential to increase the value of information derived from the image-based phenotypic traits. When studying drought conditions with automatic high throughput phenotyping facilities, caution must be exercised in the irrigation process, as the frequency and amount of irrigation can impact plant growth and development, although this in itself can provide valuable insights for further study.

Additionally, the timing of measurements must be carefully considered, as it varies among different plant species and the question to be addressed. For instance, research on rice biomass estimation revealed that the accuracy of the model decreased when rice was grown for more than six weeks due to the development of a more complex structure (Elangovan et al., 2023). Pea plants, with their climbing characteristics and ability to grow in multiple directions, possess a more complex structure. Understanding the architecture of pea plants and determining suitable ages or times for imaging could help alleviate the issue of leaf overlapping and enhance the usefulness of image-based phenotypic traits.

While this study showed slight improvement in biomass estimation by combining different types of image-based phenotypic traits in a model, it indicates a positive direction for enhancing the

capabilities of image-based phenotypic traits, leaving room for further improvement. Exploring combinations of traits not included in this experiment, such as vegetation indices, thermal imaging, hyperspectral reflectance, 3D imaging, and short-wave infrared (SWIR) spectrum, could show the potential for more accurate biomass estimation. Thermal and SWIR spectra are closely related to temperature and water content in plants, providing plants water status and ability to cope drought stress which are useful for studying drought. Integrating these traits with existing data has the potential to enhance biomass estimation.

Another suggestion is to consider accumulated data rather than specific data since biomass represents the total accumulation of plant growth until the measurement date. Research by others on peas demonstrated that forming biomass estimation models based on the summation of vegetation indices (VIs) yielded higher accuracy compared to averaging or using single time points of VIs (Vargas et al., 2019). Therefore, acquiring image data with time series information could further improve biomass estimation models, enabling the monitoring of growth and development. In addition, time series data could improve understanding plant responses and detect early responses to drought, which help in the selection of appropriate time points and traits for accurate biomass estimation.

Chapter 5: Conclusion

In conclusion, the high-throughput plant phenotyping platform holds great promise for addressing the bottleneck of phenotyping in plant breeding programs. Image-based phenotyping provides valuable insights and enormous data with rapid results, reduced labor, and non-destructive harvesting. This study detected the effects of drought stress on pea growth using multimodal camera systems. Significant differences in image-based phenotypic traits were observed between pea plants grown under control and drought conditions. Moreover, the biomass estimation model formulated by RGB sideview traits, which relates to morphological and architectural traits, exhibited the highest accuracy compared to other models. The combination of spectral imaging traits, and chlorophyll fluorescence, or both, can further enhance the accuracy of biomass estimation models. Although the improvement in model accuracy was marginal, it represents a positive step towards interpreting the vast amount of data derived from image-based phenotypic traits. By exploring new trait combinations, such as vegetation indices and thermal imaging, we can unlock the potential for more robust biomass estimation, particularly in studying stress conditions. Continued advancements in high-throughput phenotyping techniques and the refinement of biomass estimation models are crucial for enhancing agricultural practices and increasing crop yield. Improving phenotyping models is important in agriculture as biomass production correlates with crop yield. By using the power of these innovative technologies, we pave the way for more efficient approaches to plant breeding and crop production.

For further study of this pea experiment, the biomass estimation model developed in this study can be applied to estimate the biomass of pea plants that were not harvested during this experiment. Furthermore, the concept of combining different types of image-based phenotypic traits can be extended to the study of other plant traits, such as leaf area. Image-based phenotypic traits can be used to investigate variations among pea genotypes in terms of their morphological and physiological characteristics under drought stress. Examining the relationship between these traits and biomass could enhance our understanding of how different genotypes of plants respond to drought. Moreover, image-based traits can serve as supporting evidence for quantifying productive genotypes, enabling us to identify the specific traits that should be focused on to improve drought tolerance in pea breeding programs.

References

Anandan, A., Mahender, A., Sah, R., Bose, L., Subudhi, H., Meher, J., . . . Ali, J. (2020). Non-destructive phenotyping for early seedling vigor in direct-seeded rice. *Plant Methods*, 16:127. doi:10.1186/s13007-020-00666-6

Ang, Y., Norasma, N., Ya, C., Roslin, N., Ismail, M., & Che'Ya, N. (2020). Rice chlorophyll content monitoring using vegetation indices from multispectral aerial imagery. *Pertanika Journal of Science and Technology*, *28*(3), 779-795.

Baker, N. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annual Review of Plant Biology, 59*, 89-113. doi:10.1146/annurev.arplant.59.032607.092759

Banerjee, B., Joshi, S., Thoday-Kennedy, E., Pasam, R., Tibbits, J., Hayden, M., . . . Kant, S. (2020). High-throughput phenotyping using digital and hyperspectral imaging-derived biomarkers for genotypic nitrogen response. *Journal of Experimental Botany*, *71*(15), 4604–4615. doi:10.1093/jxb/eraa143

Barboza, T., Ardigueri, M., Souza, G., Ferraz, M., Gaudencio, J., & Santos, A. (2023). Performance of vegetation indices to estimate green biomass accumulation in common bean. *AgriEngineering, 5*, 840–854. doi:10.3390/agriengineering502

Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, *114*, 1312-1323.

Bhandari, A. (2023, April 26). *Multicollinearity | Causes, Effects and Detection Using VIF (Updated 2023)*. Retrieved from Analytics Vidhya: https://www.analyticsvidhya.com/blog/2020/03/what-is-multicollinearity/

Blum, A. (2005). Drought resistance, water use efficiency, and yield potential—Are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research, 56*, 1159-1168. doi:10.1071/AR05069

Bukhov, N., & Carpentier, R. (2004). Effects of water stress on the photosynthetic efficiency of plants. *Advances in Photosynthesis and Respiration*, 19. doi:10.1007/978-1-4020-3218-9_24

Campbell, M., Knecht, A., Berger, B., Brien, C., Wang, D., & Walia, H. (2015). Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiology, 168*, 1476-1489.

Campbell, Z., Acosta-Gamboa, L., Nepal, N., & Lorence, A. (2018). Engineering plants for tomorrow: how high throughput phenotyping is contributing to the development of better crops. *Phytochemistry reviews*, *17*(6), 1329-1343.

Carneiro, F., Oliveira, M., Almeida, S., Brito Filho, A., Furlani, C., Rolim, G., . . . Silva, R. (2022). Biophysical characteristics of soybean estimated by remote sensing associated with artificial intelligence. *Bioscience Journal*, e38024. doi:10.14393/BJ-v38n0a2022-55925

Cernay, C., Ben-Ari, T., Pelzer, E., Meynard, J.-M., & Makowski, D. (2015). Estimating variability in grain legume yields across Europe and the Americas. *Scientific Reports, 5:11171*. doi:10.1038/srep11171

Chaerle, L., Valcke, R., & Van Der Straeten, D. (2003). Imaging techniques in plant physiology and agronomy: From simple to multispectral approaches. In *In Plant Physiology and Plant Molecular Biology in the New Millennium; Hemantaranjan* (Vol. 5, pp. 135-155). Jodhpur, India: A., Ed.; Scientific Publisher.

Chaves, M., Maroco, J., & Pereira, J. (2003). Understanding plant responses to drought — from genes to the whole plant. *Functional Plant Biology, 30*, 239–264. doi:10.1071/FP02076

Coelho, A., Rosalen, D., & de Faria, R. (2018). Vegetation indices in the prediction of biomass and grain yield of white oat under irrigation levels. *Pesquisa Agropecuária Tropical, 48*(2), 109-117.

Cotrozzi, L., & Couture, J. (2020). Hyperspectral assessment of plant responses to multi-stress environments: Prospects for managing protected agrosystems. *Plants People Planet, 2*, 244-258. doi:10.1002/ppp3.10080

Couchond, M., Salon, C., Girodet, S., Jeudy, C., Vernoud, V., & Prudent, M. (2020). Pea efficiency of post-drought recovery relies on the strategy to fine-tune nitrogen nutrition. *Frontiers in Plant Science, 11*. doi:10.3389/fpls.2020.00204

Couture, J., Serbin, S., & Townsend, P. (2013). Spectroscopic sensitivity of real-time, rapidly induced phytochemical change in response to damage. *New Phytologist, 198*, 311–319.

Crain, J., Mondal, S., Rutkoski, J., Singh, R., & Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *Plant Genome, 11*, 1-14. doi:10.3835/plantgenome2017.05.0043

Das Choudhury, S., Bashyam, S., Qiu, Y., Samal, A., & Awada, T. (2018). Holistic and component plant phenotyping using temporal image sequence. *Plant Methods*, 14:35. doi:10.1186/s13007-018-0303-x

Das Choudhury, S., Stoerger, V., Samal, A., Schnable, J., Liang, Z., & Yu, J.-G. (2016). Automated vegetative stage phenotyping analysis of maize plants using visible light images. *KDD Workshop on Data Science for Food, Energy and Water*. San Francisco, CA.

Elangovan, A., Duc, N., Raju, D., Kumar, S., Singh, B., Vishwakarma, C., . . . Chinnusamy, W. (2023). Imaging sensor-based high-throughput measurement of biomass using machine learning models in rice. *Agriculture, 13*, 852. doi:10.3390/agriculture13040852

Enders, T., Dennis, S., Oakland, J., Callen, S., Gehan, M., Miller, N., . . . Hirsch, C. (2019). Classifying cold-stress responses of inbred maize seedlings using RGB imaging. *Plant Direct, 3*(1), e00104. doi:10.1002/pld3.104

Fahad, S., Bajwa, A., Nazir, U., Anjum, S., Farooq, A., Zohaib, A., . . . J., H. (2017). Crop production under drought and heat stress: plant responses and management options. *Frontiers in Plant Science, 8*. doi:10.3389/fpls.2017.01147

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., & Basra, S. (2009). Plant drought stress: effects, mechanisms and management. *Agronomy for Sustainable Development, 29*, 185-212. doi:10.1051/agro:2008021

Feng, X., Yu, C., Chen, Y., Peng, J., Ye, L., Shen, T., . . . He, Y. (2018). Non-destructive determination of shikimic acid concentration in transgenic maize exhibiting glyphosate tolerance using chlorophyll fluorescence and hyperspectral imaging. *Frontiers in Plant Science*, 9:468. doi:10.3389/fpls.2018.00468

Furbank, R., & Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science, 16*(12), 635-644. doi:10.1016/j.tplants.2011.09.005

Genangeli, A., Avola, G., Bindi, M., Cantini, C., Cellini, F., Grillo, S., . . . Gioli, B. (2023). Low-cost hyperspectral imaging to detect drought stress in high-throughput phenotyping. *Plants, 12*, 1730. doi:10.3390/plants12081730

Genty, B., Briantais, J., & Baker, N. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta: General Subjects, 990*, 87–92. doi:10.1016/S0304-4165(89)80016-9

Gitelson, A., Gritz, Y., & Merzlyak, M. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for nondestructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology, 160*, 271–282. doi:10.1078/0176-1617-00887

Gitelson, A., Merzlyak, M., & Chivkunova, O. (2001). Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and Photobiology, 74*, 38–45. doi:10.1562/0031-86552001074<0038: opaneo<2.0.co;2

Gobron, N., Pinty, B., Verstraete, M., & Widlowski, J. (2000). Advanced vegetation indicesoptimized for up-coming sensors: design, performance, and applications. *IEEE Transactions on Geoscience and Remote Sensing, 38*(6), 2489-2505.

Golzarian, M., Frick, R., Rajendran, K., Berger, B., Roy, S., Tester, M., & Lun, D. (2011). Accurate inference of shoot biomass from high-throughput images of cereal plants. *Plant Methods*, 7:2. doi:10.1186/1746-4811-7-2

Gregersen, P., Culetic, A., Boschian, L., & Krupinska, K. (2013). Plant senescence and crop productivity. *Plant Molecular Biology, 82*, 603-622. doi:10.1007/s11103-013-0013-8

Guo, X., Zhang, X., & Huang, Z. (2010). Drought tolerance in three hybrid poplar clones submitted to different watering regimes. *Journal of Plant Ecology, 3*(2), 79-87. doi:10.1093/jpe/rtq007

Hati, A., & Singh, R. (2023). AI-Driven Pheno-Parenting: A deep learning based plant phenotyping trait analysis model on a novel soilless farming dataset. *IEEE Access, 99*, 1-1. doi:10.1109/ACCESS.2023.3265195

Jaghdani, S. J., Jahns, P., & Tränkner, M. (2021). The impact of magnesium deficiency on photosynthesis and photoprotection in *Spinacia oleracea*. *Plant Stress, 2*, 100040. doi:10.1016/j.stress.2021.100040

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). Linear model selection and regularization. In *An introduction to statistical learning with applications in R (2nd ed.)* (pp. 225-236). Springer.

Javornik, T., Carović-Stanko. K, Gunjača, J., Vidak, M., & L. B. (2023). Monitoring drought stress in common bean using chlorophyll fluorescence and multispectral imaging. *Plants, 21*(12(6)), 1386. doi:10.3390/plants12061386

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality and Quantity, 52*(4), 957-1976. doi:10.1007/s11135-017-0584-6

Joshi, S., Thoday-Kennedy, E., Daetwyler, H., Hayden, M., Spangenberg, G., & Kant, S. (2021). High-throughput phenotyping to dissect genotypic differences in safflower for drought tolerance. *PLoS ONE, 16*(7), e0254908. doi:10.1371/journal.pone.0254908

Kalaji, H. M., Schansker, G., Brestic, M., Bussotti, F., Calatayud, A., Ferroni, L., . . . Bąba, W. (2017). Frequently asked questions about chlorophyll fluorescence, the sequel. *Photosynthesis Research, 132*, 13–66. doi:10.1007/s11120-016-0318-y

Kim, S., Kim, N., Lee, H., Lee, E., Cheon, K., Kim, M., . . . Kim, K. (2020). High-throughput phenotyping platform for analyzing drought tolerance in rice. *Planta, 252*, 38. doi:10.1007/s00425-020-03436-9

Kim, Y., Glenn, D., Park, J., Ngugi, H., & Lehman, B. (2011). Hyperspectral image analysis for water stress detection of apple trees. *Computers and Electronics in Agriculture, 77*, 155-160. doi:10.1016/j.compag.2011.04.008

Koh, J., Banerjee, B., Spangenberg, G., & Kant, S. (2022). Automated hyperspectral vegetation index derivation using a hyperparameter optimisation framework for high-throughput plant phenotyping. *New Phytologist, 233*, 2659–2670. doi:10.1111/nph.17947

Kumar, P., Eriksen, R., Simko, I., & Mou, B. (2021). Molecular mapping of water-stress responsive genomic loci lettuce (*Lactuca spp.*) using kinetics chlorophyll fluorescence, hyperspectral imaging and machine learning. *Frontiers in Genetics*, 12:634554,.

Lazarević, B., Kontek, M., Carović-Stanko, K., Clifton-Brown, J., Hassan, M., Trindade, L., & Jurišić, V. (2022). Multispectral image analysis detects differences in drought responses in novel seeded *Miscanthus sinensis* hybrids. *GCB Bioenergy, 14*(11), 1219-1234. doi:10.1111/gcbb.12999

Lazarević, B., Šatović, Z., Nimac, A., Vidak, M., Gunjača, J., Politeo, O., & Carović-Stanko, K. (2021). Application of phenotyping methods in detection of drought and salinity stress in basil (*Ocimum basilicum L.*). *Frontiers in Plant Science*, 12:629441. doi:10.3389/fpls.2021.629441

Leister, D., Varotto, C., Pesaresi, P., Niwergall, A., & Salamini, F. (1999). Large-scale evaluation of plant growth in *Arabidopsis thaliana* by non-invasive image analysis. *Plant Physiology and Biochemistry, 37*(9), 671-678. doi:10.1016/S0981-9428(00)80097-2

Li, G., Wan, S., Zhou, J., Yang, Z., & Qin, P. (2010). Leaf chlorophyll fluorescence, hyperspectral reflectance, pigments content, malondialdehyde and proline accumulation responses of castor bean *Ricinus communis L*. seedlings to salt stress levels. *Industrial Crops and Products, 31*(1), 13-19.

Li, L., Zhang, Q., & Huang, D. (2014). A review of imaging techniques for plant phenotyping. *Sensors (Switzerland*, 20078–20111. doi:10.3390/s141120078

Lichtenthaler, H., & Rinderle, U. (1988). The role of chlorophyll fluorescence in the detection of stress conditions in plants. *C R C Critical Reviews in Analytical Chemistry*, 19:S29–S85. doi:10.1080/15476510.1988.10401466

Lu, C., & Zhang, J. (1999). Effects of water stress on photosystem II photochemistry and its thermostability in wheat plants. *Journal of Experimental Botany, 50*, 1199–1206. doi:10.1093/jxb/50.336.1199

Mahlein, A., Alisaac, E., Masri, A., Behmann, J., Dehne, H., & Oerke, E. (2019). Comparison and combination of thermal, fluorescence, and hyperspectral imaging for monitoring fusarium head blight of wheat on spikelet scale. *Sensors*, 19:2281.

Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., . . . Fritschi, F. (2017). Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *Journal of Photogrammetry and Remote Sensing, 134*, 43-58. doi:10.1016/j.isprsjprs.2017.10.011

Massacci, A., Nabiev, S., Pietrosanti, L., Nematov, S., Chernikova, T., Thor, K., & Leipner, J. (2008). Response of the photosynthetic apparatus of cotton (*Gossypium hirsutum*) to the onset of drought stress under field conditions studied by gas-exchange analysis and chlorophyll fluorescence imaging. *Plant physiology and biochemistry, 46*, 189-195. doi:10.1016/j.plaphy.2007.10.006

Maxwell, K., & Johnson, G. (2000). Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany, 51*, 659–668. doi:10.1093/jexbot/51.345.659

Merzlyak, M., Gitelson, A., Chivkunova, O., & Rakitin, V. (1999). Nondestructive optical detection of pigment changes during leaf senescence and fruit ripening. *Plant Physiology, 106*, 135-141. doi:10.1034/j.1399-3054.1999.106119.x

Michelon, T., Vieira, E., & Panobianco, M. (2023). Spectral imaging and chemometrics applied at phenotyping in seed science studies: a systematic review. *Seed Science Research*, 1-14. doi:10.1017/S0960258523000028

Moshoua, D., Bravoa, C., Oberti, R., West, J., Bodria, L., McCartney, A., & Ramon, H. (2005). Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps. *Real-Time Imaging2, 11*, 75-83.

Murchie, E., & Lawson, T. (2013). Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *Journal of Experimental Botany, 64*, 3983–3998. doi:10.1093/jxb/ert208

Narina, S., Phatak, S., & Bhardwaj, H. (2014). Chlorophyll fluorescence to evaluate pigeonpea breeding lines and mungbean for drought tolerance. *Journal of Agricultural Science, 6*(11). doi:10.5539/jas.v6n11p238

Nemeskéri, E., Molnar, K., Vígh, R., Nagy, J., & Dobos, A. (2015). Relationships between stomatal behaviour, spectral traitsand water use and productivity of green peas (*Pisum sativum L.*) in dry seasons. *Acta Physiologiae Plantarum*, 37:34. doi:10.1007/s11738-015-1776-0

Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., . . . Kilian, B. (2015). Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. *Plant, Cell & Environment, 38*(10), 1980-1996. doi:10.1111/pce.12516

Nguyen, G., Norton, S., Rosewarne, G., James, L., & Slater, A. (2018). Automated phenotyping for early vigour of field pea seedlings in controlled environment by colour imaging technology. *PLoS ONE, 13*(11), e0207788. doi:10.1371/journal.pone.0207788

Nottingham, U. o. (2023, June 20). *Colour spaces*. Retrieved from A description of commonly used colour space and their applications: https://www.futurelearn.com/info/courses/introduction-to-image-analysis-for-plant-phenotyping/0/steps/297747

Pabuayon, I., Sun, Y., Guo, W., & Ritchie, G. (2019). High-throughput phenotyping in cotton: a review. *Journal of Cotton Research, 2*, 18. doi:10.1186/s42397-019-0035-0

Pandey, A., Rubiales, D., Wang, Y., Fang, P., Sun, T., Liu, N., & Xu, P. (2021). Omics resources and omics-enabled approaches for achieving high productivity and improved quality in pea (*Pisum sativum L.*). *Theoretical and Applied Genetics, 134*, 755-776. doi:10.1007/s00122-020-03751-5

Peñuelas, J., Baret, F., & Filella, L. (1995). Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosythetica, 31*, 221-230.

Polder, G., & Gowen, A. (2021). The hype in spectral imaging. *SPECTROSCOPYEUROPE, 33*(3), 12-14.

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science, 360*, 987-992. doi:10.1126/science.aaq021

Porcar-Castell, A., Tyystjärvi, E., Atherton, J., Van der Tol, C., Flexas, J., Pfündel, E., . . . Berry, J. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *Journal of Experimental Botany, 65*, 4065-4095. doi:10.1093/jxb/eru191

Rahaman, M., Ahsan, A., Gillani, Z., & Chen, M. (2017). Digital biomass accumulation using high-throughput plant phenotype data analysis. *Journal of Integrative Bioinformatics, 14*(3), 20170028. doi:10.1515/jib-2017-0028

Reynolds, D., Baret, F., Welcker, C., Bostrom, A., Ball, J., Cellini, F., . . . Tardieu, F. (2019). What is cost-efficient phenotyping? Optimizing costs for different scenarios. *Plant Science, 282*, 14-22. doi:10.1016/j.plantsci.2018.06.015

Rezzouk, F., Gracia-Romero, A., Kefauver, S., Gutiérrez, N., Aranjuelo, I., Serret, M., & Araus, J. (2020). Remote sensing techniques and stable isotopes as phenotyping tools to assess wheat yield performance: Effects of growing temperature and vernalization. *Plant Science, 295*, 110281.

Roitsch, T., Cabrera-Bosquet, L., Fournier, A., Ghamkhar, K., Jiménez-Berni, J., & Francisco, P. (2019). Review: new sensors and data-driven approaches—a path to next generation phenomics. *Plant Science*, 282:2–10.

Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., . . . Plümer, L. (2012). Early drought stress detection in cereals: simplex volume maximization for hyperspectral image analysis. *Functional Plant Biology, 39*, 878-890. doi:10.1071/fp12060

Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). Monitoring vegetation systems in the great plains with ERTS. *Proceedings of the NASA Goddard Space Flight Center 3d ERTS-1 Symp. Sect. A, 1*, pp. 309–317. Washington, DC: NASA.

RStudioTeam. (2020). *RStudio: Integrated Development for R.* Retrieved from RStudio: http://www.rstudio.com

Rubiales, D., Fondevilla, S., Chen, W., Gentzbittel, L., Higgins, T., Castillejo, M., . . . Rispail, N. (2015). Achievements and challenges in legume breeding for pest and disease resistance. *Critical Reviews in Plant Sciences, 34*, 195-236. doi:10.1080/07352689.2014.898445

Serbin, S., Singh, A., Desai, A., Dubois, S., Jablosnki, A., Kingdon, C., . . . Townsend, P. (2015). Remotely estimating photosynthetic capacity, and its response to temperature, in

vegetation canopies using imaging spectroscopy. *Remote Sensing of Environment, 167*, 78-87.

Singh, A., Jones, S., Ganapathysubramanian, B., Sarkar, S., Mueller, D., Sandhu, K., & Nagasubramanian, K. (2021). Challenges and opportunities in machine-augmented plant stress phenotyping. *Trends in Plant Science, 26*(1), 53-69.

Sommer, S., Han, E., Li, X., Rosenqvist, E., & Liu, F. (2023). The chlorophyll fluorescence parameter Fv/Fm correlates with loss of grain yield after severe drought in three wheat genotypes grown at two CO_2 concentrations. *Plants, 12*(3), 436. doi:10.3390/plants12030436

Song, P., Wang, J., Guo, X., Yang, W., & Zhao, C. (2021). High-throughput phenotyping: breaking through the bottlenecks in future crop breeding. *Crop Journal, 9*(3), 633-645.

Tackenberg, O. (2007). A new method for non-destructive measurement of biomass, growth rates, vertical biomass distribution and dry matter content based on digital image analysis. *Annals of Botany, 99*(4), 777-83. doi:10.1093/aob/mcm009

Tafesse, E., Warkentin, T., Shirtliffe, S., Noble, S., & Bueckert, R. (2022). Leaf pigments, surface wax and spectral vegetation indices for heat stress resistance in pea. *Agronomy, 12*, 739. doi:10.3390/agronomy12030739

Tribulato, A., Toscano, S., Di Lorenzo, V., & Romano, D. (2019). Effects of water stress on gas exchange, water relations and leaf structure in two ornamental shrubs in the Mediterranean area. . *Agronomy-Basel*, 9:381. doi:10.3390/agronomy9070381

Van Kooten, O. (1990). The use of chlorophyll fluorescence nomenclature in plant stress physiology. *Photosynthesis Research*, 25:147-150. doi:10.1007/BF00033156

Vargas, J., Zhang, C., Smitchger, J., McGee, R., & Sankaran, S. (2019). Phenotyping of plant biomass and performance traits using remote sensing techniques in pea (*Pisum sativum, L.*). *Sensors, 9*, 2031. doi:10.3390/s19092031

Wang, C., Nie, S., Xi, X., Luo, S., & Sun, X. (2016). Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote sensing, 9*(11). doi:10.3390/rs9010011

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal, 4*(3), 212-219. doi:10.1016/j.cj.2016.01.008

Winkel-Shirley, B. (2015). *Current Opinion in Plant Biology, 5*, 218.

Xu, Z., Zhou, G., & Shimizu, H. (2010). Plant responses to drought and rewatering. *Plant Signaling & Behavior, 5*, 649-654. doi:10.4161/psb.5.6.11398

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J., Batchelor, W., . . . Yan, J. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges and future perspectives. *Molecular Plant, 13*(2), 187-214.

Yao, J., Sun, D., Cen, H., Xu, H., Weng, H., Yuan, F., & He, Y. (2018). Phenotyping of arabidopsis drought stress response using kinetic chlorophyll fluorescence and multicolor fluorescence imaging. *Frontiers in Plant Science, 9*. doi:10.3389/fpls.2018.00603

Zhang, C., Zhou, L., Xiao, Q., Bai, X., Wu, B., Wu, N., . . . Feng, L. (2022). End-to-end fusion of hyperspectral and chlorophyll imaging to identify rice stresses. *Plant Phenomics*, 2022:9851096. doi:10.34133/2022/9851096.

Appendix

Appendix 3A.1: the effect of drought on harvested biomass and image-based phenotypic traits

Table 3A.1.1 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for biomass-related traits. (F and *p* values significant at a level of $p < 0.05$).

Trait	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
Total fresh weight (g)	2.339 (<0.001)	*	371.687 (<0.001)	*	0.927 (0.655)	ns
Total dry weight (g)	2.397 (<0.001)	*	266.483 (<0.001)	*	1.105 (0.331)	ns

[1] The symbols showing in the table are significant level according to *p* value, '*' is significant, 'ns' is not significant.

Table 3A.1.2 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for biomass image-based traits. (F and *p* values significant at a level of $p < 0.05$).

RGB sideview traits (#)	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
Top-view digital biomass	4.386 (<0.001)	*	230.502 (<0.001)	*	1.397 (0.002)	*
Side-view digital biomass	12.24 (<0.001)	*	1161.14 (<0.001)	*	1.56 (<0.001)	*
Side-view convex hull	12.60 (<0.001)	*	814.97 (<0.001)	*	12.60 (<0.001)	*
Side-view height	10.36 (<0.001)	*	266.70 (<0.001)	*	1.59 (<0.001)	*
Side-view solidity	15.77 (<0.001)	*	27.07 (<0.001)	*	1.53 (<0.001)	*

[1] The symbols showing in the table are significant level according to *p* value, '*' is significant, 'ns' is not significant.

Table 3A.1.3 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for color-based and spectral traits. (F and *p* values significant at a level of $p < 0.05$).

Color-based and spectral traits	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
Rred	5.153 (<0.001)	*	63.805 (<0.001)	*	1.439 (0.001)	*
Rgreen	6.668 (<0.001)	*	97.783 (<0.001)	*	1.393 (0.0012)	*
Rblue	5.081 (<0.001)	*	7.768 (0.006)	*	1.478 (<0.001)	*
HUE	7.578 (<0.001)	*	153.806 (<0.001)	*	1.498 (<0.001)	*
SAT	13.805 (<0.001)	*	297.635 (<0.001)	*	1.242 (0.029)	*
VAL	6.479 (<0.001)	*	92.563 (<0.001)	*	1.384 (0.002)	*

[1] The symbols showing in the table are significant level according to *p* value, '*' is significant, 'ns' is not significant.

Table 3A.1.4 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for chlorophyll fluorescence traits. (F and *p* values significant at a level of $p < 0.05$).

Chlorophyll fluorescence traits	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
F_0	7.331 (< 0.0001)	*	5.878 (0.0156)	*	1.060 (0.3030)	ns
F_m	5.174 (< 0.0001)	*	0.170 (0.6799)	ns	1.188 (0.0666)	ns
F_v/F_m	2.176 (< 0.0001)	*	24.954 (< 0.0001)	*	1.203 (0.0528)	ns

[1] The symbols showing in the table are significant level according to *p* value, '*' is significant, 'ns' is not significant.

Appendix 3A.2: the effect of drought on each side-view pixel traits

Table 3A.2.1 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for side-view digital biomass. (F and *p* values significant at a level of $p < 0.05$).

Digital biomass	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
Side-view at 0º	11.36 (<0.001)	*	1080.80 (<0.001)	*	1.61 (<0.001)	*
Side-view at 60º	10.83 (<0.001)	*	1002.77 (<0.001)	*	1.53 (<0.001)	*
Side-view at 120º	10.41 (<0.001)	*	967.00 (<0.001)	*	1.35 (0.039)	*
Side-view at 180º	10.49 (<0.001)	*	934.92 (<0.001)	*	1.44 (0.006)	*
Side-view at 240º	10.73 (<0.001)	*	1004.36 (<0.001)	*	1.55 (<0.001)	*
Side-view at 300º	11.22 (<0.001)	*	1106.15 (<0.001)	*	1.48 (<0.001)	*
Average side-view	12.24 (<0.001)	*	1161.14 (<0.001)	*	1.56 (<0.001)	*
Median of side-view	12.16 (<0.001)	*	1146.82 (<0.001)	*	1.58 (<0.001)	*
Minimum side-view	11.46 (<0.001)	*	1071.14 (<0.001)	*	1.51 (0.001)	*
Maximum side-view	11.82 (<0.001)	*	1137.71 (<0.001)	*	1.50 (0.002)	*

[1] The symbols showing in the table are significant level according to P value, '*' is significant, 'ns' is not significant.

Table 3A.2.2 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for side-view convex hull. (F and *p* values significant at a level of $p < 0.05$).

Convex hull	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
At 0º	9.28 (<0.001)	*	588.62 (<0.001)	*	1.67 (0.001)	*
At 60º	9.47 (<0.001)	*	583.56 (<0.001)	*	1.45 (0.005)	*
At 120º	9.16 (<0.001)	*	569.53 (<0.001)	*	1.45 (0.006)	*
At 180º	9.27 (<0.001)	*	579.93 (<0.001)	*	1.67 (<0.001)	*
At 240º	9.42 (<0.001)	*	597.13 (<0.001)	*	1.58 (<0.001)	*
At 300º	9.52 (<0.001)	*	601.00 (<0.001)	*	1.58 (<0.001)	*
Average	12.60 (<0.001)	*	814.97 (<0.001)	*	12.60 (<0.001)	*
Median	12.38 (<0.001)	*	804.04 (<0.001)	*	12.38 (<0.001)	*
Minimum	10.54 (<0.001)	*	694.69 (<0.001)	*	1.71 (<0.001)	*
Maximum	11.18 (<0.001)	*	684.45 (<0.001)	*	1.68 (<0.001)	*

[1] The symbols showing in the table are significant level according to P value, '*' is significant, 'ns' is not significant.

Table 3A.3 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for side-view plant height. (F and *p* values significant at a level of $p < 0.05$).

Height	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
At 0º	10.36 (<0.001)	*	261.86 (<0.001)	*	1.58 (<0.001)	*
At 60º	10.22 (<0.001)	*	257.83 (<0.001)	*	1.57 (<0.001)	*
At 120º	10.15 (<0.001)	*	263.38 (<0.001)	*	1.58 (<0.001)	*
At 180º	10.12 (<0.001)	*	268.57 (<0.001)	*	1.56 (<0.001)	*
At 240º	10.15 (<0.001)	*	266.77 (<0.001)	*	1.60 (<0.001)	*
At 300º	10.41 (<0.001)	*	271.02 (<0.001)	*	1.60 (<0.001)	*
Average	10.36 (<0.001)	*	266.70 (<0.001)	*	1.59 (<0.001)	*
Median	10.38 (<0.001)	*	265.33 (<0.001)	*	1.59 (<0.001)	*
Minimum	10.60 (<0.001)	*	262.74 (<0.001)	*	1.60 (<0.001)	*
Maximum	9.74 (<0.001)	*	266.14 (<0.001)	*	1.54 (<0.001)	*

[1] The symbols showing in the table are significant level according to P value, '*' is significant, 'ns' is not significant.

Table 3A.4 A two-way ANOVA of genetic effect, treatment, and their two-way interaction for side-view solidity. (F and *p* values significant at a level of $p < 0.05$).

Solidity	F-value (Probability of significance)					
	Genetic effect (G)		Treatment (T)		Interaction (G x T)	
At 0º	8.46 (<0.001)	*	12.39 (0.005)	*	1.24 (0.028)	*
At 60º	9.05 (<0.001)	*	16.17 (<0.001)	*	1.54 (<0.001)	*
At 120º	8.33 (<0.001)	*	12.12 (0.005)	*	0.89 (0.821)	ns
At 180º	8.30 (<0.001)	*	16.62 (<0.001)	*	1.22 (0.040)	*
At 240º	9.16 (<0.001)	*	18.94 (<0.001)	*	1.41 (0.001)	*
At 300º	9.41 (<0.001)	*	12.37 (0.005)	*	1.03 (0.378)	ns
Average	15.77 (<0.001)	*	27.07 (<0.001)	*	1.53 (<0.001)	*
Median	14.38 (<0.001)	*	20.70 (<0.001)	*	1.38 (0.002)	*
Minimum	11.25 (<0.001)	*	4.91 (0.027)	*	1.32 (0.007)	*
Maximum	14.24 (<0.001)	*	47.43 (<0.001)	*	1.65 (<0.001)	*

[1] The symbols showing in the table are significant level according to P value, '*' is significant, 'ns' is not significant.

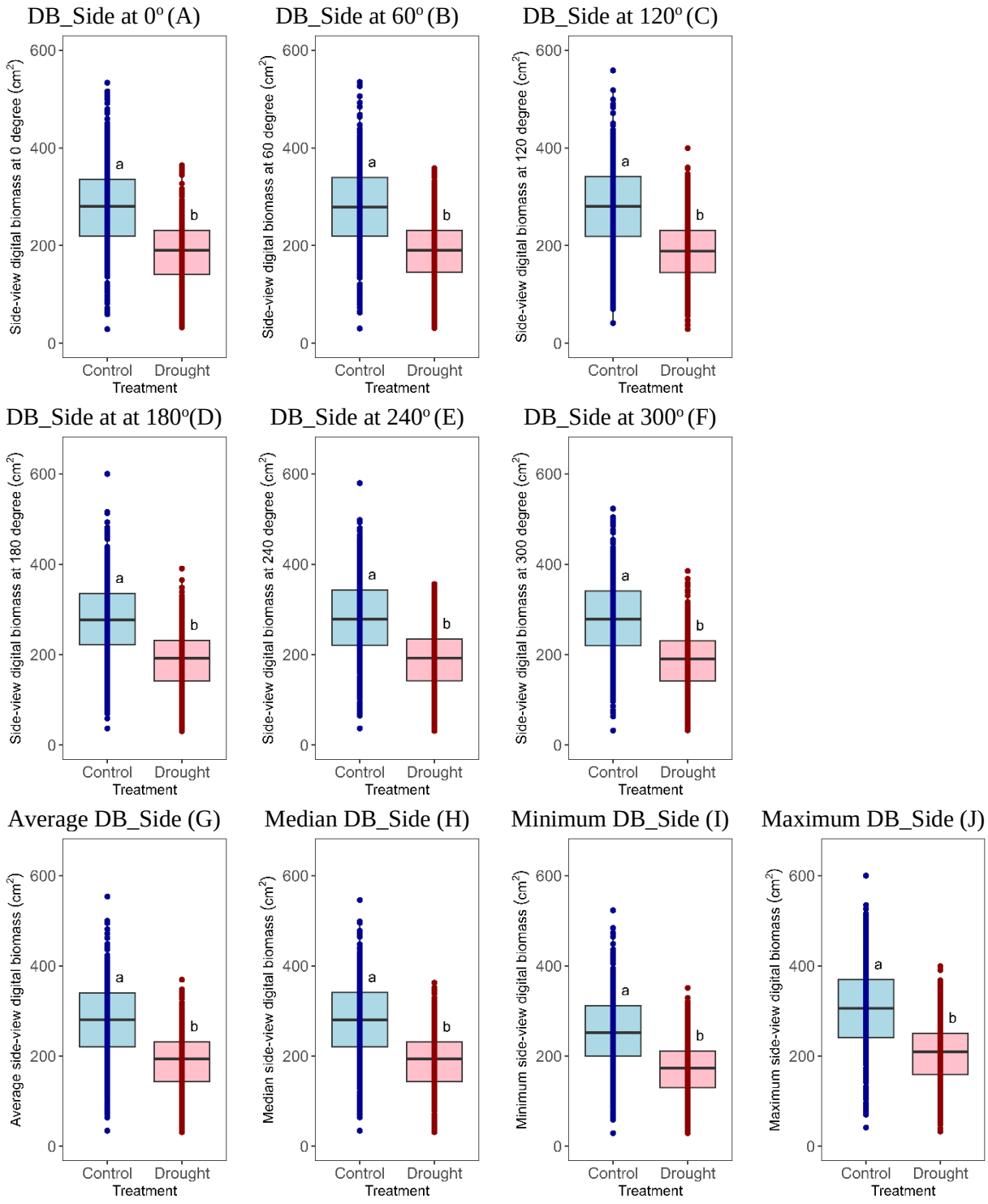


Figure 3A.2.1 Box plots of each side-view digital biomass (DB) in cm^2 of pea grown under controlled (blue) and drought-stressed (red) conditions at six different angles; 0° (A), 60 ° (B), 120 ° (C), 180 ° (D), 240 ° (E), 300 ° (F), average (G), median (H), minimum (I), and maximum (J). The letter above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

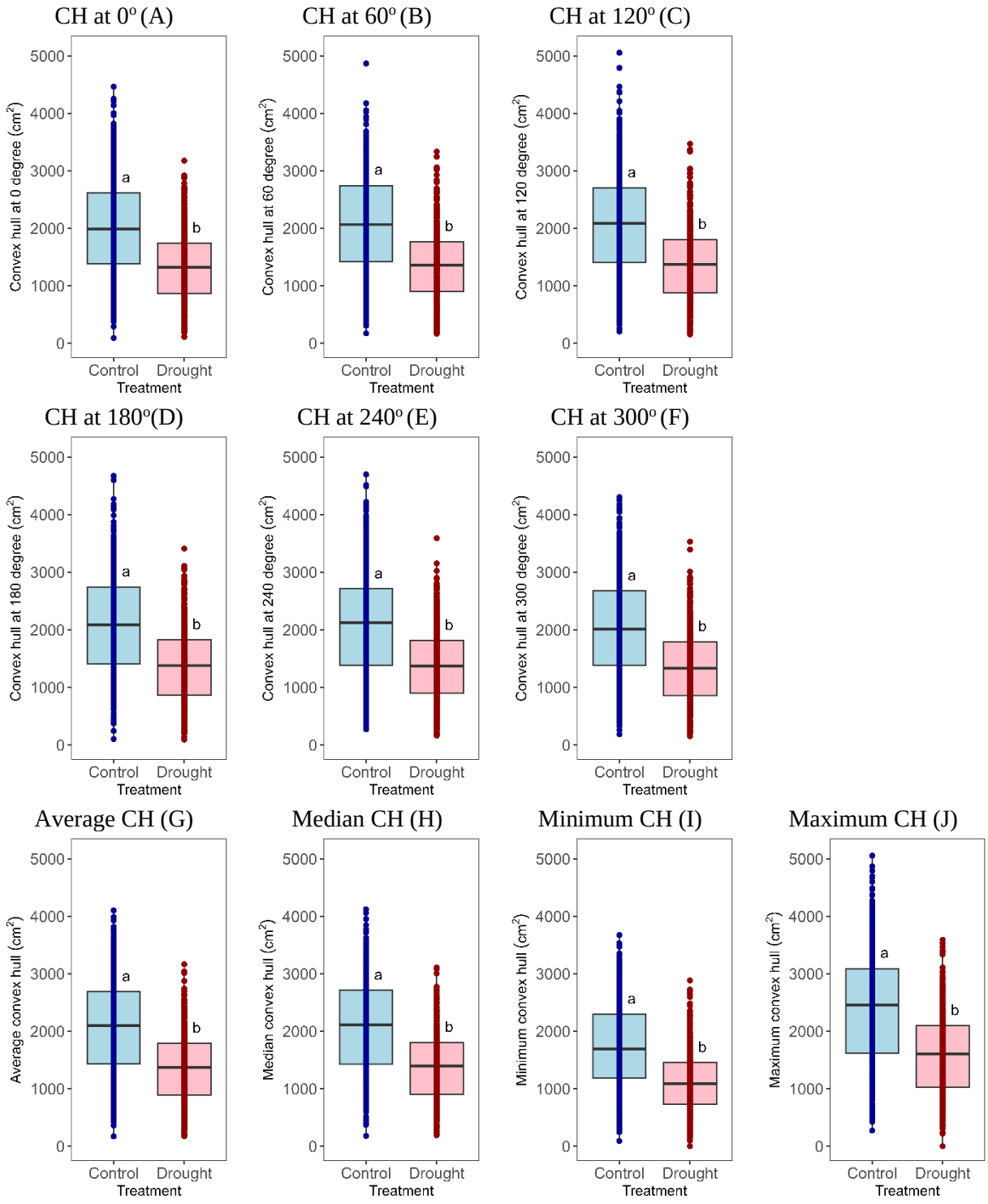


Figure 3A.2.2 Box plots of each side-view convex hull in cm^2 of pea grown under controlled (blue) and drought-stressed (red) conditions at six different angles; 0º (A), 60 º (B), 120 º (C), 180 º (D), 240 º (E), 300 º (F), average (G), median (H), minimum (I), and maximum (J). The letter above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

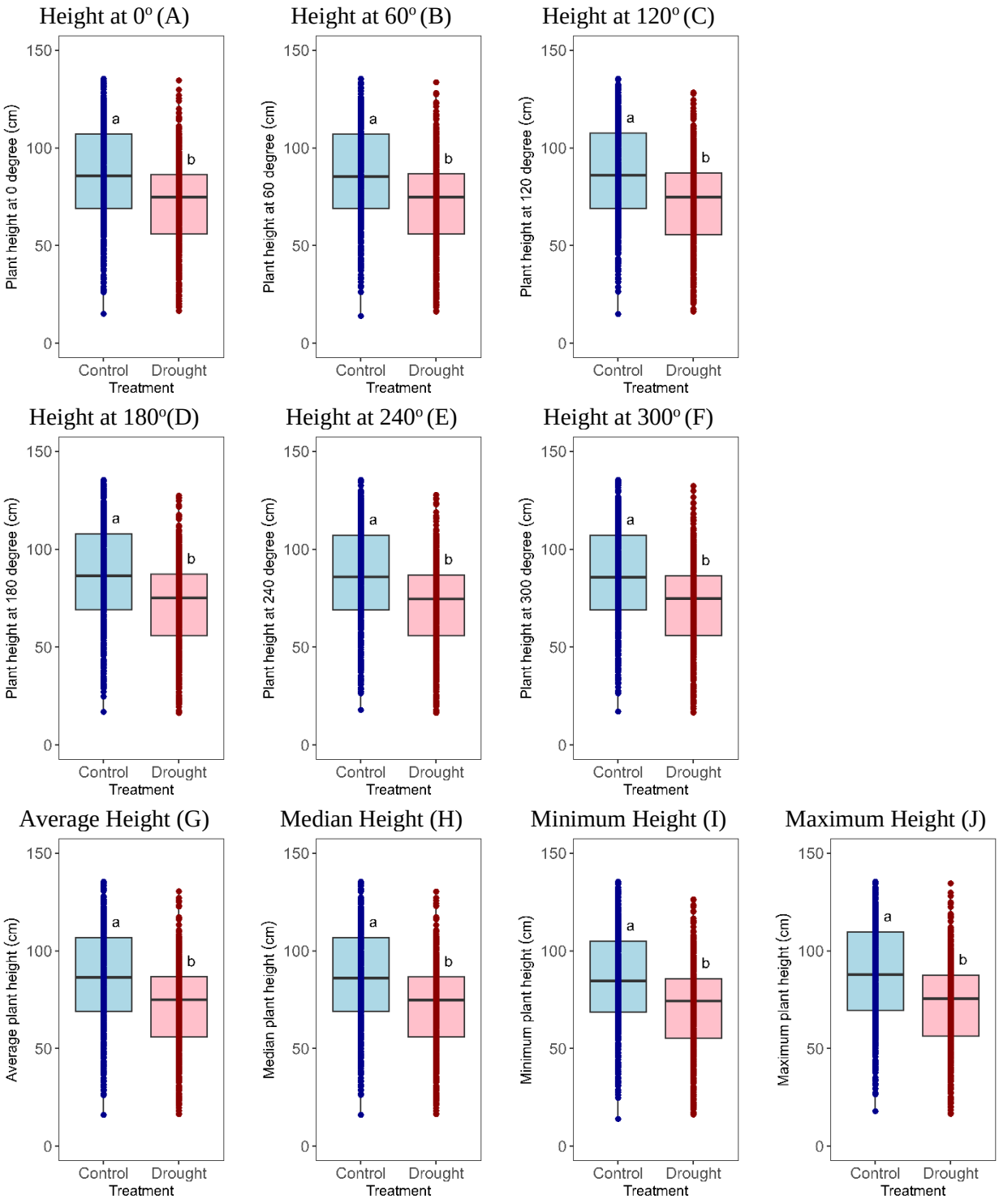


Figure 3A.2.3 Box plots of each side-view plant height in cm of pea grown under controlled (blue) and drought-stressed (red) conditions at six different angles; 0º (A), 60 º (B), 120 º (C), 180 º (D), 240 º (E), 300 º (F), average (G), median (H), minimum (I), and maximum (J). The letter above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

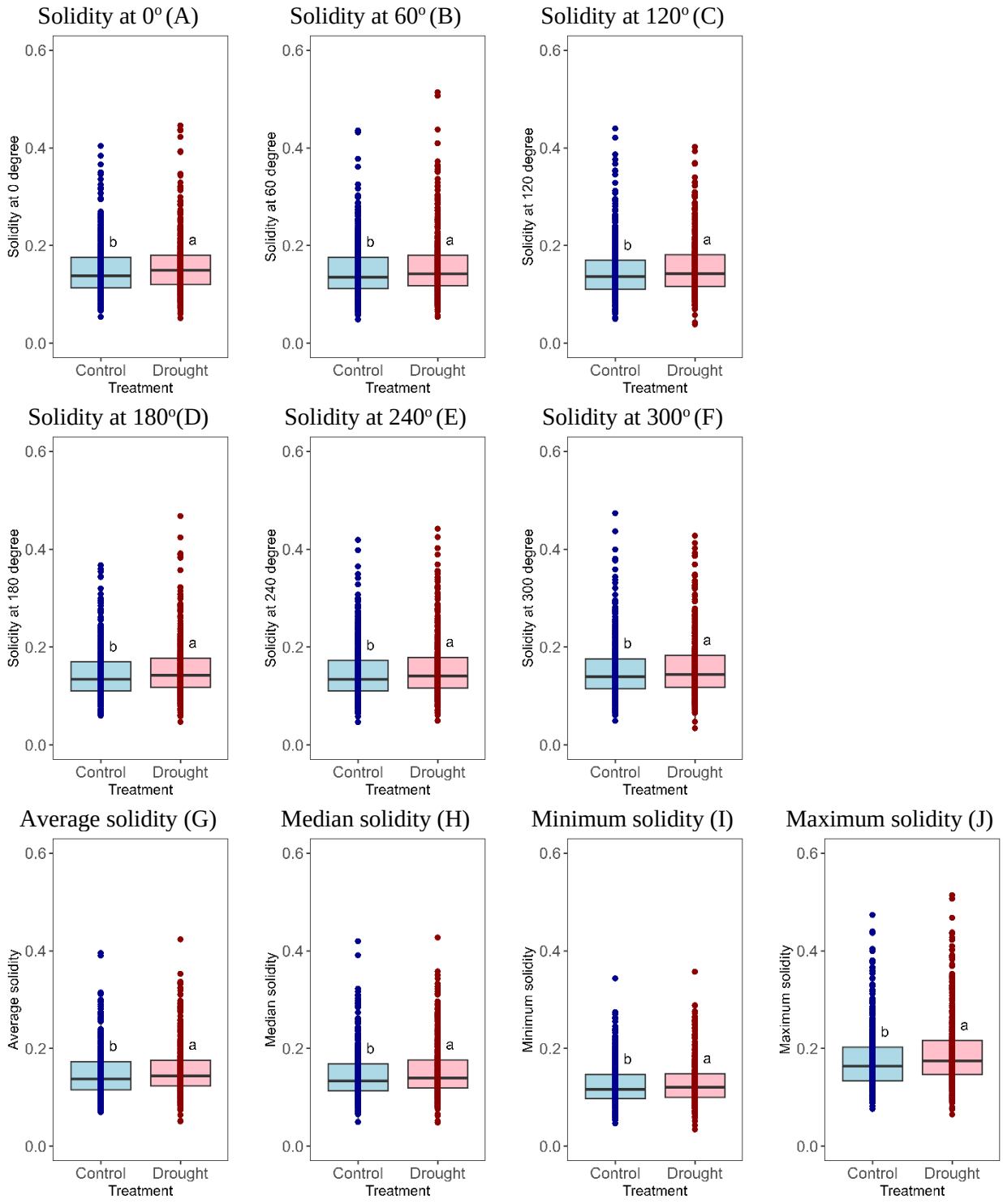


Figure 3A.2.4 Box plots of each side-view solidity of pea grown under controlled (blue) and drought-stressed (red) conditions at six different angles; 0° (A), 60 ° (B), 120 ° (C), 180 ° (D), 240 ° (E), 300 ° (F), average (G), median (H), minimum (I), and maximum (J). The letter above the box plots indicate significant differences when comparing between treatment ($p < 0.05$) using Fisher's protected LSD test.

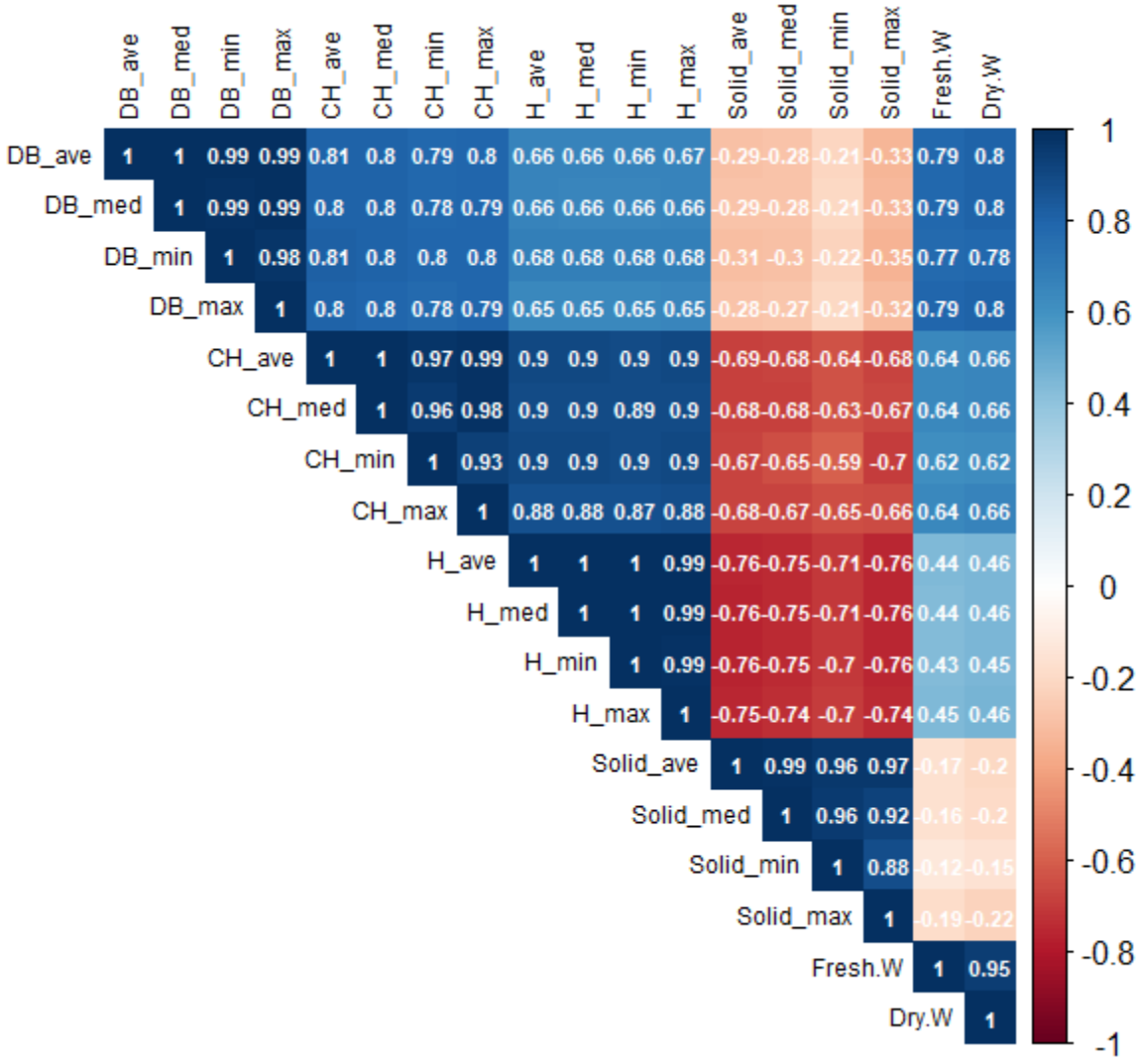


Figure 3A.2.5 Correlation matrix for among all side-view morphological traits of pea grown under control and drought conditions. Positive and negative correlations are illustrated in blue and red colors, respectively. Color shade is proportional to the correlation coefficients, with their value shown in the color intensity bar; Average (ave), median (med), minimum (min), and maximum (max) of digital biomass (DB), convex hull (CH), height (H), solidity (Solid), respectively. and total dry weight (Dry.W).

Appendix 3A.3: Criteria for evaluating the correlation coefficient.

Table 3A.3.1 Criteria for evaluating the correlation coefficient, where r is correlation coefficient.

Values	Interpretation
0.90 < r <= 1.00	Very high positive (negative) correlation
0.70 < r <= 0.90	High positive (negative) correlation
0.50 < r <= 0.70	Moderate positive (negative) correlation
0.30 < r <= 0.50	Low positive (negative) correlation
0.00 < r <= 0.30	Very low positive (negative) correlation

Appendix 3A.4: Loading scores for each principal component.

Table 3A.4.1 Trait loading scores of image-based phenotyping traits for each principal component.

Traits	Factor loadings	
	Dim1	Comp.2
DB_Top[1]	0.069	**0.422**
DB_Side	0.177	**0.232**
CH	0.254	**0.266**
Height	**0.275**	0.236
Solidity	**-0.264**	-0.167
RRed	**0.366**	-0.230
RGreen	**0.374**	-0.115
RBlue	**0.319**	-0.269
HUE	**-0.335**	0.082
SAT	0.200	**0.242**
VAL	**0.373**	-0.123
F_0	0.127	**0.361**
F_m	-0.003	**0.475**
F_v/F_m	**-0.266**	0.200
Standard deviation	1.490	0.708
Proportion of Variance	0.741	0.167
Cumulative proportion	0.741	0.908

[1] top-view digital biomass (DB_Top), side-view digital biomass (DB_Side), reflectance of red color (RRed), reflectance of green color (RGreen), reflectance of blue color (RBlue), hue, saturation, value, minimum fluorescence of dark-adapted leaves (F_0), maximum fluorescence of dark-adapted leaves (F_m), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m).

[2] Bold number indicated the maximum magnitude of loading score among two principal components.

Appendix 3A.5: Parameter selection for biomass estimation model with single type of traits

To select parameters for biomass estimation models using a single sensor, multicollinearity was considered among the parameters obtained from the same sensor. Two criteria were used: correlation coefficients and variance inflation factor (VIF). The correlation coefficients were examined first, and parameters with a correlation coefficient higher than 0.7 (Songara, 2022) were identified as having a strong relationship with others, indicating multicollinearity. In such cases, only one parameter was selected, while the others were excluded. The remaining parameters were then used to form multiple linear regression models, taking into account the VIF values. Parameters with a VIF greater than 2.5 showed multicollinearity properties and were excluded (Johnston et al., 2018). One parameter was selected based on its model performance, particularly its higher correlation of determination. This process was repeated until all VIF values for the parameters in the models were lower than 2.5. Finally, the set of parameters for the biomass estimation model formulation was determined.

Regarding RGB sideview traits, there were five parameters: the top-view pixel measured by the CropReporter camera (DB_Top), the average of side-view pixels imaged by RGB sideview cameras (DB_Side), convex hull (CH), height, and solidity. The correlation between solidity and height was lower than -0.70, showing high negative correlation, hence, solidity was eliminated since it also had lower correlation to fresh (FW) and dry weight (DW) than height (Figure 3.7). Meanwhile, height had a very high positive with convex hull area and had lower correlation coefficient with FW and DW compared with convex hull, then height was also excluded from the model formulation (Figure 3.7). Lastly, convex hull area had a high positive correlation with DB_Side and lower correlation coefficient with FW and DW than in DB_Side, then convex hull was not considered in the model. Therefore, there were only two traits using for model formulation from RGB sideview traits which were DB_Top and DB_Side, and there were no multicollinearity problems found after formulation (Table 3A.5.1).

In the case of spectral imaging (SI) traits, the reflectance of red (RRed), green (RGreen), blue (RBlue), and value (VAL) parameters showed a strong correlation among each other, with correlation coefficients higher than 0.7 (Figure 3.7). Among them, RGreen exhibited the strongest correlation with FW and DW compared to the others, leading to its selection as the primary parameter. The remaining three parameters, RGreen, HUE, and saturation (SAT), were assessed for VIF values. The results indicated that RGreen and HUE had VIF values higher than 2.5 (2.50 and 3.06, respectively), while the VIF value of SAT was 1.67 (Table 3A.4). Consequently, two multiple linear regression models were formulated: one comprising RGreen and SAT, and the other involving HUE and SAT. Neither model exhibited multicollinearity properties, as their VIF values were lower than 2.5 (Table 3A.4). To determine the most appropriate parameters, the correlation of determination (R^2) was considered. The results revealed that the model with HUE had a higher R^2 value compared to the one with RGreen (0.10 and 0.07, respectively). Therefore, SAT and HUE were included in the biomass estimation model based on SI traits.

Finally, for chlorophyll fluorescence (CF) traits, three parameters were considered: F_0, F_m, and F_v/F_m. The correlation coefficient between F_0 and F_m was found to be very strong, with a value of 0.85 (Figure 3.7). Consequently, F_m was selected as it displayed a higher correlation with FW and DW compared to F_0. Additionally, the VIF values indicated

no multicollinearity properties, with a value of 1.09. Hence, F_m and F_v/F_m were included in the biomass estimation model derived from CF traits.

Table 3A.5.1 Variance inflation factor (VIF) values for each variable used in biomass estimation models with single type of traits, its correlation of determination (R^2), and remark from each model.

Biomass estimation model with single type of traits	Variance inflation factor (VIF) values							R^2	Remarks
	DB_Top	DB_Side	RGreen	HUE	SAT	F_m	F_v/F_m		
RGB sideview traits									
2 variables	1.77	1.77	-	-	-	-	-	0.76	Selected[1]
Spectral imaging traits									
3 variables	-	-	2.50	3.06	1.67	-	-	0.11	Colinear
2 variables	-	-	1.36	-	1.36	-	-	0.07	
2 variables	-	-	-	1.67	1.67	-	-	0.10	Selected
Chlorophyll fluorescence traits									
2 variables	-	-	-	-	-	1.09	1.09	0.19	Selected

[1] Remarks indicated status of the considering model, 'Colinear' means the model with that set of parameters had multicollinearity problem, and 'Selected' means the model with that set of parameters was selected for biomass estimation model formulation; top-view digital biomass (DB_Top), side-view digital biomass (DB_Side), hue (HUE), saturation (SAT), maximum fluorescence of dark-adapted leaves (F_m), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m).

Appendix 3A.6: Parameter selection for biomass estimation model with multiple types of traits

Biomass estimation models were formulated using multiple types of image-based phenotypic traits in four scenarios: (i) a combination of RGB sideview, and spectral imaging (SI) traits, (ii) a combination of RGB sideview and chlorophyll fluorescence (CF) traits, (iii) a combination of SI and CF traits, and (iv) a combination of all traits. The initial models were formulated by combining parameters from the models with single types of phenotypic traits. Similar to the models with single types of phenotypic traits, the parameters used in each model were selected by considering the correlation among traits and variance inflation factor (VIF) values. One parameter was excluded from the model until there were no multicollinearity issues.

The model derived from the combination of RGB sideview, and SI traits initially included four variables: top-view digital biomass (DB_Top), side-view digital biomass (DB_Side), HUE, and saturation (SAT), as there were no multicollinearity issues. These four parameters were then included in the model formulation (Table 3A.6.1).

For the model with the combination of RGB sideview and CF traits, the model started with four variables: DB_Top, DB_Side, maximum fluorescence of dark-adapted leaves (F_m), and and the maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m). However, there was a multicollinearity issue due to a higher VIF value (>2.5) in DB_Top (Table 3A.6.1). Therefore, one of the two parameters with the highest VIF value was eliminated. Based on the correlation of determination (R^2), the model excluded F_m, which provided a higher R^2. Thus, three variables, excluding F_m, were selected for the model with the combination of RGB sideview and CF traits (Table 3A.6.1). Similarly, the combination of SI and CF traits model indicated multicollinearity with four variables (Table 3A.6.1). Excluding HUE from the model provided a higher R^2 compared to excluding SAT and resolved the multicollinearity problem. Therefore, HUE, F_m, and F_v/F_m were used to formulate the biomass estimation model in this case (Table 3A.6.1).

Lastly, the model with the combination of all traits included all six parameters, but multicollinearity was observed in four of them (Table 3.7). Even after eliminating one of these parameters, multicollinearity still persisted. Consequently, the model with five parameters, excluding SAT, was used as a base model for further parameter reduction, as it showed the highest R^2 among the models with five variables (Table 3A.6.1). Either DB_Top or F_m was eliminated from the model since these two parameters had the highest VIF values (Table 3A.6.1). The results indicated that removing F_m from the model provided a higher R^2. Therefore, the model with the combination of all traits consisted of four parameters: DB_Top, DB_Side, HUE, and F_v/F_m (Table 3A.6.1).

Table 3A.6.1 Variance inflation factor (VIF) values for variables used in biomass estimation models with multiple type of traits, its correlation of determination (R^2), and remark from each model.

Biomass estimation model with multiple types of traits	Variance inflation factor (VIF) values						R^2	Remarks
	DB_Top	DB_Side	HUE	SAT	F_m	F_v/F_m		
RGB sideview, and spectral imaging traits								
4 variables	2.11	2.01	1.93	2.03	-	-	0.755	Selected[1]
RGB sideview and chlorophyll fluorescence traits								
4 variables	3.47	1.97	-	-	2.22	1.10	0.777	Colinear
3 variables	-	1.12	-	-	1.21	1.10	0.720	
3 variables	1.90	1.83	-	-	-	1.07	0.774	Selected
Spectral imaging, and chlorophyll fluorescence traits								
4 variables	-	-	2.88	3.30	2.11	1.65	0.157	Colinear
3 variables	-	-	-	1.74	1.79	1.38	0.157	Selected
3 variables	-	-	1.52	-	1.11	1.65	0.154	
All traits								
6 variables	3.64	2.21	3.08	3.44	3.74	1.67	0.797	Colinear
5 variables	-	1.25	3.06	3.30	2.21	1.67	0.742	Colinear
5 variables	3.62	2.17	-	1.91	3.22	1.38	0.779	Colinear
5 variables	3.49	2.17	1.71	-	2.23	1.67	0.794	Colinear & Check
5 variables	2.15	2.01	2.65	2.06	-	1.62	0.792	Colinear
4 variables	-	1.25	1.70	-	1.21	1.67	0.742	
4 variables	1.90	2.00	1.70	-	-	1.60	0.792	Selected

[1] Remarks indicate the status of the considered models. "Colinear" means that the model with that set of parameters had a multicollinearity problem. "Selected" means that the model with that set of parameters was chosen for biomass estimation model formulation. "Colinear & Check" means that the model with that set of parameters had a multicollinearity problem, but it was the best model with a certain number of variables. In this case, the model was used to further check multicollinearity by reducing one variable: top-view digital biomass (DB_Top), side-view digital biomass (DB_Side), hue (HUE), saturation (SAT), maximum fluorescence of dark-adapted leaves (F_m), and maximum quantum yield of photosystem II (PSII) photochemistry (F_v/F_m).

Appendix 4A.1: Irrigation data in the experiment

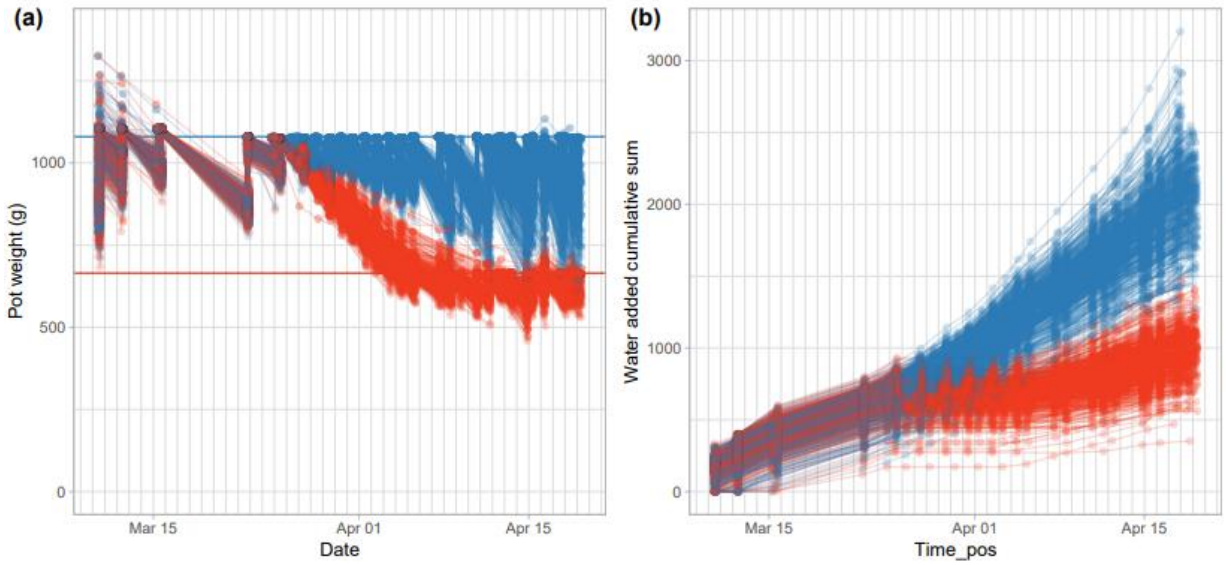


Figure 4A.1.1 The pot weight (A) and cumulative weight of water added (B) during the whole experiment; Blue and red horizontal line represented threshold of weight at 70% of field capacity (FC) and 30% FC for control and drought treatment, respectively.

Appendix 4A.2: Flowering pea plant during the experiment

Figure 4A.2.1 Pea plants with flowers and peas were observed just one week before the imaging measurement.